

Prévision des Rendements des Actifs Financiers en Utilisant les Régimes Économiques

Otman

June 12, 2024

Remerciements

Je tiens à exprimer ma profonde gratitude à toutes les personnes qui ont contribué à la réalisation de ce projet. Ce travail n'aurait pas été possible sans le soutien et l'encouragement de plusieurs individus précieux.

Tout d'abord, je souhaite remercier mon professeur encadreur, **Réné Garcia**, pour son encadrement exceptionnel, sa patience et ses conseils avisés. Votre expertise et votre dévouement ont été des piliers essentiels dans l'accomplissement de ce projet. Votre confiance en mes capacités m'a permis de repousser mes limites et d'atteindre mes objectifs.

Je tiens également à exprimer ma reconnaissance à mes professeurs, dont les enseignements ont largement contribué à enrichir mes connaissances et compétences.

Un remerciement spécial va à ma famille pour leur soutien tout au long de ce parcours. Parents, frères et sœurs pour leur patience, leurs encouragements et leur amour inconditionnel.

Je souhaite également remercier mes amis et collègues pour leur soutien moral, leurs conseils et leurs encouragements. Votre présence à mes côtés a été une grande source de motivation et de réconfort.

Enfin, je n'oublie pas toute personne qui, de près ou de loin, a contribué à la réalisation de ce projet. Que ce soit par des discussions constructives, des conseils pratiques ou simplement par leur présence bienveillante, chacune de vos contributions a été précieuse.

À toutes et à tous, merci du fond du cœur.

Otman

Contents

1	Revue de la Littérature	3
1.1	Prévision des Rendements des Actifs	3
1.2	Régimes économiques	4
1.3	Analyse en Composantes Principales (PCA)	4
1.4	Filtre de Hamilton	5
2	Présentation des données	7
2.1	Caractéristiques des Actifs Financiers	7
2.2	Données Macro-économiques	8
3	PCA et filtre de Hamilton: Donnée macro	9
3.1	Analyse en Composantes Principales (PCA)	9
3.1.1	Étapes de la PCA	9
3.1.2	Résultats de la PCA	10
3.2	Modèle de changement de Régimes	11
3.2.1	Représentation État-Espace	11
3.2.2	Filtre de Kalman	12
3.2.3	Changement régimes de Hamilton	13
4	Utilisation des Réseaux de Neurones pour la Prédiction des Rendements des Actifs	16
4.1	Introduction aux Réseaux de Neurones	16
4.2	Structure et Fonctionnement d'un Réseau de Neurones	16
4.2.1	Couche d'entrée	17
4.2.2	Couches Cachées	17
4.2.3	Couche de Sortie	17
4.3	Entraînement du Modèle	17
4.3.1	Propagation Avant (Forward Propagation)	18
4.3.2	Calcul de la Perte (Loss Calculation)	18

4.3.3	Rétro propagation (Backpropagation)	18
4.3.4	Itérations et Convergence	18
4.4	Implémentation	19
4.4.1	Préparation des Données	19
4.4.2	Construction du Modèle	19
4.4.3	Entraînement du Modèle	19
4.4.4	Évaluation et Prédiction	20
5	Résultats	21
5.1	Probabilités de Régimes Économiques	21
5.2	Performance du Modèle de Prédiction	23
6	Discussion	27
6.1	Implications des Résultats	27
6.2	Limitations de l'étude	27
6.3	Suggestions	27
7	Conclusion	28

Abstract

Ce projet se concentre sur l'analyse et la prévision des rendements des actifs financiers en combinant les caractéristiques des actifs avec les régimes économiques obtenus à partir des données macroéconomiques. L'idée de ce projet est inspirée des travaux de Chen et al. et Gu et al., qui ont démontré que l'information macroéconomique est indispensable dans la prévision des rendements. En conséquence, nous avons souhaité créer un modèle plus simple qui prend en compte cette information macroéconomique. La méthode proposée combine l'analyse en composantes principales (ACP) et le filtre de Hamilton pour identifier les probabilités des régimes économiques. Ces probabilités sont ensuite intégrées avec les caractéristiques des actifs financiers pour modéliser et prédire les rendements.

Introduction

La prédiction des rendements constitue un défi majeur pour les investisseurs et les analystes financiers, car elle leur permet de prendre des décisions éclairées sur la gestion de leur portefeuille et d'anticiper les mouvements du marché. Pour les investisseurs, il est crucial de pouvoir estimer avec précision les rendements futurs des actifs financiers afin de maximiser les gains potentiels tout en minimisant les risques associés.

Dans cette quête de prévision des rendements, une gamme de modèles théoriques ont vu le jour. Le CAPM est l'un des premiers modèles, il essaie d'évaluer le rendement d'un actif en fonction de son risque systématique, mesuré par bêta. Plus tard Fama et French proposent les modèles à plusieurs facteurs. Chacun de ces modèles cherchaient à expliquer les variations des rendements en fonction de différents facteurs économiques et financiers, mais tous ont des limites à prédire avec précision les mouvements du marché.

L'émergence du machine learning a ouvert de nouvelles possibilités dans le domaine de la prédiction des rendements. Ces modèles se basent sur des algorithmes qui exploitent la puissance de calcul des machines pour découvrir des schémas complexes dans les données. Les études ont pu montrer que des modèles sont plus précis, même si leur problème d'interprétabilité persiste. En effet, les recherches telles que celles de Gu et al. et Chen et al., ont pu montrer une amélioration significative par rapport aux méthodes traditionnelles de prévision des rendements.

Dans cette étude, nous proposons une approche pour estimer les rendements des actifs financiers en intégrant les probabilités de régimes économiques avec les caractéristiques des actifs dans un modèle de réseau de neurones. On utilise le filtre de Hamilton pour extraire les régimes économiques à partir des données historiques, puis on combine ces informations avec les caractéristiques des actifs pour obtenir une matrice enrichie de données. Cette méthodologie permet de capturer les variations des rendements des actifs dans différents régimes économiques.

D'abord, les données macroéconomiques sont classées en groupes et une analyse de composantes principales est effectuée pour chaque groupe. La première com-

posante principale (PC1) de chaque groupe, représentant la majorité de la variance, est extraite et utilisée pour capturer les dynamiques macro-économiques pertinentes.

Ensuite, les séries temporelles de PC1 sont passées à travers le filtre de Hamilton pour identifier les régimes économiques et calculer les probabilités associées à chaque régime.

Enfin, un modèle de régression (Réseau de neurones) est ensuite utilisé pour estimer les rendements des actifs en fonction de la matrice de données enrichies.

Ce projet vise à démontrer que l'intégration des régimes économiques, obtenus via une combinaison de PCA et du filtre de Hamilton, peut améliorer la compréhension et la prévision des rendements des actifs financiers. En utilisant les informations macroéconomiques dynamiques et en les combinant avec les caractéristiques spécifiques des actifs, il est anticipé que les modèles de prévision pourront capturer plus efficacement les dynamiques économiques complexes et fournir des estimations de rendement plus précises.

Chapter 1

Revue de la Littérature

1.1 Prédiction des Rendements des Actifs

L'estimation des rendements des actifs financiers a toujours été un sujet central en finance. Les modèles traditionnels, tels que le modèle d'évaluation des actifs financiers (CAPM) , le modèle à trois facteurs (1993) et le modèle à cinq facteurs (2014) de Fama et French , se concentrent principalement sur les caractéristiques spécifiques des actifs et les facteurs de marché (Common risk factors in the returns on stocks and bonds, A five-factor asset pricing model).

Avec l'avènement des techniques de machine learning, de nouvelles approches ont émergé pour améliorer la prédiction des rendements des actifs. Gu et al. (2020), dans leur étude "Empirical Asset Pricing via Machine Learning", démontrent des gains économiques significatifs pour les investisseurs utilisant des prévisions basées sur le machine learning, doublant parfois les performances des stratégies de régression traditionnelles.

Chen et al. (2023), dans leur étude "Deep Learning in Asset Pricing", utilisent des réseaux de neurones profonds et un pour estimer un modèle de tarification des actifs pour les rendements des actions individuelles. Leur approche utilise une vaste quantité d'informations de conditionnement et la condition fondamentale de non-arbitrage comme fonction de critère pour construire les actifs de test les plus informatifs avec une approche "adversariale" (Réseau génératif antagoniste).

Ces avancées dans l'utilisation de l'apprentissage machine pour la prédiction des rendements des actifs montrent que les modèles traditionnels peuvent être considérablement améliorés en incorporant des techniques modernes d'apprentissage automatique, ouvrant ainsi de nouvelles perspectives pour la recherche en finance.

1.2 Régimes économiques

Les régimes économiques jouent un rôle crucial dans l'estimation des rendements des actifs financiers, car ils permettent de capturer les changements structurels dans l'économie qui peuvent influencer les rendements. Les régimes économiques peuvent être modélisés à l'aide de processus de changement de régime, tels que les modèles de Markov switching, qui permettent de passer entre différents états de l'économie de manière probabiliste.

L'utilisation des régimes économiques dans l'analyse financière est bien illustrée par les travaux de René Garcia et Pierre Perron (1996) dans leur étude "An Analysis of Real Interest Rates Under Regime Shifts". Ils montrent comment les taux d'intérêt réels peuvent être influencés par des changements de régime, ce qui permet une meilleure compréhension et prévision de leurs dynamiques. De même, l'étude "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process" explore comment les processus de Markov switching peuvent capturer les distributions conditionnelles des taux d'intérêt, offrant ainsi une modélisation plus précise des taux en fonction des différents régimes économiques.

Les implications des régimes économiques sur les rendements des actifs ont également été étudiées dans diverses recherches. Par exemple, Ang et Bekaert (2002) dans leur article "International Asset Allocation with Regime Shifts" examinent comment les régimes économiques internationaux peuvent affecter l'allocation d'actifs et les rendements. De même, Hamilton (1989), dans son étude pionnière "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", démontre l'importance des régimes économiques pour comprendre les cycles économiques et leurs impacts sur les rendements des actifs.

Ces études montrent que les régimes économiques, et en particulier les modèles de changement de régime de Markov, sont des outils puissants pour améliorer la prévision et l'estimation des rendements des actifs financiers en tenant compte des changements structurels et des dynamiques économiques complexes.

1.3 Analyse en Composantes Principales (PCA)

L'Analyse en Composantes Principales (PCA) est une technique statistique utilisée pour réduire la dimensionnalité des jeux de données tout en préservant autant que possible la variance des données initiales. En finance, la PCA est particulièrement utile pour décomposer les mouvements des marchés financiers en un ensemble de

facteurs orthogonaux, permettant ainsi de simplifier l'analyse des risques et des rendements.

La PCA trouve de nombreuses applications en finance, notamment pour la gestion de portefeuille, l'analyse de risque, et la détection de structures cachées dans les données de marché. En réduisant la dimensionnalité des données, la PCA aide à identifier les principaux moteurs des variations des prix des actifs et à construire des modèles plus robustes et interprétables.

Parmi les travaux notables utilisant la PCA en finance, Connor et Korajczyk (1988) dans leur étude "Risk and Return in an Equilibrium APT: Application of a New Test Methodology" appliquent la PCA pour tester le modèle de tarification par arbitrage (APT). Ils montrent que la PCA peut identifier les facteurs de risque systématiques qui influencent les rendements des actifs. De même, le travail de Litterman et Scheinkman (1991) dans "Common Factors Affecting Bond Returns" utilise la PCA pour analyser les facteurs communs influençant les rendements obligataires, démontrant son utilité pour décomposer les risques liés aux taux d'intérêt.

En résumé, la PCA est une méthode puissante pour analyser et simplifier des jeux de données complexes en finance, aidant ainsi à une meilleure compréhension des facteurs sous-jacents qui influencent les rendements des actifs.

1.4 Filtre de Hamilton

Le Filtre de Hamilton, développé par James D. Hamilton, est une technique statistique utilisée pour extraire les composantes non observées d'une série temporelle, notamment les cycles économiques et les tendances. Introduit dans son étude pionnière "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle" (1989), le Filtre de Hamilton permet de modéliser des séries temporelles en tenant compte des changements de régime, souvent représentés par un processus de Markov.

L'utilité du Filtre de Hamilton en finance réside dans sa capacité à capturer les dynamiques non linéaires et les changements structurels dans les données économiques et financières. Cette technique est particulièrement précieuse pour l'analyse des cycles économiques, la détection des points de retournement et la prévision des tendances économiques.

Dans "Time Series Analysis" (1994), Hamilton approfondit l'application du filtre et de la modélisation des changements de régime à divers problèmes économiques, offrant un cadre robuste pour analyser les séries temporelles économiques et financières.

Le Filtre de Hamilton est couramment utilisé pour :

Détection des Régimes : Identifier les périodes de croissance et de récession dans les données macro-économiques. Prévion Économique : Améliorer les prévisions économiques en tenant compte des changements de régime. Analyse des Cycles Économiques : Comprendre les cycles économiques et leur impact sur les rendements des actifs financiers.

Chapter 2

Présentation des données

Pour cette étude, nous avons utilisé les données que Chen et al. (2023), disponible sur le drive de Chen ([lien ici](#)) qui ont été collectées et traitées par eux dans le cadre de leur article "Deep Learning in Asset Pricing".

2.1 Caractéristiques des Actifs Financiers

Les données comprennent les rendements mensuels des actions américaines, collectées à partir de la base de données CRSP (Center for Research in Security Prices). La période d'échantillonnage s'étend de janvier 1967 à décembre 2016, couvrant ainsi 50 ans.

Chen et al. ont divisé l'ensemble des données en trois échantillons : un échantillon d'entraînement de 20 ans (1967-1986), un échantillon de validation de 5 ans (1987-1991), et un échantillon de test hors échantillon de 25 ans (1992-2016).

En plus des rendements des actions, Chen et al. ont collecté 46 caractéristiques spécifiques aux entreprises, soit listées dans la bibliothèque de données de Kenneth French, soit utilisées par Freyberger et al. (2020). Ces variables sont construites à partir des données comptables de la base de données CRSP/Compustat ou des rendements passés de CRSP. Les variables mises à jour annuellement le sont à la fin de chaque mois de juin, conformément à la convention de Fama-French, tandis que les variables mises à jour mensuellement le sont à la fin de chaque mois pour une utilisation le mois suivant.

Le nombre total d'actions disponibles dans CRSP est d'environ 31 000. Cependant, en suivant les méthodologies de Kelly et al. (2019) et Freyberger et al. (2020),

Chen et al. se sont limités aux rendements des actions pour lesquelles toutes les informations caractéristiques de l'entreprise étaient disponibles à un moment donné, réduisant l'ensemble à environ 10 000 actions.

2.2 Données Macro-économiques

Enfin, Chen et al. ont également collecté 178 séries temporelles macro-économiques provenant de trois sources : 124 prédicteurs macro-économiques de la base de données FRED-MD (détaillés par McCracken et Ng, 2016), les séries temporelles médianes transversales pour chacune des 46 caractéristiques des entreprises, et huit prédicteurs macro-économiques supplémentaires suggérés par Welch et Goyal (2007).

En outre, pour faciliter l'interprétation de l'impacte des différentes données macro-économiques sur l'estimation des rendements des actifs, on a classé les 178 séries temporelles en neuf catégories distinctes. Ces catégories, basées sur la classification de la base de données FRED-MD, sont les suivantes : Production et revenus, Travail et marché, Logement, Consommation, monnaie et crédit, taux d'intérêt et de change, Prix, Marchés boursiers et autres variables . Cette catégorisation permet de regrouper les séries temporelles similaires et de mieux comprendre les relations entre les différentes variables macroéconomiques. La description des différentes variables et des catégories est accessible en ligne dans le travail de McCracken et Ng (2020) sur FRED-MD, lien ([ici](#)),

Chapter 3

PCA et filtre de Hamilton: Donnée macro

Cette section détaille la méthodologie employée pour classer les données macro, effectuer l'analyse en composantes principales (PCA) et appliquer le filtre de Hamilton pour obtenir les probabilité des régimes.

3.1 Analyse en Composantes Principales (PCA)

L'Analyse en Composantes Principales (PCA) est une technique de réduction de dimensionnalité largement utilisée en statistique et en apprentissage automatique. Elle permet de transformer un ensemble de variables corrélées en un ensemble de variables non corrélées appelées composantes principales, tout en conservant autant que possible de la variance présente dans les données d'origine.

La PCA vise à identifier les directions dans lesquelles les données varient le plus. Ces directions sont représentées par des vecteurs appelés composantes principales. L'objectif est de réduire la complexité des données tout en conservant leur structure essentielle.

3.1.1 Étapes de la PCA

1. **Standardisation des données** : Pour garantir que chaque variable contribue équitablement à l'analyse, il est courant de standardiser les données, en les

centrant (soustraction de la moyenne) et en les réduisant (division par l'écart-type).

2. **Calcul de la matrice de covariance** : La matrice de covariance, notée Σ , mesure les relations de covariance entre les différentes variables. Pour un ensemble de données X avec n observations et p variables, la matrice de covariance est calculée comme suit :

$$\Sigma = \frac{1}{n-1} X^T X$$

3. **Calcul des vecteurs propres et des valeurs propres** : Les vecteurs propres (\mathbf{v}) et les valeurs propres (λ) de la matrice de covariance sont calculés pour déterminer les directions principales et l'importance de ces directions respectivement. Ces vecteurs propres forment les composantes principales et les valeurs propres indiquent la quantité de variance expliquée par chaque composante. Mathématiquement, cela se traduit par :

$$\Sigma \mathbf{v} = \lambda \mathbf{v}$$

4. **Projection des données** : Les données d'origine sont projetées dans le nouvel espace formé par les composantes principales, transformant ainsi les données initiales en un ensemble de nouvelles variables non corrélées :

$$X' = X \mathbf{V}$$

où \mathbf{V} est la matrice des vecteurs propres.

3.1.2 Résultats de la PCA

- **Variance expliquée** : Les valeurs propres (λ) de la matrice de covariance représentent la quantité de variance expliquée par chaque composante principale. La somme des valeurs propres est égale à la variance totale des données d'origine, et le pourcentage de variance expliquée par chaque composante principale est donné par :

$$\text{Variance expliquée par } \mathbf{v}_i = \frac{\lambda_i}{\sum_{j=1}^p \lambda_j}$$

- **Choix des composantes principales** : Les composantes principales sont ordonnées par la quantité de variance qu'elles expliquent. En général, seules les premières composantes principales (celles avec les plus grandes valeurs propres) sont conservées, car elles capturent la majorité de la variance dans les données.

- **Dé-corrélation des variables** : Les composantes principales sont orthogonales entre elles, ce qui signifie qu'elles sont non corrélées. Cela simplifie l'analyse et permet d'éviter la redondance d'information.

Dans le contexte de ce projet , on regroupe les données macro-économiques en plusieurs groupes, puis on effectue une PCA sur chacun de ces groupes. Cela permet de :

- **Réduire la dimensionnalité** : En résumant chaque groupe de variables macro-économiques par un petit nombre de composantes principales on passe de 178 variables 8 groupes, cela nous facilite l'utilisation du filtre de Hamilton sans perdre d'information cruciale, autrement le bruit de ces données rendra le filtre inutilisable (dans le filtre il y a une partie d'optimisation ou on doit minimiser une fonction objective, avec 178 variables le nombre de paramètres de la fonction explose et trouver un optimum devient presque impossible)
- **Éliminer les redondances** : Les composantes principales étant non corrélées, elles éliminent la redondance présente dans les données initiales, ce qui peut améliorer la performance des algorithmes de prédiction.
- **Améliorer l'interprétabilité** : Avec un nombre limité de groupe de données macro l'interprétation devient plus facile que les 178 variables d'origine, facilitant ainsi l'analyse des résultats.

En résumé, on groupe nos données en neuf catégories et on ajoute un dixième groupe qui contient toutes les données et on applique la PCA sur chaque groupe, on garde que la première composante de chaque groupe. On obtient ainsi les séries chronologique de chaque groupe de données macro économiques.

3.2 Modèle de changement de Régimes

3.2.1 Représentation État-Espace

Un système linéaire peut être représenté en forme état-espace avec les équations suivantes :

Équation d'observation

$$\mathbf{y}_t = \mathbf{A}'\mathbf{x}_t + \mathbf{H}'\xi_t + \omega_t \quad (3.1)$$

avec $\mathbf{E}[\omega_t\omega_t] = \mathbf{R}$ de dimension $n \times n$

où :

- \mathbf{y}_t est le vecteur d'observation à l'instant t (dimension $n \times 1$).
- \mathbf{H}' est la matrice de coefficient appelée matrice d'émission (dimension $n \times r$).
- ξ est le vecteur d'états du système (dimension $r \times 1$).
- \mathbf{x}_t vecteur de variables exogènes prédéterminées (dimension $k \times 1$).
- \mathbf{A}' est la matrice de coefficients pour les variables exogènes (dimension $n \times k$).
- ω_t un vecteur d'erreur de mesure ou de bruit (dimension $n \times 1$).

Équation d'état

$$\xi_{t+1} = \mathbf{F}\xi_t + \mathbf{v}_{t+1} = \quad (3.2)$$

avec $\mathbf{E}[\mathbf{v}_t\mathbf{v}_t] = \mathbf{Q}$ de dimension $r \times r$

où :

- ξ_t est le vecteur d'état à l'instant t (dimension $n \times 1$).
- \mathbf{F} est la matrice de transition d'état (dimension $n \times n$).
- \mathbf{v}_t est le bruit de processus (dimension $n \times 1$).

3.2.2 Filtre de Kalman

Le filtre de Kalman est une méthode récursive pour estimer l'état d'un processus dynamique en minimisant l'erreur quadratique moyenne. Il fonctionne en deux étapes principales : la prédiction et la mise à jour.

Étape de mise à jour

$$\begin{aligned}\mathbf{K}_t &= \mathbf{P}_{t|t-1} \mathbf{H} (\mathbf{H}' \mathbf{P}_{t|t-1} \mathbf{H} + \mathbf{R})^{-1} \\ \hat{\xi}_{t|t} &= \hat{\xi}_{t|t-1} + \mathbf{K}_t (\mathbf{y}_t - \mathbf{A}' \mathbf{x}_t - \mathbf{H}' \hat{\xi}_{t|t-1}) \\ \mathbf{P}_{t|t} &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}') \mathbf{P}_{t|t-1}\end{aligned}$$

où :

- \mathbf{K}_t est le gain de Kalman.
- $\hat{\xi}_{t|t}$ est l'estimation a posteriori de l'état à l'instant t .
- $\mathbf{P}_{t|t}$ est la matrice de covariance de la distribution conditionnelle de ξ_t (covariance a posteriori de l'erreur d'estimation).

Étape de prédiction

$$\hat{\xi}_{t+1|t} = \mathbf{F} \hat{\xi}_{t|t} \tag{3.3}$$

$$\mathbf{P}_{t+1|t} = \mathbf{F} \mathbf{P}_{t|t} \mathbf{F}' + \mathbf{Q} \tag{3.4}$$

où :

- $\hat{\xi}_{t+1|t}$ est l'estimation a priori de l'état à l'instant $t + 1$.
- $\mathbf{P}_{t+1|t}$ est la covariance a priori de l'erreur d'estimation.

3.2.3 Changement régimes de Hamilton

Hamilton a étendu le filtre de Kalman pour modéliser les changements de régimes économiques en utilisant des chaînes de Markov à états cachés.

Par exemple un Changement de régime AR(1) à la forme:

$$y_t = \mu_{st} + \beta_{st} y_{t-1} + \omega_{st}$$

ou $s_t = 1, 2, 3, \dots, N$ sont les régimes possibles.

Dans le filtre de Kalman les états ξ_t étaient des variables continues, quand ces états représentent des régimes économiques distincts elles sont des variables discrètes et

leur distribution conditionnelle se définie par:

$$P(s_t = i | \zeta_{t-1})$$

avec $i = 1, 2, \dots$ et F_t représente toute l'information (variables et paramètres) connue jusqu'au temps (t).

Dans le cas d'une chaîne de Markov

$$P(s_t = i | \zeta_{t-1}) = P(s_t = i | s_{t-1})$$

ce qui donne

$$P(s_t = i | s_t = j) = P_{ij}$$

Ainsi dans le cas de deux régimes la matrice des probabilités de transitions est:

$$P(s_t = j | s_{t-1} = i) = P_{ij} = \begin{bmatrix} p_{00} & p_{10} \\ p_{01} & p_{11} \end{bmatrix}$$

Hamilton propose la même technique que Kalman : prédiction puis mise à jour pour trouver:

$$\xi_{i,t} = P(s_t = i | \zeta_t)$$

ou $\zeta_t = y_t, y_{t-1}, \dots$

en se basant sur l'assomption que $f(y_t | s_t = i, F_{t-1}) \sim N(\mu_{y_t}, \sigma)$

L'algorithme proposé par Hamilton est comme suite:

Étape1: Poser

$$\xi_{i,t-1} = P(s_{t-1} = i | \zeta_{t-1})$$

Étape2: Calculer la distribution de densité conditionnelle conjointe de y_t et de s_t :

$$f(y_t, s_t = i | \zeta_{t-1}) = f(y_t | s_t = i, \zeta_{t-1}) \cdot P(s_t = i | \zeta_{t-1})$$

ou:

$$f(y_t | s_t = j, \zeta_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left[-\frac{(y_t - \mu_j - \beta_j y_{t-1})^2}{2\sigma_j^2} \right]$$

Étape3: On utilise la densité conditionnelle et les probabilité à priori pour obtenir la vraisemblance de l'observation :

$$f(y_t | \zeta_{t-1}) = \sum_{j=0}^1 f(y_t, s_t = j | \zeta_{t-1}) = \sum_{j=0}^1 f(y_t | s_t = j, \zeta_{t-1}) P(s_t = j | \zeta_{t-1})$$

$$f(y_t | \zeta_{t-1}) = \sum_{i=0}^1 \sum_{j=0}^1 f(y_t | s_t = i, \zeta_{t-1}) P_{ij} P(s_{t-1} = i | \zeta_{t-1}) = \sum_{i=0}^1 \sum_{j=0}^1 \xi_{i,t-1} p_{ij} f(y_t | s_{t-1} = i, \zeta_{t-1})$$

En détails:

$$f(y_t | \zeta_{t-1}) = \xi_{0,t-1} p_{00} \eta_{0t} + \xi_{0,t-1} p_{01} f(y_t | s_t = i, \zeta_{t-1}) + \xi_{1,t-1} p_{10} \eta_{0t} + \xi_{1,t-1} p_{11} f(y_t | s_t = i, \zeta_{t-1})$$

Étape4: On utilise la vraisemblance pour mettre à jour le nouveau état:

$$\xi_{jt} = P(s_t = j | \zeta_t) = \frac{\sum_{i=0}^1 \xi_{i,t-1} p_{ij} \eta_{jt}}{f(y_t | \zeta_{t-1})}$$

Finalement répéter de : $t = 1$ à T pour obtenir les probabilités à toutes les périodes.

Optimisation: :

Le log vraisemblance des données peut être calculer comme suit:

$$\log(f(y_T, y_{T-1}, \dots, y_1 | y_0)) = \sum_{t=1}^T \log(f(y_t | \zeta_{t-1}))$$

L'algorithme peut être commencé avec des probabilité non conditionnelle.

A partir de paramètres initiaux on utilise une optimisation numérique pour obtenir les paramètres optimaux $(\mu_0, \mu_1, \beta_0, \beta_1, \sigma_0, \sigma_1, p, q)$.

A noté qu'on peut s'en passer facilement du β pour un modèle de type:

$$y_t = \mu_{s_t} + \sigma_{s_t}$$

sans auto régression. Le code permet d'implémenter les deux cas.

Chapter 4

Utilisation des Réseaux de Neurones pour la Prédiction des Rendements des Actifs

4.1 Introduction aux Réseaux de Neurones

En apprentissage machine les réseaux de neurones artificiels (ANN - Artificial Neural Networks) sont des modèles conçus pour reconnaître des motifs complexes dans les données et les utiliser pour effectuer des tâches tel que la prédiction.

Ils sont composés de couches de neurones inter connectés qui traitent les informations de manière non linéaire, permettant ainsi de modéliser des relations complexes entre les variables d'entrée et de sortie.

4.2 Structure et Fonctionnement d'un Réseau de Neurones

Un réseau de neurones peut être décrit comme une succession de fonctions appliquées aux données d'entrée pour produire une sortie. Le but est de trouver une fonction complexe f telle que $y = f(X)$, où X représente les données d'entrée (caractéristiques des actifs et états des régimes macro par exemple) et y représente la sortie (rendement prédit dans cet exemple).

4.2.1 Couche d'entrée

- **Description** : La couche d'entrée reçoit les données initiales. Chaque neurone de cette couche représente une caractéristique spécifique de l'ensemble de données.
- **Formule** : $\mathbf{a}^{(0)} = \mathbf{x}$, où \mathbf{x} est le vecteur des caractéristiques d'entrée, la couche d'entrée ne contient ni poids ni biais car elle sert seulement à recevoir les entrées pour le modèle.

4.2.2 Couches Cachées

- **Description** : Les couches cachées sont composées de noeuds qui effectuent des transformations non linéaires sur les données. Chaque couche cachée l reçoit les activations de la couche précédente $l - 1$, les transforme, et produit des activations pour la couche suivante.
- **Formule** : $\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}$, où $\mathbf{W}^{(l)}$ est la matrice de poids, $\mathbf{b}^{(l)}$ est le vecteur de biais, et $\mathbf{z}^{(l)}$ est le vecteur de valeurs linéaires.
- **Fonction d'activation** : $\mathbf{a}^{(l)} = g(\mathbf{z}^{(l)})$, où g est une fonction d'activation non linéaire (par exemple, ReLU, Sigmoid).

4.2.3 Couche de Sortie

- **Description** : La couche de sortie produit le résultat final du réseau. Pour la prédiction des rendements des actifs, cette couche génère une valeur continue représentant le rendement prédit.
- **Formule** : $\mathbf{y}_{\text{prédit}} = \mathbf{a}^{(L)}$, où L est le nombre total de couches.

4.3 Entraînement du Modèle

L'objectif de l'entraînement est de trouver les poids \mathbf{W} et les biais \mathbf{b} qui minimisent l'erreur entre les prédictions du modèle et les valeurs réelles. Ce processus se déroule en plusieurs étapes :

4.3.1 Propagation Avant (Forward Propagation)

Les données d'entrée traversent le réseau couche par couche. Chaque neurone applique la fonction $\mathbf{z}^{(l)} = \mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)}$ suivie de l'activation $\mathbf{a}^{(l)} = g(\mathbf{z}^{(l)})$. Cela permet de calculer la sortie prédite $\mathbf{y}_{\text{prédit}}$.

4.3.2 Calcul de la Perte (Loss Calculation)

La fonction de perte $L(\mathbf{y}_{\text{réel}}, \mathbf{y}_{\text{prédit}})$ mesure la différence entre les prédictions et les valeurs réelles. Pour un problème de régression, l'erreur quadratique moyenne (Mean Squared Error, MSE) est couramment utilisée :

$$L = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

4.3.3 Rétro propagation (Backpropagation)

L'erreur calculée est propagée en arrière à travers le réseau pour mettre à jour les poids et les biais. Les gradients de la fonction de perte par rapport aux poids et aux biais sont calculés en utilisant la règle de la chaîne (A noté que les graphes de calculs permettent d'exécuter la différentiation automatique d'une manière rapide, les dérivées explicites des fonctions complexes ne sont jamais calculées). Les poids et les biais sont mis à jour en utilisant un algorithme d'optimisation de la descente de gradient:

$$\mathbf{W}^{(l)} \leftarrow \mathbf{W}^{(l)} - \eta \frac{\partial L}{\partial \mathbf{W}^{(l)}}$$

et

$$\mathbf{b}^{(l)} \leftarrow \mathbf{b}^{(l)} - \eta \frac{\partial L}{\partial \mathbf{b}^{(l)}},$$

où η est le taux d'apprentissage.

Dans ce projet on utilise l'algorithme d'optimisation WAdam.

4.3.4 Itérations et Convergence

Le processus de propagation avant et de rétro propagation est répété pour de nombreuses itérations (époques) jusqu'à ce que l'erreur soit minimisée et que le modèle converge vers une solution optimale.

4.4 Implémentation

4.4.1 Préparation des Données

Après application du filtre de Hamilton pour déterminer les probabilités de régime macro-économiques, celles ci sont combinées avec les caractéristiques des actifs pour données une matrice de caractéristiques augmentée.

En effet pour prendre en compte toutes les interactions linéaire et non linéaire entre les données on construit une matrice de caractéristiques enrichies et d'implémenter le model suivant comme Gu. et al. (2020):

$$E[r_{i,t+1}] = f(x_{it} \otimes c_t) = f(z_{it})$$

ou c_t est constitué du vecteur de probabilité augmenté d'un vecteur de uns $c_t = [I, p_t]$

4.4.2 Construction du Modèle

- Les caractéristiques augmentées servent de données d'entrée au modèle. Afin de prendre en compte la dépendance temporelle entre les données on fait une prédiction cross sectionnelle, c'est à dire que chaque entrée comprend toutes les caractéristiques qui correspond à une période spécifique et chaque erreur calculée correspond à une période.
- La fonction activation utilisée est Relu.
- : Pour évaluer l'erreur on utilise la fonction objective qui prend en compte la dépendance temporelle entre les données:

$$L(\theta) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T (r_{it+1} - f(z_{it}; \theta))$$

4.4.3 Entraînement du Modèle

Une partie des données est utilisée pour entraîner le modèle afin de trouver les paramètres optimaux (poids et biais). La rétro propagation est utilisée pour ajuster les poids en fonction de l'erreur d'entraînement.

Malgré leur puissance les modèles de réseaux de neurones peuvent souffrir de problèmes de sur entraînement (surtout si les données sont insuffisante), et leur convergence dépend aussi des points initiaux, la validation est utilisée pour trouver les

hyper paramètres optimaux dans ce cas.

Pour trouver les hyper paramètres optimaux on utilise l'ensemble de donnée de validation et on essaie un ensemble d'hyper paramètres: 2 pas d'apprentissage, 2 ensembles de paramètres initiaux et 02 'droping rates', on choisit les hyper paramètres qui donne l'erreur minimale sur parmi cet ensemble.

4.4.4 Évaluation et Prédiction

Le modèle est validé en utilisant un ensemble de données distinct pour vérifier sa performance. Les métriques telles que l'erreur quadratique moyenne (MSE) ou le coefficient de détermination (R^2) et le Ratio de Sharpe sont utilisées pour évaluer la qualité des prédictions. Une fois le modèle entraîné et validé, il est utilisé pour prédire les rendements des actifs sur de nouvelles données.

Chapter 5

Résultats

Cette section présente les résultats de l'analyse, y compris les probabilités de régimes économiques obtenues, et les performances du modèle.

5.1 Probabilités de Régimes Économiques

Dans la figure ci après on voit les probabilités des états du régimes obtenues en utilisant la composante principales de différents groupes de données y compris les cas ou on utilise tous les groupes et ou on fait le PCA directement sur toutes les données sans groupes.

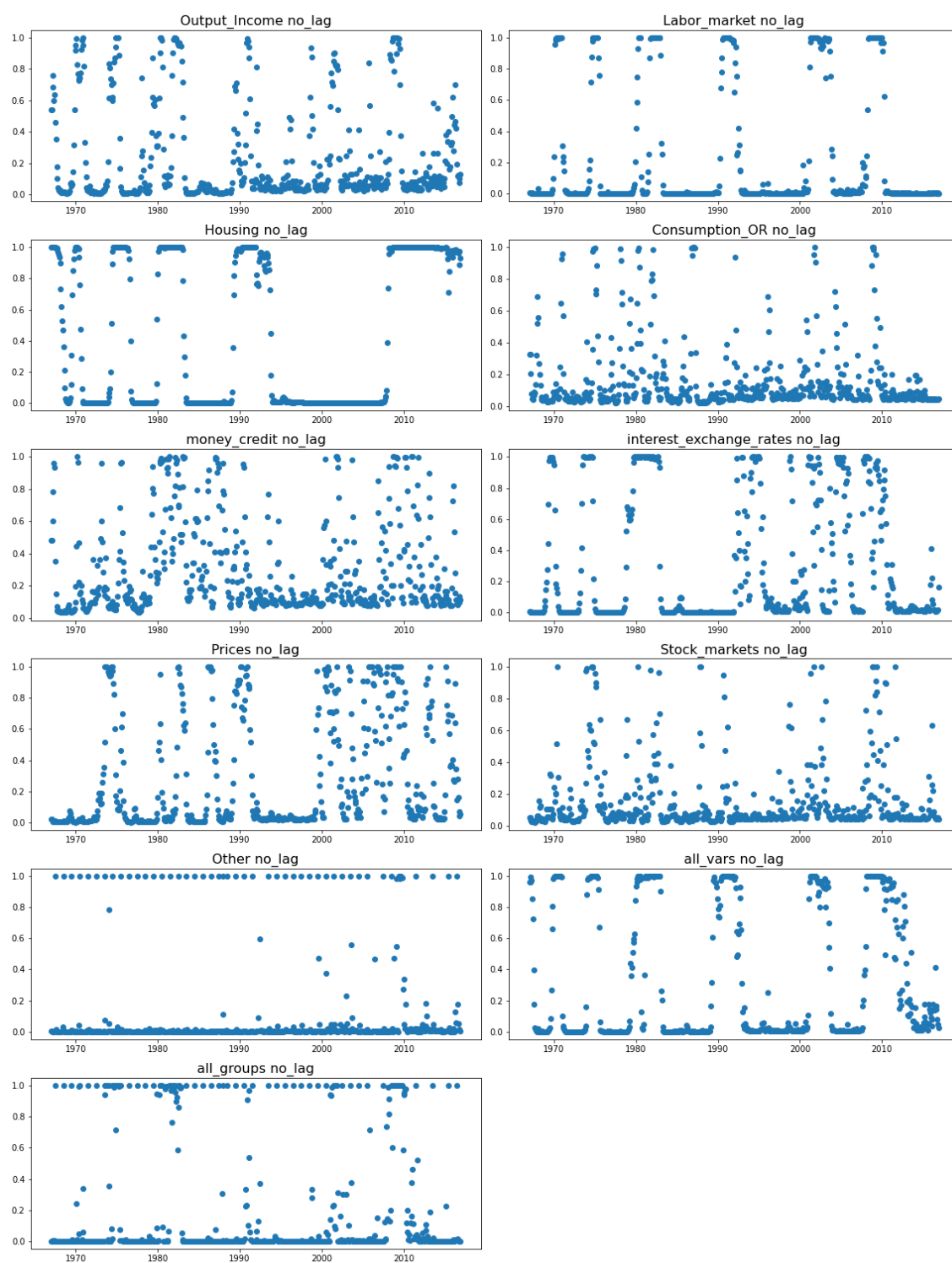


Figure 5.1: États du régime

5.2 Performance du Modèle de Prédiction

Sharp Ratio OOS L_S portf	
macro groupe	sharp
all vars	0.71
all groups	0.663
money credit	0.581
Labor market	0.575
Output Income	0.614
Housing	0.612
Consumption OR	0.588
interest exchange rates	0.594
Prices	0.573
Stock markets	0.585
Other	0.61
without macro	0.573

Figure 5.2: Sharp Ratios of Out of Simple Long short portfolios

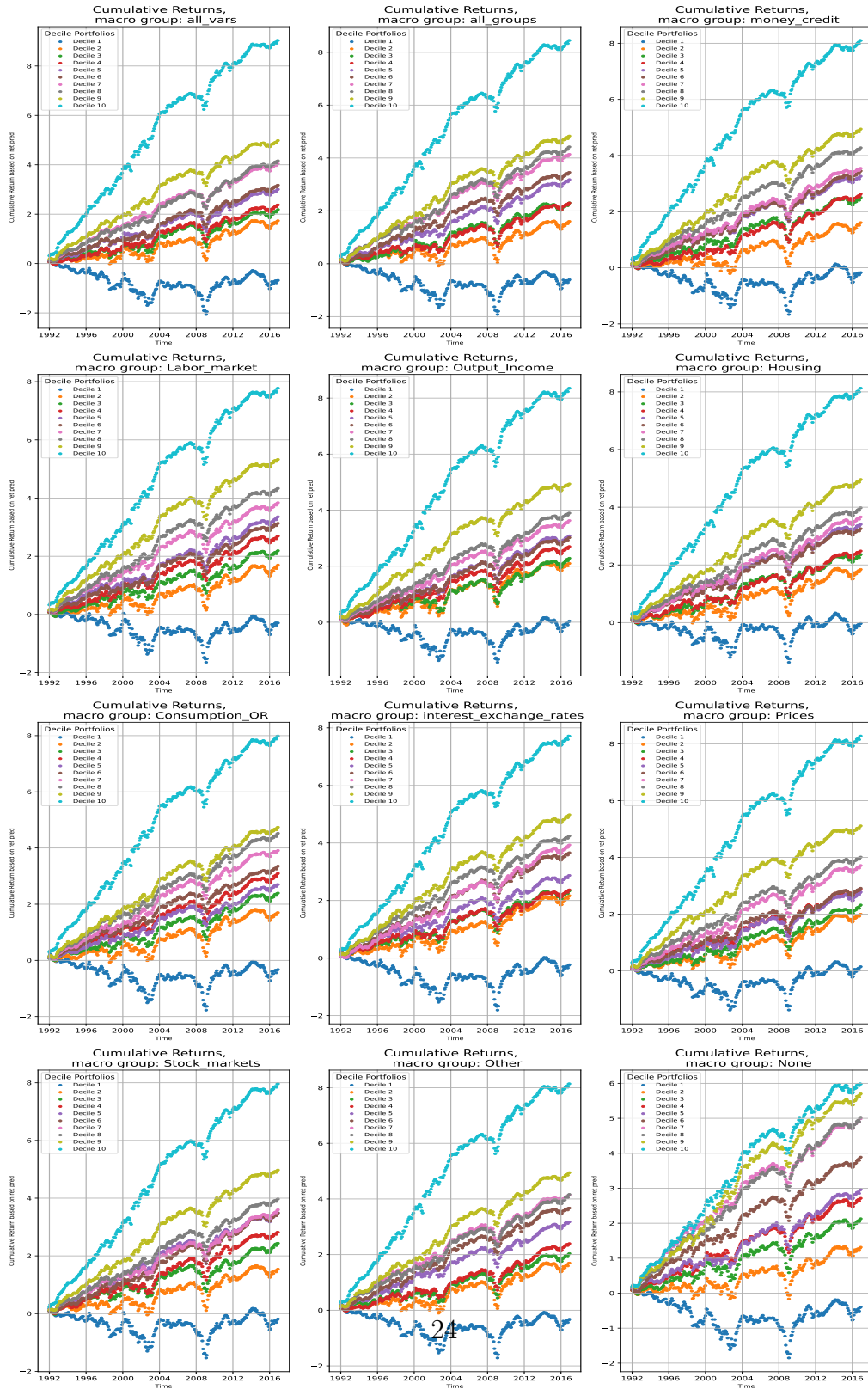


Figure 5.3: Rendement cumulatif de portefeuilles de déciles classé par prédiction de rendement

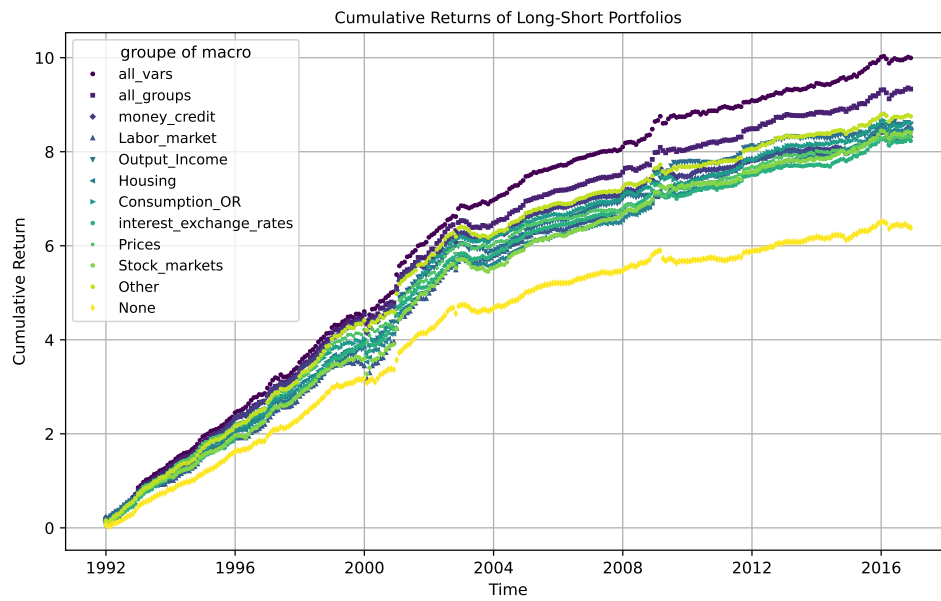


Figure 5.4: Rendement cumulatif de portefeuilles Long-court

CS_R2 deciles of momentum sorted portfolios with diff groups of macro

decile	all vars	all gr	Mny Cr	Lb Mrk	Out In	Hous	Cons Qr	Int Ex	Pce	Stk Mrt	Other	no Mac
1	0.712	0.391	0.248	0.398	-0.204	0.546	0.132	0.365	0.522	0.124	0.370	0.225
2	0.819	0.531	0.290	0.580	-0.319	0.765	0.266	0.464	0.659	0.312	0.566	0.730
3	0.864	0.613	0.401	0.688	-0.070	0.854	0.406	0.588	0.736	0.508	0.682	0.898
4	0.921	0.722	0.541	0.784	0.160	0.916	0.531	0.713	0.812	0.661	0.785	0.973
5	0.937	0.767	0.628	0.825	0.326	0.936	0.590	0.772	0.840	0.737	0.826	0.992
6	0.949	0.802	0.691	0.854	0.424	0.945	0.639	0.814	0.857	0.780	0.848	0.997
7	0.944	0.812	0.725	0.855	0.485	0.937	0.667	0.827	0.858	0.792	0.846	0.996
8	0.943	0.821	0.747	0.855	0.514	0.932	0.696	0.841	0.858	0.795	0.846	0.995
9	0.919	0.791	0.720	0.820	0.511	0.892	0.683	0.811	0.830	0.752	0.812	0.976
10	0.874	0.712	0.619	0.742	0.401	0.816	0.610	0.723	0.768	0.640	0.734	0.936
all	0.878	0.687	0.564	0.725	0.246	0.832	0.520	0.686	0.762	0.595	0.715	0.839

Figure 5.5: CSR2 de portefeuilles de déciles classé par Moment

Chapter 6

Discussion

Cette section discute des implications des résultats, des limitations de l'étude, et des suggestions pour des recherches futures.

6.1 Implications des Résultats

6.2 Limitations de l'étude

6.3 Suggestions

Chapter 7

Conclusion

Références

@article hamilton1989, title=A new approach to the economic analysis of non-stationary time series and the business cycle, author=Hamilton, James D, journal=Econometrica: Journal of the Econometric Society, pages=357–384, year=1989, publisher=JSTOR

@article fama1993common, title=Common risk factors in the returns on stocks and bonds, author=Fama, Eugene F and French, Kenneth R, journal=Journal of Financial Economics, volume=33, number=1, pages=3–56, year=1993, publisher=Elsevier