

HEC MONTRÉAL

Allocation de titres via la sélection de variables prédictives à travers l'apprentissage automatique

par

Alexis Tchuinté Tamen

**Sciences de la gestion
(Option Ingénierie Financière)**

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M. Sc.)*

Avril 2023
© Alexis Tchuinté Tamen, 2023

Résumé

La méthode des moments généralisée itérée, notée ITGMM (Iterated Generalized Method of Moments) est utilisée pour analyser le modèle non-paramétrique proposé par [Brandt \(1999\)](#). Nous reconstituons sa base de données, répliquons ses résultats et étendons l'analyse empirique sur la période post-publication. Nous constatons que les résultats de choix du portefeuille changent comparativement à ce qui est observé au développement même si le constat global reste le même, à savoir que les choix de portefeuille sont sensibles aux variables de prévision. Le paramètre identifiant le poids optimal de portefeuille à allouer aux actions, est estimé dans les cas des moments non-conditionnels et conditionnels. L'estimateur ITGMM est couplé à une méthode d'apprentissage automatique populaire ; à savoir LOOCV (leave-one-out cross validation) pour estimer les paramètres qui maximisent l'utilité espérée conditionnellement à l'état d'une variable prédictive. Le LOOCV est un variant du k-fold (ou plis) CV. L'approche utilisée est efficace, consistant et fiable, car permet par exemple de trouver la valeur exacte du résultat théorique lorsque $\gamma = 1$ (i.e, cas ayant une utilité logarithmique), ce qui n'est pas toujours le cas dans [Brandt \(1999\)](#). Le logiciel Matlab est utilisé pour coder et produire nos résultats.

Mots-clés

Méthodes des moments généralisés (GMM), Estimateur-GMM, Choix de portefeuille, Leave-one-out cross validation (LOOCV).

Abstract

The Iterated Generalized Method of Moments (ITGMM) is used to analyze the non-parametric model proposed by [Brandt \(1999\)](#). We reconstruct its database, replicate its results, and extend the empirical analysis to the post-publication period. We find that the portfolio choice results change compared to what is observed in development period although the overall finding remains the same, that is portfolio choices are sensitive to forecasting variables. The parameter identifying the optimal portfolio weight to be allocated to stocks is estimated in the cases of unconditional and conditional moments. The ITGMM estimator is coupled with a popular machine learning method; namely LOOCV (leave-one-out cross validation) to estimate the parameters that maximize the expected utility conditional on the state of a predictor variable. The LOOCV is a variant of the k-fold CV. The approach used is efficient, consistent and reliable because it allows for example to find the exact value of the theoretical result when $\gamma = 1$ (i.e., cases with logarithmic utility), which is not always the case in [Brandt \(1999\)](#). Matlab software is used to code and produce our results.

Keywords

Generalized Method of Moments (GMM), GMM-estimator, Portfolio selection, Leave-one-out cross validation (LOOCV).

Table des matières

Résumé	i
Abstract	iii
Liste des tableaux	vii
Liste des figures	ix
Liste des abréviations	xi
Remerciements	xiii
Introduction	1
Revue de la littérature	5
1 Méthodologie de recherche	9
1.1 Motivation du choix de la méthode	9
1.2 GMM itérée (ITGMM)	10
1.3 Méthode d'apprentissage automatique LOOCV	12
2 Applications	15
2.1 Description des données et construction de variables	15
2.2 Moments inconditionnel et conditionnel (Brandt (1999))	23
2.2.1 Mise en contexte	23

2.2.2	Choix du portefeuille optimal	24
2.2.3	Maximisation de l'utilité espérée	25
2.2.4	Estimateur inconditionnel GMM (Brandt (1999))	28
2.2.5	Estimateur conditionnel GMM (Brandt (1999))	29
2.2.5.1	Cas discret	29
2.2.5.2	Cas continu	30
2.2.5.2.1	Choix du kernel et sélection de la largeur de la fenêtre	31
2.3	Résultats empiriques	38
2.3.1	Illustrations : Moments inconditionnels (Brandt (1999))	38
2.3.1.1	Résultats sur les données de développement	38
2.3.1.2	Résultats empiriques sur les données post-développement	42
2.3.2	Illustrations : Moments conditionnels (Brandt (1999))	44
2.3.2.1	Résultats sur les données de développement	44
2.3.2.2	Résultats empiriques sur les données post-développement	50

Bibliographie **59**

Annexes **i**

Annexe A – Preuve du lemme 1	i
Annexe B – Preuve du théorème 1	i
Annexe C – Preuve du théorème 2	iii
Annexe D – Rappels sommaires sur les approches 1SGMM et 2SGMM	v
GMM à une étape (1SGMM)	v
GMM à deux étapes (2SGMM)	viii

Liste des tableaux

0.1	Liste sommaire d'articles académiques visant à prédire la moyenne (\times) et la variance (\checkmark) des rendements d'un marché boursier en utilisant les variables prédictives.	8
2.1	Statistiques descriptives des variables prédictives (toute l'historique)	20
2.2	Statistiques descriptives des variables (historique de développement)	21
2.3	Estimateur GMM : Choix du portefeuille et consommation en fonction de l'horizon et l'aversion au risque (Réplication Table III de Brandt (1999))	39
2.4	Estimateur GMM : Choix du portefeuille et consommation en fonction de l'horizon et l'aversion au risque (post-développement)	42
2.5	Estimateur conditionnel GMM : Choix du portefeuille en fonction des variables prédictives (Réplication Table V de Brandt (1999))	45
2.6	Estimateur conditionnel GMM : Choix du portefeuille en fonction des variables prédictives (période d'observation : janvier 1997 - décembre 2021)	50

Liste des figures

2.1	Évolution mensuelle du rendement du dividende de janvier 1947 à décembre 2021.	16
2.2	Évolution mensuelle de l'écart de défaut de janvier 1947 à décembre 2021.	17
2.3	Évolution mensuelle de l'écart de taux de janvier 1947 à décembre 2021.	18
2.4	Évolution mensuelle du rendement excédentaire retardé de l'indice NYSE de janvier 1947 à décembre 2021.	19
2.5	Histogrammes des variables prédictives : Dividend yield (figure (a)), Default premium (figure (b)), Term premium (figure (c)), Lagged excess return (figure (d)). Les valeurs illustrées sur l'axe des abscisses représentent les centres des classes de la variable. Chaque classe contient un nombre d'observations de l'échantillon de la variable. Les classes sont déterminées en utilisant les percentiles de la variable analysée (du 10 ^e au 90 ^e percentile). Au-dessus de chaque diagramme à bâton est placée la fréquence d'observations de l'échantillon dans chaque classe. L'axe des ordonnées représente le nombre d'observations de la classe.	22
2.6	(source Brandt (1999))	40
2.7	Estimateur GMM vs Figure 1 Brandt (1999). Illustrations de l'évolution du choix du portefeuille en fonction de l'horizon et l'aversion au risque	41
2.8	Évolution du poids optimal du portefeuille à allouer en actions en fonction de l'horizon sur la période post-développement du modèle de Brandt (1999).	43
2.9	Réplication de la figure 2a Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est le rendement du dividend (dividend yield). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).	46

2.10	Réplication de la figure 2b Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est la prime de défaut (default premium). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).	47
2.11	Réplication de la figure 2c Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est la prime de durée (term premium). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).	48
2.12	Réplication de la figure 2d Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est le rendement excédentaire retardé (lagged excess return). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).	49
2.13	Estimation du choix du portefeuille lorsque la variable prédictive est le rendement du dividende. Les données vont de janvier 1997 à décembre 2021 (période post-développement).	52
2.14	Estimation du choix du portefeuille lorsque la variable prédictive est la prime de défaut. Les données vont de janvier 1997 à décembre 2021.	52
2.15	Estimation du choix du portefeuille lorsque la variable prédictive est l'écart de taux. Les données vont de janvier 1997 à décembre 2021.	53
2.16	Estimation du choix du portefeuille lorsque la variable prédictive est le rendement excédentaire retardé. Les données vont de janvier 1997 à décembre 2021.	53
2.17	Données de développement : Illustration de l'évolution du choix du portefeuille en fonction de l'horizon et du rendement du dividende pour un investisseur d'aversion au risque $\gamma = 5$. La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).	54
2.18	Données de post-développement : Illustration de l'évolution du choix du portefeuille en fonction de l'horizon et du rendement du dividende pour un investisseur d'aversion au risque $\gamma = 5$. La figure est obtenue avec notre estimateur GMM conditionnel et les données sur la période allant de janvier 1997 à décembre 2021.	55

Liste des abréviations

GMM Generalized Method of Moments

HEC Hautes études commerciales

LOOCV Leave-one-out cross validation

M. Sc. Maîtrise ès sciences

Remerciements

Toute grâce excellente descend du ciel ! Je remercie Dieu pour sa bonté et la santé qu'il m'a accordées tout au long de ce travail.

Je remercie mon directeur de mémoire le Professeur Tolga Cenesizoglu pour tout son soutien, sa disponibilité, ses commentaires et son encadrement. Travailler sous votre direction m'a aidé à grandir dans le domaine de la recherche en finance. Merci pour tout !

Je dis un grand merci à ma moitié, mon épouse Darline, sans qui rien n'aurait été possible. Tu as su m'apporter ton soutien infaillible. Malgré toutes les difficultés, j'ai eu une épaule sur laquelle me pencher et repartir plus confiant. Merci pour tout chérie !

Je dis merci à mes trois enfants : Daryl Maël, Anne Keren, et Phanuella Dalexia. Vous êtes mes rayons de soleil, la boussole sans laquelle je ne tiendrais plus debout ! Je vous remercie infiniment pour la chaleur et le bonheur que vous m'apportez au quotidien.

Je dis merci chaleureusement à Madeleine Bouniol et Régine Tamen, mes deux Mams. Je n'aurai jamais été l'homme que je suis aujourd'hui sans votre accompagnement jalonné de vos conseils et sagesses qui me bercent encore aujourd'hui ! Vous êtes mes premières femmes et je vous témoigne toute ma gratitude. Merci infiniment !

Je dis un merci spécial à tous les enseignants de la M.Sc. en option Ingénierie Financière ainsi que des options Finance et Sciences de la Décision pour leurs savoirs et passions transmis. Merci pour tout !

Je ne saurais terminer sans remercier les bibliothécaires de HEC Montréal, pour leur implication, gentillesse et professionnalisme. Merci beaucoup !

*Je dédicace ce mémoire à mes trois enfants
Daryl Maël, Anne Keren, et Phanuella Da-
lexia.*

Introduction

La question de choix de portefeuille est un enjeu central pour l'investisseur. Pour prédire ses rendements futurs, l'investisseur doit analyser comment allouer sa richesse aux titres susceptibles de maximiser ses rendements. Le choix des variables prédictives est à cet effet très important dans ce processus puisqu'en fonction de la conjoncture économique, certaines variables économiques peuvent être plus importantes que d'autres et donc l'investisseur doit savoir lesquelles lui permettent de bien allouer sa richesse et dans quelles proportions.

Dans la littérature, plusieurs études se sont intéressées à ce sujet ([Merton \(1969\)](#), [Brandt \(1999\)](#), [Campbell et Viceira \(1999\)](#), [Heaton et Lucas \(2000\)](#), [Viceira \(2001\)](#), [Campbell et collab. \(2001\)](#), [Aït-Sahalia et Brandt \(2001\)](#), [Brandt et collab. \(2005b\)](#), [Brandt et Santa-Clara \(2006\)](#), [Ait-Sahalia et Brandt \(2007\)](#), [Welch et Goyal \(2008\)](#), [Cenesizoglu et Essid \(2011\)](#), [Cenesizoglu et Timmermann \(2012\)](#), [Cenesizoglu et collab. \(2017\)](#), [Di Maggio et collab. \(2020\)](#), [Connor et collab. \(2020\)](#)). Il ressort de cela que l'une des grandes difficultés est de développer une méthode simple, efficace et capable d'estimer adéquatement les paramètres des modèles formulés. La plupart des approches traditionnelles dites paramétriques supposent que l'on connaisse la distribution des variables (par exemple la loi normale). Cependant, bien que formuler des hypothèses sur la distribution des variables du modèle permette de faciliter les analyses mathématiques du modèle, cela pourrait néanmoins ne pas coïncider avec la réalité et donc avoir une incidence sur les résultats. D'où l'importance d'utiliser l'approche non-paramétrique notamment lorsque la distribution n'est pas connue.

Dans ce mémoire, nous suivons l'approche de [Brandt \(1999\)](#) et utilisons la méthode des moments généralisés itérée qui est une méthode basée uniquement sur les données et qui présente plusieurs avantages, notamment sa flexibilité, car elle permet d'estimer les paramètres d'un modèle sans faire d'hypothèses sur les distributions des variables. À notre connaissance, ce mémoire est la seule étude qui réplique les résultats de [Brandt \(1999\)](#) pour la première fois en considérant les mêmes historiques de données. Par ailleurs, notre estimateur est plus précis et fiable, car il permet par exemple de trouver la valeur exacte du résultat théorique du choix de consommation d'un investisseur ayant une utilité logarithmique (i.e., avec une aversion au risque $\gamma = 1$), ce qui n'est pas toujours le cas dans [Brandt \(1999\)](#). De plus, dans ce mémoire, nous analysons les résultats du modèle de [Brandt \(1999\)](#) sur la période post-publication. Cela nous permet de valider si les résultats du modèle changent en utilisant les données post-développement. Nos analyses sont effectuées pour les cas des moments inconditionnels et conditionnels aux variables prédictives. Pour le cas inconditionnel, nous observons que les valeurs des choix optimaux du portefeuille sont en général inférieures aux valeurs obtenues lors du développement. Par exemple, avec les données post-développement, un investisseur ayant une aversion au risque $\gamma = 5$ et un horizon d'un mois aurait été prudent en investissant environ 25% en actions au lieu de 82% en considérant les données de développement. Par ailleurs, nous remarquons que pour les horizons supérieurs à 3 mois, le cas des moments inconditionnels donne des trajectoires des choix optimaux de portefeuille similaires à celles obtenues au développement. Pour le conditionnel, nous observons quelques changements sur les résultats, même si la conclusion générale reste la même, à savoir que le choix de portefeuille est sensible aux variables prédictives. Par conséquent, l'investisseur choisirait des portefeuilles différents en fonction des états de chaque variable prédictive. Toutefois, comme avec les données de développement, nous remarquons que les courbes des choix de portefeuille sont croissantes en fonction du rendement du dividende et aussi en fonction du rendement excédentaire retardé, tandis qu'elles passent de non-linéaires à approximativement linéairement décroissantes pour les variables prime de défaut et prime de durée. Ultime-ment, conditionnellement aux variables prédictives, le portrait de décisions du portefeuille

change en passant des données de développement aux données post-développement. Toutefois, le rôle de la variable prédictive est très important ainsi que l'horizon. Par exemple, [Barberis \(2000\)](#) montre que le fait de conditionner le rendement du dividende réduit la variance des rendements cumulés prédits à long terme, ce qui conduit une allocation plus importante en actions pour les investisseurs à long terme. Pour [Barberis \(2000\)](#), l'augmentation du portefeuille en fonction de la variable "dividend yield" est linéaire, tandis que [Brandt \(1999\)](#) présente une dynamique beaucoup plus riche, à savoir une évolution non-linéaire, voire non-monotone, pour certains rendements du dividende, à l'horizon. Dans ce mémoire, nos résultats sur la période de développement sont similaires à ceux de [Brandt \(1999\)](#). Cependant, en utilisant les données post-développement, le pattern d'évolution du choix optimal du portefeuille en fonction du dividende change, même si la forme demeure non-linéaire. Nous remarquons notamment que les valeurs des pourcentages du portefeuille allouées en actions sur la période de développement sont supérieures à celles de la période post-développement. Ce qui pourrait s'expliquer par la baisse du rendement du dividende observée sur la période post-publication du modèle.

Ce mémoire est séparé en quatre parties. La première partie concerne l'état de l'art et la problématique du mémoire. La deuxième partie décrit la méthodologie de la recherche utilisée (voir Chapitre 1). Ici, nous décrivons brièvement les algorithmes d'implémentation de la méthode ITGMM et la méthode d'apprentissage automatique utilisée dans ce mémoire. La troisième partie (voir Chapitre 2) concerne les applications de la méthode sur les données empiriques. Ici, nous reconstituons la base de données et les variables utilisées, et répliquons les résultats du modèle de [Brandt \(1999\)](#). Nous étendons les analyses sur les données post-développement pour comparer et voir si les résultats du modèle ne changent pas si l'investisseur utilise les données antérieures au développement. La dernière partie concerne les discussions générales des résultats de nos travaux et la conclusion.

Revue de la littérature

Les pionniers de la littérature sur le choix dynamique de portefeuille sont [Markowitz \(1952\)](#) et [Fama \(1970\)](#) en temps discret et [Merton \(1969\)](#), [Samuelson \(1969\)](#) et [Merton \(1971\)](#) en temps continu. Cette littérature originelle a apporté de nombreuses informations importantes sur la théorie économique qui sous-tend le choix du portefeuille optimal d'un investisseur, comme les effets de la diversification et le compromis entre le rendement attendu et le risque dans le paradigme moyenne-variance de [Markowitz \(1952\)](#). Cependant, le réalisme des problèmes de choix de portefeuille étudiés a très souvent été limité par la rareté des résultats analytiques et par la puissance limitée des méthodes numériques existantes ([Brandt et collab. \(2005a\)](#)). La problématique de notre mémoire est de comprendre comment utiliser une méthode quantitative adéquate pour adresser le problème d'allocation du portefeuille en utilisant certaines variables prédictives.

Certains auteurs formulent des hypothèses sur la distribution des rendements (exemple : la loi normale ou le processus de Wiener pour la dynamique des prix) afin de résoudre analytiquement le problème de choix de portefeuille optimal ([Merton \(1969\)](#), [Samuelson \(1969\)](#), [Merton \(1971\)](#), [Kogan et Uppal \(2000\)](#)), tandis que d'autres considèrent que les rendements ont une distribution arbitraire et basent leurs analyses sur des simulations (exemple [Brandt \(1999\)](#), [Balduzzi et Lynch \(1999\)](#), [Barberis \(2000\)](#), [Brandt et collab. \(2005a\)](#)). L'avantage principal de considérer des modèles basés sur les hypothèses de distribution de certaines variables d'état est qu'elles facilitent l'étude analytique et le calcul d'une solution fermée. Cependant, les solutions de forme fermée ne sont disponibles que pour quelques paramétrages spéciaux des préférences de l'investisseur et de la dy-

namique des rendements, comme le montrent [Kim et Omberg \(1996\)](#), [Wachter \(2002\)](#) et [Liu \(2007\)](#). Dans la littérature, l'approche la plus populaire est la simulation numérique. Elle consiste à discrétiser l'espace d'état. Une fois l'espace d'état discrétisé, la fonction valeur peut être évaluée par un choix d'intégration quadratique ([Balduzzi et Lynch \(1999\)](#)), de simulations ([Barberis \(2000\)](#)), de discrétisations binomiales ([Dammon et collab. \(2001\)](#)), ou de régressions non-paramétriques ([Brandt \(1999\)](#)), puis l'optimisation dynamique peut être résolue par récursivité arrière. Cependant, excepté l'approche non-paramétrique de [Brandt \(1999\)](#), ces méthodes numériques et d'approximation de solution partagent quelques limitations importantes ([Brandt et collab. \(2005a\)](#)). Par exemple, elles supposent non-réalisme de simples distributions de rendements, et donc les solutions obtenues peuvent ne pas coïncider avec la réalité du fait des hypothèses formulées. Plus important encore, selon [Brandt et collab. \(2005a\)](#), ces méthodes ne peuvent pas traiter le grand nombre de variables d'état avec des dynamiques compliquées qui apparaissent dans de nombreux problèmes réalistes de choix de portefeuille. Dans ce mémoire, nous suivons l'approche non-paramétrique de [Brandt \(1999\)](#) qui ne présente pas de limitations issues, notamment de la considération des hypothèses de distributions des rendements ou des variables d'état. De plus, l'approche non-paramétrique a l'avantage de traiter un grand nombre de variables d'état en considérant les dynamiques linéaires et non-linéaires. Cette flexibilité est un atout dans l'analyse des problèmes complexes de choix de portefeuille.

Dans la littérature du choix optimal du portefeuille, la sélection de variables prédictives permettant de prédire les rendements est un enjeu important. En effet, il est difficile de déterminer sur quelle sélection ou combinaison de variables prédictives l'investisseur doit se concentrer ([Aït-Sahalia et Brandt \(2001\)](#)). Ainsi, plusieurs études de la littérature empirique pré-déterminent une ou plusieurs variables d'état pour mener leurs analyses (voir [Tableau 0.1](#)). La plupart de ces études cherchent à déterminer les variables qui prédisent le mieux le ratio des deux premiers moments conditionnels (moyenne et variance) de la distribution des rendements. Cependant, suivant [Aït-Sahalia et Brandt \(2001\)](#), choisir les variables qui prédisent le mieux la moyenne et la variance séparément pourrait s'avérer contre-productif et susciter de nombreuses questions. Par exemple, que faire si

une variable a un effet positif à la fois sur la moyenne (ce qui plaît à l'investisseur) et sur la variance (ce qui déplaît à l'investisseur) ? Comment saisir l'importance relative que les préférences de l'investisseur accordent aux différents moments ? Par ailleurs, même si la moyenne et la variance sont prédites simultanément, l'approche moyenne-variance présentent quelques objections. Premièrement, elle ignore toutes préférences au-delà des moments de rendements d'ordre supérieur à deux, en particulier le skewness et le kurtosis des rendements. Deuxièmement, elle représente la maximisation de l'utilité espérée pour un cas particulier de l'utilité quadratique, et donc ne permet pas de spécifier des cas plus généraux. Troisièmement, l'approche moyenne-variance est inhérente au problème de l'investisseur myope qui ne considère qu'une période, alors qu'il y a une évidence que la plupart des problèmes d'investissement impliquent des horizons plus longs avec un rééquilibrage intermédiaire du portefeuille. Toutes ces remarques et critiques suggèrent que dans un contexte de choix de portefeuille, on doit sélectionner des variables pour prédire directement les pondérations optimales du portefeuille (Brandt (1999), Aït-Sahalia et Brandt (2001), Brandt et Santa-Clara (2006)), plutôt que de sélectionner d'abord des variables pour prédire les caractéristiques distinctes de la distribution des rendements et d'explorer ensuite leurs implications pour l'allocation d'actifs. Par conséquent, dans ce mémoire, nous utilisons l'approche non-paramétrique suivant Brandt (1999) qui permet de formuler explicitement le problème de choix de portefeuille optimal comme un problème de maximisation de l'utilité espérée inter-temporelle. En outre, l'approche non-paramétrique a pour avantages : (i) d'éviter toute hypothèse sur la distribution des rendements, (ii) de considérer plusieurs variables prédictives, (iii) d'estimer directement les poids des portefeuilles sans passer par les caractéristiques des distributions des rendements, (iv) d'être utilisée dans les cas complexes de dynamiques de portefeuille, par exemple les problèmes non-linéaires de choix de portefeuille. Tous ces critères sont cruciaux dans la résolution de problèmes réalistes de choix de portefeuille et permettent de relier la formulation théorique du problème et sa solution aux données.

TABLE 0.1 – Liste sommaire d’articles académiques visant à prédire la moyenne (×) et la variance (✓) des rendements d’un marché boursier en utilisant les variables prédictives.

Auteurs	Variables prédictives														
	TS	DY	LR	TBY	Def	Trend	LV	LSR	DER	PS	ratios				
											LnDP	ITC	DE	CW	BM
Fama et Schwert (1977)				×											
Engle (1982)							✓								
Bollerslev (1986)			✓												
Keim et Stambaugh (1986)					×	×									
Campbell (1987)	×	✓													
French et collab. (1987)							✓								
Campbell et Shiller (1988a)		×													
Fama et French (1989)	×	×			×										
Schwert (1989)					✓		✓		✓						
Cochrane (1991)												×			
Ferson et Harvey (1991)	×	×	×	×	×										
Whitelaw (1994)				✓	✓		✓			✓					
Lamont (1998)													×		
Pontiff et Schall (1998)															×
Harvey (2001)	✓	✓		✓	✓		✓		✓						
Lettau et Ludvigson (2001)															×
Aït-Sahalia et Brandt (2001)	×				×	×					×				

"TS" i.e. term spread, "DY" i.e. dividend yield, "LR" i.e. lagged return, "TBY" i.e. treasury bill yield, "Def" i.e. default spread,

"LV" i.e. lagged variance, "LSR" i.e. lagged squared returns, "DER" i.e. debt-to-equity, "PS" i.e. paper spread, "LnDP" i.e. log dividend price,

"ITC" i.e. investment-to-capital, "DE" i.e. to dividend-to-earnings, "CW" i.e. consumption-to-wealth, "BM" i.e. book-to-market.

Chapitre 1

Méthodologie de recherche

La méthode que nous utilisons dans ce mémoire est la Méthode des Moments Généralisés (GMM) itérée.

1.1 Motivation du choix de la méthode

L'estimation GMM est une méthode semi-paramétrique qui a été formalisée par [Hansen \(1982\)](#) et est depuis devenue l'une des méthodes d'estimation les plus utilisées pour les modèles en économie, en finance et dans plusieurs autres domaines (voir la table 1.1, p16-17 dans [Hall \(2005\)](#) pour une revue de la littérature sur les applications de la GMM dans divers champs). En économie, par exemple, il permet d'estimer un modèle structurel ; équation par équation. En finance, il permet d'estimer la plupart des données telles que les rendements boursiers. L'estimateur GMM est utilisé pour les modèles linéaires et non-linéaires. Contrairement à l'estimation par maximum de vraisemblance (MLE), l'estimation par GMM ne nécessite pas une connaissance complète de la distribution des données. Seuls les moments spécifiés dérivés d'un modèle sous-jacent sont nécessaires pour l'estimation GMM. En plus d'être utilisé lorsque la distribution est inconnue, l'estimateur par GMM permet aussi d'estimer les modèles pour lesquels il y a plus de conditions de moment que de paramètres. Par conséquent, il est le moins susceptible d'être mal spécifié. Il s'agit d'une caractéristique qui est unique à l'estimation GMM.

Suivant ([Chaussé \(2010\)](#)), la méthode des moments généralisés peut être vue comme une généralisation de plusieurs autres méthodes classiques telles que moindres carrés ordinaires (OLS), variables instrumentales (IV) ou le maximum de vraisemblance (MLE). Les propriétés des estimateurs de OLS dépendent de l'exogénéité des régresseurs et de la circularité des résidus, tandis que celles de MLE dépendent du choix de la fonction de vraisemblance. L'estimateur GMM est plus flexible puisqu'il ne nécessite que quelques hypothèses sur les conditions du moment.

Dans la littérature, il existe quatre approches de l'estimateur GMM : (1) l'estimateur GMM à une étape noté 1SGMM, (2) l'estimateur GMM à deux étapes noté 2SGMM, (3) l'estimateur GMM itéré noté ITGMM, et (4) l'estimateur continument actualisé noté CUE.

En général, l'estimateur efficient GMM est soit le 2SGMM ou ITGMM ou CUE. L'estimateur 1SGMM dépend du choix arbitraire de la matrice de poids initiale et donc peut être sous-optimal dans les cas où la matrice de poids est mal choisie. En pratique, l'itération est requise pour obtenir l'estimateur optimal ([Hall \(2005\)](#)). Dans ce document, nos analyses se baseront sur l'estimateur itéré ITGMM. La sous-section 1.2 présente une brève description de l'approche ITGMM. Néanmoins, pour le lecteur, les approches 1SGMM et 2SGMM sont présentées en Annexe D (2.3.2.2) puisqu'elles sont en effet les étapes 1 et 2 de l'ITGMM. Par ailleurs, le lecteur peut se référer à [Hayashi \(2000\)](#) et [Hall \(2005\)](#) pour plus de détails sur les approches de GMM.

1.2 GMM itérée (ITGMM)

La méthode consiste à estimer un vecteur de paramètres $\theta^* \in \mathbb{R}^p$ à partir d'un modèle basé sur le vecteur de $q \times 1$ moments inconditionnels suivant :

$$E[\psi(\theta^*, x)] = 0, \tag{1.1}$$

où le vecteur $\psi(\cdot, \cdot) \in \mathcal{M}_{q \times 1}$ est appelé la fonction moment, et $x \in \mathcal{M}_{m \times 1}$ est un vecteur de variables explicatives qui peuvent être les données transversales (cross-sectional),

séries temporelles ou les deux. Pour que GMM produise des estimations cohérentes à partir des conditions ci-dessus en (1.1), θ^* doit être l'unique solution de $E[\psi(\theta, x)] = 0$ et appartenir à un espace compact. D'autres hypothèses sont requises telles que : la fonction $\psi(\theta, x)$ doit être deux fois continument différentiable par rapport à θ et mesurable en x , l'espérance du produit extérieur notée $\Delta = E[\psi(\theta^*, x)\psi(\theta^*, x)'] \in \mathcal{M}_{q \times q}$ et la matrice des dérivées partielles par rapport à θ notée $\Gamma = E\left[\frac{\partial \psi}{\partial \theta'}(\theta^*, x)\right] \in \mathcal{M}_{q \times p}$ doivent être de rang maximal par rapport à la colonne (i.e., son rang est p , le nombre de colonnes). Toutefois, une des caractéristiques importantes de l'estimateur GMM est qu'il n'impose aucune condition sur la distribution de x . Pour plus de détails techniques et d'autres conditions de régularités voir [Newey et McFadden \(1994\)](#), [Newey et Smith \(2004\)](#) et [Hall \(2005\)](#).

Remarque 1 *Une condition nécessaire pour l'identification de θ^* est la condition d'ordre*

$$q \geq p, \tag{1.2}$$

puisque si $q < p$, alors $\text{rang}(\Gamma) < p$ ce qui contredit l'hypothèse selon laquelle $\text{rang}(\Gamma)$ doit être au moins égale au nombre de colonnes p .

Définition 1 *i) Lorsque le nombre de conditions de moment (q) est égale au nombre de paramètres à estimer (p), le système est bien identifié (just identified).*

ii) Lorsque $q > p$, le système est sur-identifié (overidentified).

iii) Lorsque $q < p$ (i.e., la condition d'ordre n'est pas satisfaite), alors le système n'est pas identifiable (underidentified).

Définition 2 — *Le processus qui permet de dire comment étant donné un échantillon on obtient une valeur approximative de θ est appelé un estimateur (exemple l'estimateur GMM).*

— *Le résultat pour un échantillon donné est appelé une estimation de θ .*

L'estimateur GMM itéré utilise l'estimateur GMM à deux étapes. Le processus consiste à répéter les étapes du 2SGMM jusqu'à ce que l'estimateur $\hat{\theta}$ ne change plus entre deux itérations consécutives. L'algorithme est le suivant :

1. Calculer $\theta_0 = \arg \min_{\theta} J_{I_q, n}(\theta)$.
2. Calculer $\hat{\Delta}^{-1} = \left[\frac{1}{n} \sum_{i=1}^n \psi(\theta_0, x_i) \psi(\theta_0, x_i)' \right]^{-1}$.
3. Calculer $\hat{\theta}_1 = \arg \min_{\theta} J_{\hat{\Delta}^{-1}, n}(\theta)$.
4. Si $\|\hat{\theta}_1 - \theta_0\| < \varepsilon$, alors passer à l'étape 5, sinon faire $\theta_0 = \hat{\theta}_1$ et retourner à l'étape 1.
5. L'estimateur GMM itéré $\hat{\theta}_{itgmm} = \hat{\theta}_1$,

où $\varepsilon \rightarrow 0$, lorsqu'on augmente la précision. [Hamilton \(1994\)](#) (p 413) montre que comparativement à l'estimateur en deux étapes, l'estimateur GMM itéré présente un avantage pratique, à savoir que les estimations résultantes sont invariantes par rapport à l'échelle des données et à la matrice de pondération initiale \hat{W} .

1.3 Méthode d'apprentissage automatique LOOCV

Pour estimer les paramètres du modèle conditionnel, nous utilisons la méthode d'apprentissage automatique Leave-One-Out Cross Validation, nommée LOOCV. La méthode LOOCV est basée sur les données. Cette approche minimise l'erreur quadratique moyenne dans la prédiction hors échantillon (out-of-sample). La validation croisée de type "leave-one-out" se résume en trois grandes étapes :

1. Pour chaque observation $\{R_{t+1}, z_t\}$, utiliser le reste d'observations de l'échantillon $\{R_{j+1}, z_j\}_{j=1, j \neq t}^T$ pour estimer $\{\alpha(z_t, \tau), q(z_t, \tau)\}$.
2. Calculer l'erreur dans la prédiction hors échantillon

$$v_{t+1} = \{\alpha_{t+1} - \hat{\alpha}(z_t, \tau), q_{t+1} - \hat{q}(z_t, \tau)\} \quad (1.3)$$

3. Choisir λ qui minimise la somme des carrés des erreurs prédicés, i.e.,

$$\hat{\lambda} = \arg \min_{\lambda} \sum_{t=1}^T \mathbf{v}'_{t+1} \mathbf{v}_{t+1}. \quad (1.4)$$

Pour plus de détails sur les applications de l'approche de la validation croisée de type "leave-one-out" à la régression non-paramétrique, le lecteur peut se référer aux publications suivantes : ([Härdle, 1984](#)), [Härdle et Marron \(1985\)](#), [Hall et collab. \(1995\)](#), [Wu \(1997\)](#) et [Lee et Lemieux \(2010\)](#).

Chapitre 2

Applications

2.1 Description des données et construction de variables

Nous construisons notre base de données pour les mêmes variables utilisées par [Brandt \(1999\)](#), à savoir : le rendement du dividende (dividend yield), l'écart de défaut (term spread), l'écart de taux (term spread), et le rendement excédentaire retardé (lagged excess return). L'idée est de reproduire une base de données similaire à celle de [Brandt \(1999\)](#) afin de comparer nos résultats sur la même période et ensuite tester l'approche sur une période post-développement du modèle de [Brandt \(1999\)](#). Cependant, la disponibilité des données reste évidemment un grand challenge pour mener des études comparatives. Dans ce mémoire, nous avons eu recours aux données qui viennent de plusieurs sources fiables comme la FRED (Federal Reserve Economic Data) et le CRSP (Center for Research in Security Prices) à travers la banque de données WRDS (Wharton Research Data Services) à laquelle est affiliée HEC de Montréal.

Nous avons étendu la base de données de [Brandt \(1999\)](#) en considérant un historique de données allant de janvier 1947 à décembre 2021.

La variable prédictive "dividend yield" est la somme des dividendes réinvestis versés sur l'indice NYSE au cours des 12 derniers mois. Nous calculons cette variable en utilisant les données du prix de l'indice NYSE extrait de WRDS. La figure [2.1](#) présente son

évolution sur la période allant de janvier 1947 à décembre 2021.

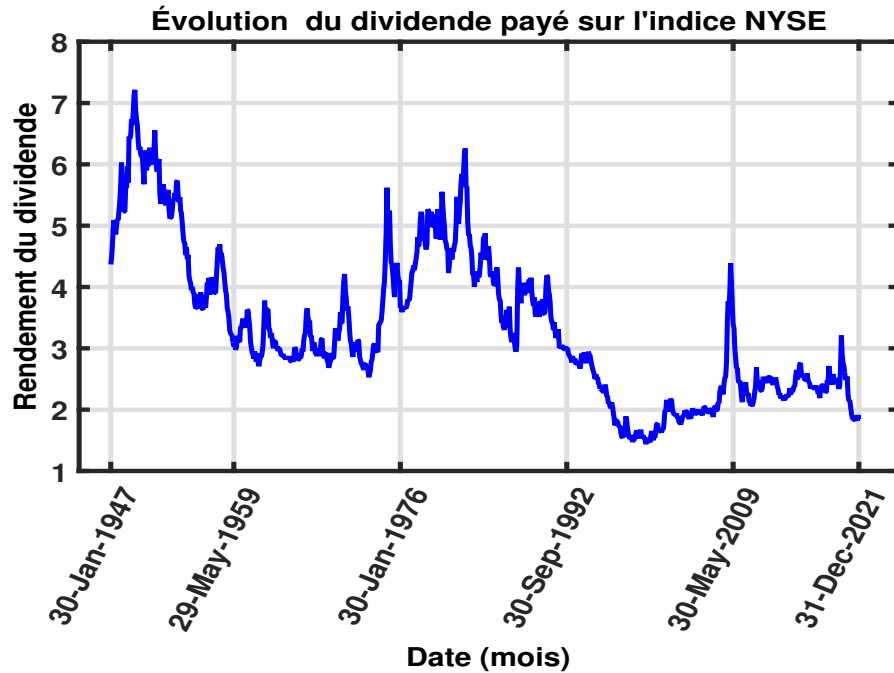


FIGURE 2.1 – Évolution mensuelle du rendement du dividende de janvier 1947 à décembre 2021.

La variable prédictive "default spread" se calcule comme étant la différence entre le rendement des entreprises notées BAA par le rendement des entreprises notées AAA. Cette variable prédictive est un indicateur de risque. Les données de notation des entreprises notées BAA et AAA proviennent du site de la FRED. La figure 2.2 présente l'évolution de la variable "default spread" sur la période allant de janvier 1947 à décembre 2021.

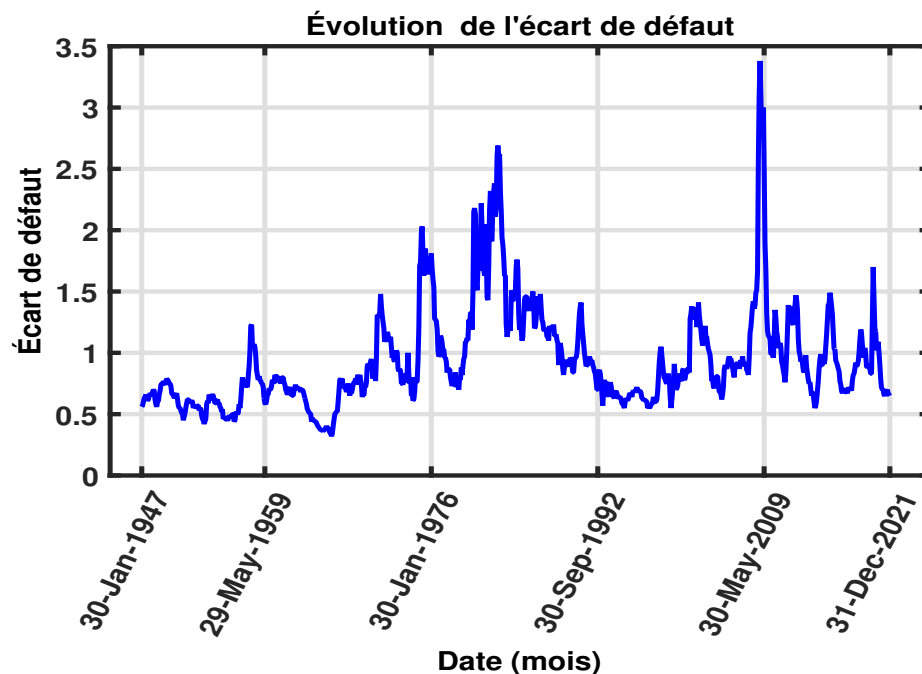


FIGURE 2.2 – Évolution mensuelle de l'écart de défaut de janvier 1947 à décembre 2021.

La variable prédictive "term spread" est la différence entre le rendement des obligations gouvernementales à long terme par le rendement du bon du Trésor 90 jours (90 days treasury bill yield). Les données du rendement des obligations gouvernementales à long terme viennent du site de Amit Goyal (voir [Welch et Goyal \(2008\)](#)) et celles du rendement du bon du Trésor 90 jours proviennent du site de la FRED. La figure 2.3 présente l'évolution de la variable "term spread" sur la période allant de janvier 1947 à décembre 2021.

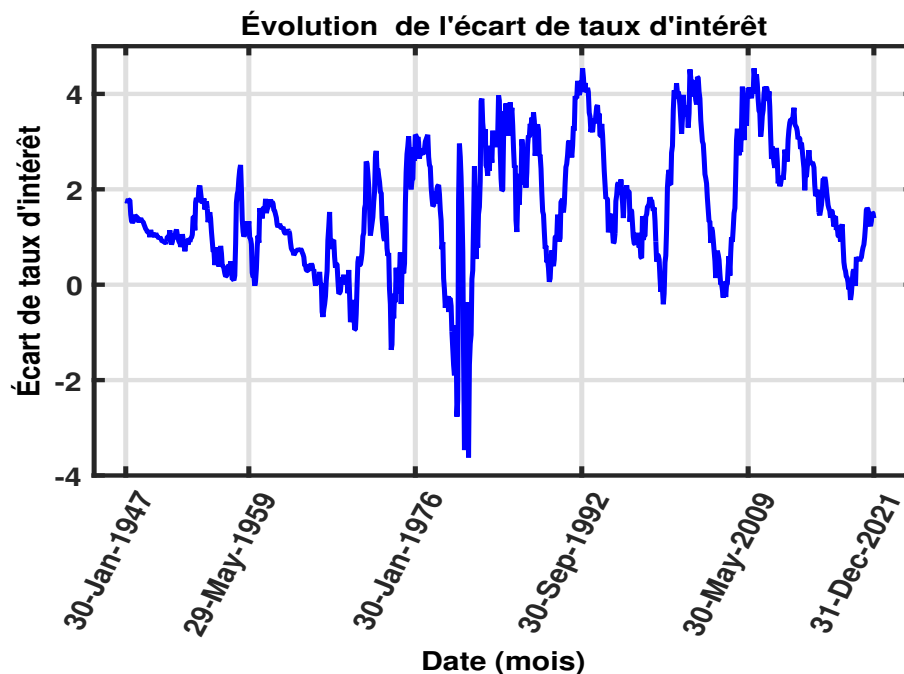


FIGURE 2.3 – Évolution mensuelle de l'écart de taux de janvier 1947 à décembre 2021.

La variable prédictive "lagged excess return" ou rendement excédentaire retardé est le rendement excédentaire du NYSE mesuré au cours de la période de rééquilibrage précédente. On utilise les données du rendement de l'indice NYSE et du taux sans risque (taux de rendement du bon du Trésor 30 jours) pour calculer cette variable. Le rendement de l'indice NYSE et le taux sans risque sont extraits de WRDS. La figure 2.4 présente l'évolution de la variable "lagged excess return" sur la période allant de janvier 1947 à décembre 2021.

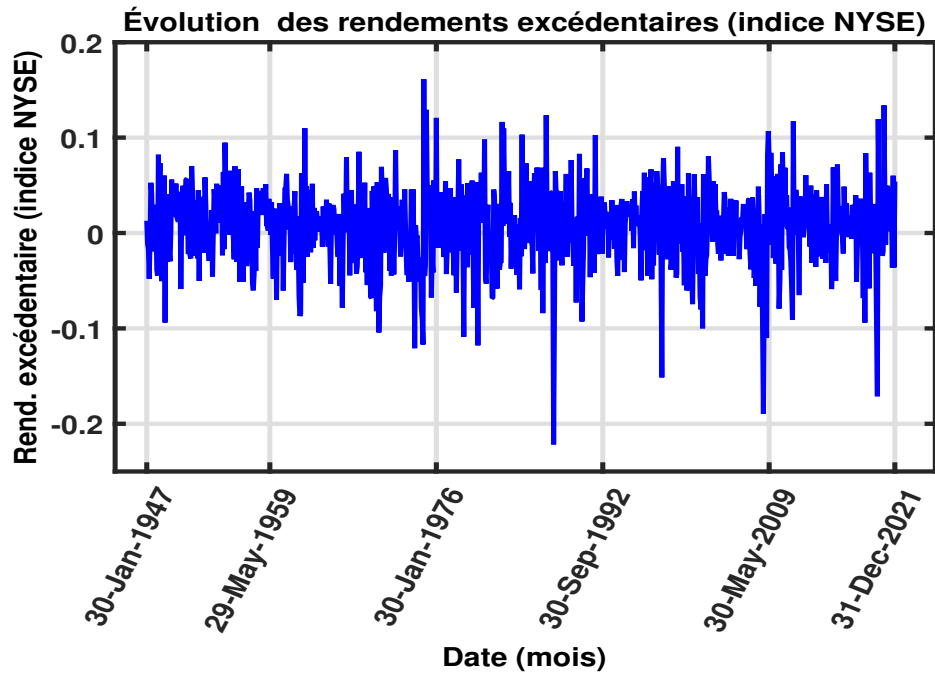


FIGURE 2.4 – Évolution mensuelle du rendement excédentaire retardé de l'indice NYSE de janvier 1947 à décembre 2021.

La description des données sur toute l'historique est présentée dans le Tableau 2.1.

TABLE 2.1 – Statistiques descriptives des variables prédictives (toute l’historique)

Statistiques	Variables			
	Div	Def	Term	Ret
Moyenne	3.3529	0.9436	1.6341	0.6474
Écart-type	1.2392	0.4240	1.3384	4.1418
Skewness	0.6923	1.9204	-0.0093	-0.5198
Kurtosis	2.7715	8.2560	2.9972	5.1936
Percentile				
5%	1.6808	0.4800	-0.2786	-6.1305
25%	2.3999	0.6700	0.7215	-1.8259
50%	3.0341	0.8300	1.4694	0.9784
75%	4.1306	1.1200	2.6640	3.2957
95%	5.7075	1.7650	3.9779	6.5092
Corrélation avec				
Div	1.0000			
Def	0.1468	1.0000		
Term	-0.2024	0.2605	1.0000	
Ret	-0.0382	0.0157	0.0666	1.0000

Nous séparons les données en deux : données de développement du modèle de [Brandt \(1999\)](#) (janvier 1947 à décembre 1996) et données post-développement (janvier 1997 à décembre 2021). Les données de développement nous permettent de répliquer les résultats de [Brandt \(1999\)](#) en utilisant la méthode GMM et les données post-développement nous permettent d’analyser si les résultats obtenus lors du développement ont changé ou pas.

La description des données est présentée dans le [Tableau 2.2](#).

TABLE 2.2 – Statistiques descriptives des variables (historique de développement)

Statistiques	Variables			
	Div	Def	Term	Ret
Moyenne	3.9450	0.9190	1.4113	0.6483
Écart-type	1.0725	0.4300	1.2894	4.0259
Skewness	0.6965	1.4843	-0.1193	-0.3750
Kurtosis	2.6644	5.1719	3.5207	5.1017
Percentile				
5%	2.6917	0.4500	-0.4850	-5.9399
25%	3.0056	0.6400	0.6300	-1.8115
50%	3.7430	0.7700	1.2700	0.9175
75%	4.6729	1.1200	2.2650	3.1920
95%	6.0229	1.8100	3.6050	6.4379
Corrélation avec				
Div	1.0000			
Def	0.2033	1.0000		
Term	-0.1218	0.2062	1.0000	
Ret	-0.0406	0.0709	0.1083	1.0000

Note : Les données des rendements proviennent du CRSP à partir de WRDS.

Les statistiques descriptives de la Table 2.2 sont similaires à ceux de la Table II de [Brandt \(1999\)](#), excepté celles de la variable "term spread" qui sont légèrement différentes. Notons que pour calculer la variable "term spread", [Brandt \(1999\)](#) utilise les rendements annualisés d'un portefeuille d'obligations gouvernementales à long-terme sans nommer spécifiquement s'il s'agit par exemple d'une obligation 10–an ou 20–an ou 30–an des États-Unis. Pour nos analyses, pour être cohérent avec ce qui est utilisé très souvent dans la littérature, nous avons choisi les obligations gouvernementales à long-terme 10–an des États-Unis. Par ailleurs, observons que puisque nos valeurs des statistiques descriptives de la variable "term spread" et celles de [Brandt \(1999\)](#) sont proches, la différence notée pourrait être due au fait que certaines banques de données mettent à jour leurs données régulièrement, et donc les valeurs de celles-ci sont potentiellement réajustées. Les rendements des obligations gouvernementales de [Brandt \(1999\)](#) sont ex-

traits de la plateforme CITIBASE, tandis que les nôtres proviennent du site de Amit Goyal (<http://www.hec.unil.ch/agoyal/>, Welch et Goyal (2008)).

Les histogrammes des variables prédictives sont présentés à la Figure 2.5.

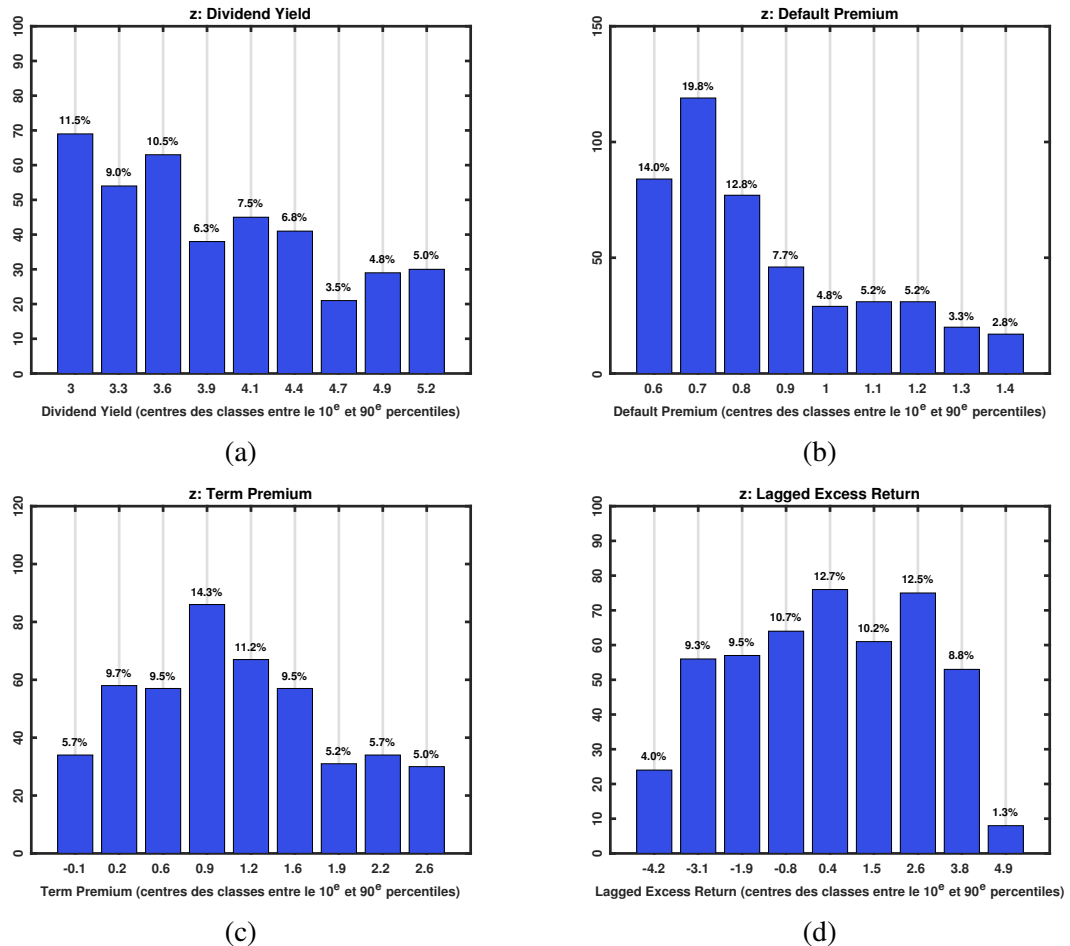


FIGURE 2.5 – Histogrammes des variables prédictives : Dividend yield (figure (a)), Default premium (figure (b)), Term premium (figure (c)), Lagged excess return (figure (d)). Les valeurs illustrées sur l'axe des abscisses représentent les centres des classes de la variable. Chaque classe contient un nombre d'observations de l'échantillon de la variable. Les classes sont déterminées en utilisant les percentiles de la variable analysée (du 10^e au 90^e percentile). Au-dessus de chaque diagramme à bâton est placée la fréquence d'observations de l'échantillon dans chaque classe. L'axe des ordonnées représente le nombre d'observations de la classe.

L'investisseur a le choix entre 2 titres : la valeur pondérée de l'indice NYSE et un bon du Trésor 30-jours. On applique la déflation sur les rendements nominaux en utilisant le taux de changement sur l'indice du prix à la consommation, nommé CPI (Consumer Price Index). Cela permet d'enlever l'effet de l'inflation sur le rendement.

Brandt (1999) présente les résultats empiriques pour un investisseur ayant une aversion au risque relative γ compris entre 1 (cas de l'utilité logarithmique) et 10. L'horizon varie entre 1 mois et 20 ans, et la période de re-balancement (temps qui s'écoule entre deux décisions d'investissements consécutives) varie entre 1 mois et 4 ans. Le coefficient de préférence temporelle est arbitrairement fixé à $\beta = 0.99$.

2.2 Moments inconditionnel et conditionnel (Brandt (1999))

Nous traitons deux cas de figure : (1) la méthode des moments inconditionnels c-à-d où les opportunités d'investissement sont constantes, et (2) la méthode des moments conditionnels où les opportunités d'investissement dépendent des états des variables prédictives. Avant de présenter les résultats empiriques, nous décrivons ces deux cas de figure. Ces étapes sont nécessaires pour l'implémentation de la méthode. Nous utilisons le logiciel Matlab pour coder et produire nos résultats.

2.2.1 Mise en contexte

Brandt (1999) utilise une approche non-paramétrique pour examiner comment le portefeuille et le choix de consommation dépendent des variables prédictives qui varient dans le temps. Il estime notamment un portefeuille à une période et un portefeuille à plusieurs périodes ainsi que les règles de consommation pour un investisseur ayant une aversion au risque constante (CRRA i.e., constant relative risk aversion) et un horizon d'investissement compris entre 1 mois et 20 ans.

Dans cette sous-section, nous allons nous intéresser au portefeuille à une période pour lequel l'investisseur choisit d'allouer une fraction α de sa richesse aux actions et le reste de sa richesse dans un titre sans risque. Les actions donnent un rendement excédentaire incertain \tilde{R}_{t+1}^e par rapport au taux sans risque R^f . Toutefois, considérer qu'il existe des opportunités d'investissement signifie que la distribution des rendements peut varier à

travers le temps. Si c'est le cas, les rendements pourraient dépendre d'une ou plusieurs variables prédictives observables z_t .

2.2.2 Choix du portefeuille optimal

L'équation d'Euler qui caractérise le choix du portefeuille comme une fonction de la variable prédictive est donnée par :

$$\alpha(z) = \left\{ \alpha/E \left[u' (R^f + \alpha \tilde{R}_{t+1}^e) \tilde{R}_{t+1}^e | z_t = z \right] = 0 \right\}, \quad (2.1)$$

où $u(W)$ est la fonction d'utilité.

L'équation (2.1) signifie que le choix optimal du portefeuille $\alpha(z)$ en un état donné z , c'est la décision α qui rend d'espérance de l'utilité marginale de l'investisseur égale à zero sachant que la variable prédictive observable $z_t = z$.

Remarque 2 (*Brandt (1999)*)

Si les opportunités d'investissement sont constants, alors le choix du portefeuille est indépendant de l'état z .

La remarque 2 implique que la règle de décision est constante, i.e., $\alpha(z) = \alpha$. Ainsi par la loi des espérances itérées, le choix du portefeuille est aussi identifié par l'espérance inconditionnelle de l'équation (2.1). On a :

$$\alpha_T = \left\{ \alpha/E \left[u' (R^f + \alpha \tilde{R}_{t+1}^e) \tilde{R}_{t+1}^e \right] = 0 \right\}. \quad (2.2)$$

En remplaçant l'espérance inconditionnelle de l'équation (2.2) par une moyenne empirique, on obtient un estimateur consistant de α donné par :

$$\alpha_T = \left\{ \alpha / \frac{1}{T} \sum_{t=1}^T u' (R^f + \alpha \tilde{R}_{t+1}^e) \tilde{R}_{t+1}^e = 0 \right\}. \quad (2.3)$$

La motivation économique de *Brandt (1999)* découle des preuves évidentes de la variabilité dans le temps des primes de risque et de la volatilité. Dans la littérature, certaines

études (exemples [Campbell et Viceira \(1999\)](#), [Brennan et collab. \(1997\)](#)) modélisent les moments conditionnels des rendements comme fonctions des variables prédictives et spécifient la dynamique des opportunités d'investissement. Ensuite, ils résolvent les décisions d'un investisseur en utilisant des estimations de la distribution conditionnelle implicite des rendements. Cependant, [Brandt \(1999\)](#) estime directement le portefeuille et le choix de consommation à partir des données en utilisant la méthode des moments généralisés.

Dans [Brandt \(1999\)](#), l'investisseur alloue une fraction de sa richesse à l'indice NYSE (New York Stock Exchange) et le reste au bon du Trésor 30-jours. Les décisions sont faites conditionnellement au rendement du dividende (dividend yield), la prime de défaut (default premium), la prime de durée (term premium) et le rendement excédentaire différé.

2.2.3 Maximisation de l'utilité espérée

[Brandt \(1999\)](#) considère que l'investisseur a un horizon fini et qu'il maximise l'utilité attendue de sa consommation tout au long de sa vie (i.e., lifetime consumption). Ainsi, un investisseur maximise l'espérance conditionnelle de :

$$u(C_0, C_1, \dots, C_T) = \sum_{t=1}^T \beta^t u(C_t), \quad (2.4)$$

avec,

$$u(C_t) = \begin{cases} \frac{C_t^{1-\gamma}}{1-\gamma} & \text{for } \gamma > 1, \\ \ln(C_t) & \text{for } \gamma = 1, \end{cases} \quad (2.5)$$

où,

- β : est le facteur d'actualisation et mesure la patience, c-à-d la volonté de renoncer à la consommation aujourd'hui pour consommer demain,
- γ : est le coefficient d'aversion au risque. Il capture la réticence à échanger la consommation pour un pari équitable sur la consommation actuelle.

À chaque période t , l'investisseur consomme une fraction q_t de la richesse W_t et alloue son épargne à deux classes d'actifs : un titre sans risque dont le rendement est $R^f \equiv \exp(r^f)$ et N titres risqués ayant des rendements excédentaires $\tilde{R}_{t+1}^e \equiv \exp(\tilde{r}_{t+1}) - \exp(r^f)$, où \tilde{r}_{t+1} et r^f sont des taux de rendements composés continus (i.e., $\tilde{r}_{t+1} = \ln\left(\frac{P_{t+1}}{P_t}\right)$).

En allouant des fractions α_t de son épargne aux titres risqués, l'investisseur est confronté à la contrainte budgétaire intertemporelle suivante :

$$\tilde{W}_{t+1} = (1 - q_t)W_t \tilde{R}_{t+1}^p, \quad (2.6)$$

où,

$$\tilde{R}_{t+1}^p \equiv R^f + \alpha_t' \tilde{R}_{t+1}^e, \quad (2.7)$$

est le rendement brute sur la richesse générée par le choix du portefeuille.

Remarque 3 (*Brandt (1999)*)

Pour les préférences avec CRRA (i.e., constant relative risk aversion), le choix du portefeuille et des règles de consommation sont fonctions de l'état de nature actuelle et de l'horizon.

Ainsi, étant donné un vecteur k -dimensionnel de variables prédictives z_t qui capturent l'état actuel, les règles de décisions sont :

$$\begin{cases} \alpha_t = \alpha(z_t, T - t), \\ q_t = q(z_t, T - t). \end{cases} \quad (2.8)$$

Brandt (1999) considère qu'il y'a aucun legs, donc l'investisseur utilise toute sa richesse à la fin de l'horizon. Ainsi,

$$\begin{cases} \alpha(z, 0) = 0, \\ q(z, 0) = 1. \end{cases} \quad (2.9)$$

À moins que l'investisseur soit myope (cas avec utilité logarithmique), le choix du portefeuille et de la consommation anticipe les changements futurs des opportunités d'investissement. Pour formaliser cette idée, [Brandt \(1999\)](#) écrit l'utilité indirecte de l'investisseur comme suit :

$$V(z_t, W_t, T - t) = \max_{\{\alpha_\tau, q_\tau\}_{\tau=t}^T} E \left[\sum_{\tau=t}^T \beta^{\tau-t} u(C_\tau) / z_t \right], \quad (2.10)$$

sujet à la contrainte budgétaire (2.6).

Lemme 1 *On montre que l'utilité indirecte devient :*

$$V(z_t, W_t, T - t) = \max_{\alpha_t, q_t} u(q_t W_t) + \beta E [V(z_{t+1}, W_{t+1}, T - (t + 1)) / z_t]. \quad (2.11)$$

Preuve : Voir Annexe A à la sous-section [2.3.2.2](#).

Remarque 4 *L'équation (2.11) permet d'interpréter l'utilité indirecte V comme étant l'utilité attendue du portefeuille actuel et futur ainsi que des choix de consommation compte tenu de l'état actuel z_t , de la richesse et de l'horizon.*

Une caractéristique importante de la fonction utilité CRRA u , est son homothétie (i.e., propriété de fonction homogène de degré 1). Par conséquent, l'utilité indirecte de l'investisseur peut s'écrire comme suit :

$$V(z_t, W_t, T - t) = \psi(z_t, T - t) u(W_t), \quad (2.12)$$

où,

$$\psi(z_t, T - t) = \begin{cases} q(z_t, T - t)^{-\gamma} & \text{for } \gamma > 1, \\ 1 & \text{for } \gamma = 1. \end{cases} \quad (2.13)$$

Notons que l'expression de ψ découle de la condition d'enveloppe $\frac{\partial V}{\partial W} = \frac{\partial u}{\partial C}$ appliquée aux équations (2.12) et (2.5).

Remarque 5 *L'utilité CRRA u et l'utilité indirecte V sont globalement concaves. Par conséquent, il existe une solution intérieure au portefeuille et choix de consommation.*

Théorème 1 *En appliquant conditions de premier ordre de maximisation de l'utilité espérée de l'investisseur, les règles de décision de l'investisseur sont caractérisées par les $N + 1$ équations d'Euler conditionnelles suivantes :*

$$\begin{bmatrix} \alpha(z_t, T-t) \\ q(z_t, T-t) \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ q \end{bmatrix} : E[m_{t+1}(\alpha, q, T-t) | z_t] = 0 \right\}, \quad (2.14)$$

où,

$$m_{t+1}(\alpha, q, T-t) = \begin{bmatrix} \psi(z_{t+1}, T-t-1) u'(\tilde{W}_{t+1}) \tilde{R}_{t+1}^e \\ \beta \psi(z_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p - 1 \end{bmatrix}, \quad (2.15)$$

avec \tilde{W}_{t+1} , \tilde{R}_{t+1}^p et ψ donnés aux équations (2.6), (2.7) et (2.13) respectivement.

Preuve : Voir Annexe B à la sous-section 2.3.2.2.

2.2.4 Estimateur inconditionnel GMM (Brandt (1999))

Ici, on suppose que les opportunités d'investissement sont constantes. Ce qui implique que le choix du portefeuille et consommation est indépendant des variables prédictives et de l'horizon. Ainsi, les paramètres $\{\alpha, q\}$ dans (2.14) pour lesquels l'espérance conditionnelle de m_{t+1} est nulle, sont aussi les paramètres pour lesquels l'espérance inconditionnelle de m_{t+1} est nulle. Par ailleurs, tant que les rendements sont stationnaires, le remplacement cette espérance inconditionnelle par une moyenne historique, permet d'obtenir un estimateur consistant donné par :

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{q}_T \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ q \end{bmatrix} : \frac{1}{T} \sum_{t=1}^T m_{t+1}(\alpha, q, \tau) = 0 \right\}. \quad (2.16)$$

L'application de l'estimateur inconditionnel GMM sur les données réels du marché nord-américain est présentée en section 2.3.1.

2.2.5 Estimateur conditionnel GMM (Brandt (1999))

L'idée de base de la méthode des moments conditionnels est que les paramètres estimés dépendent des opportunités d'investissement variant dans le temps. Pour mieux comprendre le principe de la méthode des moments conditionnels, Brandt (1999) propose de considérer le cas où l'évolution des opportunités d'investissement est discrète et le cas où cette évolution est continue.

2.2.5.1 Cas discret

Supposons que les opportunités d'investissement ne prennent qu'un seul des S états suivants : $z = \{z_1, z_2, \dots, z_S\}$. Les règles de décision du portefeuille et de consommation sont les S ensembles $\{\alpha(z_s, \tau), q(z_s, \tau)\}_{s=1}^S$ qui rendent nul les espérances conditionnelles $E[m_{t+1}|z = z_s]$. En remplaçant ces moments théoriques par les moyennes empiriques, chacune construite uniquement avec des rendements observés à l'état s , on obtient l'estimateur suivant :

$$\begin{bmatrix} \hat{\alpha}_T(z, \tau) \\ \hat{q}_T(z, \tau) \end{bmatrix} = \left\{ \begin{array}{l} \begin{bmatrix} \alpha \\ q \end{bmatrix} : \frac{1}{T_1} \sum_{\{t: z_t = z_1\}} m_{t+1}(\alpha, q, \tau) = 0 \quad \text{si } z = z_1, \\ \\ \begin{bmatrix} \alpha \\ q \end{bmatrix} : \frac{1}{T_2} \sum_{\{t: z_t = z_2\}} m_{t+1}(\alpha, q, \tau) = 0 \quad \text{si } z = z_2, \\ \\ \dots \\ \\ \begin{bmatrix} \alpha \\ q \end{bmatrix} : \frac{1}{T_S} \sum_{\{t: z_t = z_S\}} m_{t+1}(\alpha, q, \tau) = 0 \quad \text{si } z = z_S, \end{array} \right. \quad (2.17)$$

où T_s est le nombre de fois que l'état s est observé dans l'échantillon de taille $T = \sum_{s=1}^S T_s$. Cet exemple montre comment la méthode des moments conditionnels fonctionne dans le cas discret. À chaque état de la nature s , [Brandt \(1999\)](#) considère le choix de portefeuille et de consommation comme étant indépendant des autres et applique la méthode des moments traditionnelle donnée en (2.16), uniquement avec les rendements observés à l'état s . Cela permet d'obtenir les estimations consistantes du portefeuille et du choix de consommation à l'état s . Collectivement, les S estimations trouvées révèlent comment les règles de décisions dépendent des variables prédictives. Ainsi, la méthode des moments conditionnels donne une estimation ponctuelle, ou non-paramétrique, des règles de décision du portefeuille et du choix de consommation.

2.2.5.2 Cas continu

En principe, la même approche expliquée dans le cas discret s'applique lorsque l'évolution des opportunités d'investissement est continue. Selon [Brandt \(1999\)](#), le seul inconvénient est que pour chaque état z , l'échantillon fini qu'on dispose peut ne pas contenir

assez d'observations de l'état z pour calculer adéquatement une moyenne empirique de m_{t+1} avec des rendements observé uniquement en cet état de la nature z . [Brandt \(1999\)](#) propose une solution intuitive qui consiste à estimer le choix de portefeuille et de consommation à l'état z , pas seulement avec les rendements observés à l'état z , mais aussi avec les rendements observés aux états "similaires" à z . Dans cette sous-section nous présentons cette solution.

Définissons une fonction poids (nommée noyau ou kernel) donnée par $\omega\left(\frac{z-z_t}{h_T}\right)$ pour mesurer comment un état observé z_t est similaire à un état de référence z . Le paramètre h_T , appelé largeur de la bande (ou bandwidth), est strictement positif. Il permet de normaliser la distance entre z_t et l'état de référence z . Par conséquent, l'estimateur prend la forme suivante :

$$\begin{bmatrix} \hat{\alpha}_T(z, \tau) \\ \hat{q}_T(z, \tau) \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ q \end{bmatrix} : \frac{1}{Th_T^K} \sum_{t=1}^T \omega\left(\frac{z-z_t}{h_T}\right) m_{t+1}(\alpha, q, \tau) = 0 \right\}, \quad (2.18)$$

où le facteur Th_T^K assure que la moyenne pondérée de m_{t+1} soit non-dégérée lorsque $h_T \rightarrow 0$.

Une façon de comprendre l'estimateur est de penser à la fonction poids comme une fenêtre de données centrée en z et le paramètre h_T comme étant la largeur de la fenêtre. Un tel estimateur utilise une observation t seulement si z_t entre dans la fenêtre. Si l'observation z_t tombe hors de la fenêtre, alors l'état z_t est jugé trop "dis-similaire" à z et l'estimateur ignore cette observation dans le processus.

2.2.5.2.1 Choix du kernel et sélection de la largeur de la fenêtre

Suivant [Brandt \(1999\)](#), une bonne fonction de poids utilise normalement toutes les observations, mais place une emphase sur les données des états qui sont très similaires à l'état de référence. Une fonction de poids qui accomplit cela est le produit de K fonctions de densité de la loi normale centrée réduite :

$$\omega(u) = \prod_{i=1}^K k(u_i) = \sum_{i=1}^K \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u_i^2\right). \quad (2.19)$$

Notons que le produit des fonctions de densités de loi normale (voir équation (2.19)), n'est pas la seule fonction de poids qui donne des estimations cohérentes. Les fonctions de poids admissibles se présentent sous diverses formes. Leur caractéristique commune est que la moyenne pondérée de m_{t+1} converge uniformément vers l'espérance conditionnelle de m_{t+1} .

Bien que l'efficacité de l'estimateur dépende de la forme du kernel, des études de simulation pour d'autres estimateurs non-paramétriques montrent que la forme du kernel est beaucoup moins importante que le choix des largeurs des fenêtres en fonction de la taille de l'échantillon (voir [Härdle et Marron \(1985\)](#), [Silverman \(1986\)](#), et [Härdle \(1990\)](#) pour un résumé sur la littérature concernant les fonctions de poids optimaux pour les estimateurs de régression non-paramétrique).

Notons que l'utilisation d'un produit de densités normales standard comme choix du kernel est pratique car elle est symétrique, uniformément positive et efficace sur le plan calculatoire.

Toutefois, le choix des largeurs de fenêtres est le véritable obstacle à l'application de la méthode de moments conditionnels. Pour mieux comprendre l'importance du choix de la largeur de fenêtre, [Brandt \(1999\)](#) examine comment les propriétés de l'estimateur changent lorsqu'il fait varier les largeurs de la fenêtre. Il trouve que : d'une part, s'il augmente h_T , et par conséquent élargit le spectre de la fonction de poids, l'estimateur utilise plus de données dans l'ensemble, mais distingue moins les données des états différents. Cela diminue la variance de l'estimateur, mais augmente le biais potentiel en faisant la moyenne des observations qui sont moins "semblables". D'autre part, s'il diminue h_T , et donc réduit la dispersion de la fonction de poids, l'estimateur distingue davantage les données des états différents, mais utilise moins de données dans l'ensemble. Ceci diminue le biais potentiel de l'estimateur, mais augmente la variance. Donc, les largeurs de fenêtres optimales sont celles qui permettent d'atteindre un équilibre souhaitable entre le biais et

la variance.

Remarque 6 (*Brandt (1999)*)

Puisque h_T est interprétée comme la largeur de la fenêtre (window width) :

- *h_T très grand, signifie qu'une observation à l'état $z_t \neq z$ est moins pondérée à la baisse pour être différente de z .*

- *h_T très petit, signifie que l'observation à l'état $z_t \neq z$ est davantage pondérée à la baisse.*

- *Aux extrêmes, $h_T = \infty$ correspond à l'estimateur inconditionnel (voir équation (2.16)) et $h_T = 0$ correspond au cas discret (voir équation (2.17)).*

Interpréter h_T comme largeur de la fenêtre aide aussi à clarifier comment le paramètre optimal h_T^* varie avec la taille de l'échantillon et à travers les variables explicatives.

Remarque 7 (*Brandt (1999)*)

Un petit échantillon ou une variable de prévision très dispersée nécessite une fenêtre plus large qu'un grand échantillon ou une variable de prévision moins dispersée.

Remarque 8 (*Brandt (1999)*)

Comme toute méthode statistique, la méthode des moments conditionnels a ses avantages et ses inconvénients.

1. Ses avantages.

- *Un de ses grands avantages est qu'elle est non-paramétrique. Elle ne nécessite pas de formuler des hypothèses de distribution, comme par exemple la normalité.*

- *Elle peut permettre d'estimer les paramètres même si le modèle ne peut être résolu analytiquement à partir des conditions du premier ordre.*

- L'estimateur est moins biaisé que les estimateurs paramétriques mal spécifiés du fait que la distribution n'est pas connue.
- L'estimateur est au moins consistant.

2. Ses désavantages.

- Coût de perte d'observations dans l'estimation.
- La variance de l'estimateur conditionnel excède celle des estimateurs paramétriques correctement spécifiés avec la connaissance de la distribution.

Toutefois, h_T est déterminé afin d'obtenir un équilibre optimal de l'erreur quadratique moyenne entre le biais et la variance pour un échantillon de taille T .

À présent nous donnons quelques propriétés asymptotiques. Intuitivement selon [Brandt \(1999\)](#), la méthode des moments conditionnels conduit aux estimations consistantes des règles de décisions car la moyenne pondérée de m_{t+1} converge uniformément vers l'espérance conditionnelle de m_{t+1} pour tout $\{\alpha, q\}$, τ et z . Ce qui signifie que :

$$\frac{1}{Th_T^K} \sum_{t=1}^T \omega\left(\frac{z - z_t}{h_T}\right) m_{t+1}(\alpha, q, \tau) \xrightarrow{T \rightarrow \infty} E[m_{t+1}(\alpha, q, \tau) | z_t = z]. \quad (2.20)$$

Il s'en suit que les choix de portefeuille et de consommation qui mettent à zéro la régression non-paramétrique convergent vers ceux qui mettent à zéro l'espérance conditionnelle correspondante. On a les propriétés asymptotiques suivantes :

Propriété 1 : La fonction kernel $\omega(u)$ est un produit de K fonctions continues et bornées $k(u_i)$ qui satisfont :

$$\int_{\mathbb{R}} k(u) du = 1, \int_{\mathbb{R}} uk(u) du = 0, \text{ et } \int_{\mathbb{R}} u^2 k(u) du < +\infty. \quad (2.21)$$

Propriété 2 : La largeur de la fenêtre (bandwidth), notée h_T satisfait :

$$h_T \longrightarrow 0, Th_T^{K+4} \longrightarrow 0, \text{ et } Th_T^K \longrightarrow \infty \text{ quand } T \longrightarrow \infty. \quad (2.22)$$

Pour plus de détails sur les propriétés asymptotiques faites sur la fonction kernel, et sur la largeur de la fenêtre, le lecteur peut se référer à [Härdle et Marron \(1985\)](#), [Härdle \(1990\)](#), [Gourieroux et collab. \(1994\)](#) et [Brandt \(1999\)](#).

[Brandt \(1999\)](#) montre que sous certaines conditions de régularité, on a :

$$\sqrt{Th_T^K} \left(\begin{bmatrix} \hat{\alpha}_T(z, \tau) \\ \hat{q}_T(z, \tau) \end{bmatrix} - \begin{bmatrix} \alpha_T(z, \tau) \\ q_T(z, \tau) \end{bmatrix} \right) \longrightarrow \mathcal{N}(0, \Sigma(z, \tau)), \quad (2.23)$$

ce qui signifie la convergence en distribution vers un vecteur aléatoire de loi normale multivariée de moyenne 0 et de matrice de variance-covariance :

$$\Sigma(z, \tau) = \frac{D(z, \tau)^{-1} V(z, \tau) D(z, \tau)^{-1}}{f_z(z)} \int_{\mathbb{R}^K} \omega^2(u) du, \quad (2.24)$$

où,

$f_z(z)$ est la fonction de densité inconditionnelle de z ,

$$D(z, \tau) = E \left[\frac{\partial m_{t+1}(\alpha, q, \tau)}{\partial [\alpha', q]} \middle| z \right], \quad (2.25)$$

et,

$$V(z, \tau) = E \left[m_{t+1}(\alpha, q, \tau) m_{t+1}(\alpha, q, \tau)' \middle| z \right]. \quad (2.26)$$

Intuitivement, les éléments dans la matrice D mesurent la sensibilité d'un moment particulier par rapport aux petites variations des paramètres appartenant $\{\alpha, q\}$. Si la sensibilité par rapport à un paramètre donné de $\{\alpha, q\}$ est grande, de petits changements dans ce paramètre entraîneront des changements relativement importants dans la fonction objectif et le paramètre en question est estimé de manière relativement précise.

La propriété asymptotique donnée en (2.23) permet d'obtenir alternativement une interprétation pertinente de l'estimateur conditionnel. Supposons qu'on observe au temps $t + 1$ une mesure du bruit du choix de portefeuille et de consommation au temps t . Soit,

$$\begin{bmatrix} \tilde{\alpha}_{t+1} \\ \tilde{q}_{t+1} \end{bmatrix} = \begin{bmatrix} \alpha(z_t, \tau) \\ q(z_t, \tau) \end{bmatrix} + \bar{v}_{t+1}, \quad (2.27)$$

où,

$$v = D(z_t, \tau)^{-1} m_{t+1}(\alpha(z_t, \tau), q(z_t, \tau), \tau) \quad (2.28)$$

est le terme d'erreur de moyenne nulle.

Notons que la métrique conventionnelle pour le trade-off (l'équilibre) entre biais et variance consiste à minimiser l'erreur quadratique moyenne (EQM ou MSE : mean squared error) des estimations. Partant de l'équation (2.27) qui est une interprétation de la régression transformée de l'estimateur conditionnel, Brandt (1999) argumente que les largeurs de fenêtres qui minimisent la moyenne des EQM des estimations sont de la forme suivante :

$$h_T = \lambda \sigma(z) T^{-(1/(K+4))}, \quad (2.29)$$

où, $\sigma(z)$ est le vecteur des écart-types inconditionnels des variables de prédictives. Le paramètre λ dépend de $\Sigma(z, \tau)$ qui dépend également de $f_z(z)$. Cependant, aucune de ces quantités n'est connue.

En pratique, il existe plusieurs façons de choisir λ . Brandt (1999) utilise la validation croisée de type "leave-one-out", nommée LOOCV (voir sous-section 1.3).

Remarque 9 (Brandt (1999))

Puisque l'observation $\{\alpha_{t+1}, q_{t+1}\}$ n'est pas connue, l'erreur prédictive ne peut être construite à partir de (1.3).

Théorème 2 *L'erreur prédictive peut alternativement être construite par :*

$$v = \hat{D}(z_t, \tau)^{-1} \bar{m}_{t+1}, \quad (2.30)$$

où, \bar{m}_{t+1} est donnée par :

$$\bar{m}_{t+1} = \frac{1}{h_T^K} \omega\left(\frac{z - z_t}{h_T}\right) m_{t+1}, \quad (2.31)$$

et

$$\hat{D} = [\hat{D}_1 \ \hat{D}_2; \ \hat{D}_3 \ \hat{D}_4], \quad (2.32)$$

est une estimation de la matrice $D(z, \tau)$, avec

$$\begin{cases} \hat{D}_1 &= E \left[-\gamma(1-q)^{-\gamma} \times (\tilde{R}_{t+1}^e)^2 \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma-1} \times N(z_t, \tau) | z \right], \\ \hat{D}_2 &= E \left[\gamma(1-q)^{-\gamma-1} \times \tilde{R}_{t+1}^e \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma} \times N(z_t, \tau) | z \right], \\ \hat{D}_3 &= E \left[\beta(1-\gamma) \times \left(\frac{1-q}{q}\right)^{-\gamma} \times \tilde{R}_{t+1}^e \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma} \times N(z_t, \tau) | z \right], \\ \hat{D}_4 &= E \left[\frac{\beta\gamma}{q(1-q)} \times \left(\frac{1-q}{q}\right)^{-\gamma} \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{1-\gamma} \times N(z_t, \tau) | z \right], \end{cases} \quad (2.33)$$

où,

$$N(z_t, \tau) = \frac{1}{h_T^K} \omega\left(\frac{z - z_t}{h_T}\right), \quad (2.34)$$

et,

$$m_{t+1}(\alpha, q, T-t) = \begin{bmatrix} \psi(\tilde{z}_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(W_t)} \tilde{R}_{t+1}^e \\ \beta \psi(\tilde{z}_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p - 1 \end{bmatrix}. \quad (2.35)$$

Preuve : Voir Annexe C à la sous-section 2.3.2.2.

Sachant que

$$m_{t+1}(\alpha, q, 1) = \begin{bmatrix} m_{t+1}^1(\alpha, q, 1) \\ m_{t+1}^2(\alpha, q, 1) \end{bmatrix} = \begin{bmatrix} \frac{u'(\tilde{W}_{t+1})}{u'(W_t)} \tilde{R}_{t+1}^e \\ \beta \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p - 1 \end{bmatrix}, \quad (2.36)$$

il est facile de montrer que :

$$v_{t+1} = \frac{N(z_t, \tau)}{\det(\hat{D})} \begin{pmatrix} \hat{D}_4 m^1 - \hat{D}_3 m^2 \\ \hat{D}_1 m^2 - \hat{D}_2 m^1 \end{pmatrix}. \quad (2.37)$$

2.3 Résultats empiriques

2.3.1 Illustrations : Moments inconditionnels (Brandt (1999))

2.3.1.1 Résultats sur les données de développement

Dans cette sous-section, on se propose d'appliquer notre technique d'estimation GMM pour répliquer les résultats de la table III de Brandt (1999). L'objectif est de trouver les règles de décisions pour un portefeuille à une seule période. C'est-à-dire trouver la fraction de l'épargne allouée aux actions (i.e., choix du portefeuille), ainsi que la fraction de richesse consommée (i.e., le choix de consommation). Brandt (1999) suppose que l'horizon et les périodes de re-balancement du portefeuille varie entre 1 mois et 4 ans. Nos résultats sont présentés dans le Tableau 2.3 suivant :

TABLE 2.3 – Estimateur GMM : Choix du portefeuille et consommation en fonction de l’horizon et l’aversion au risque (Réplication Table III de [Brandt \(1999\)](#))

Aversion au risque	Horizon					
	1 ^{er} –mois	3 ^e –mois	6 ^e –mois	1 ^{ère} –année	2 ^e –année	4 ^e –année
Choix du portefeuille						
$\gamma = 1$	3.4033	2.9657	2.7267	2.1830	2.0291	3.4637
$\gamma = 2$	1.9742	1.8459	1.6946	1.5510	1.5045	2.2571
$\gamma = 5$	0.8226	0.7875	0.7081	0.6606	0.6509	0.8422
$\gamma = 10$	0.4134	0.3970	0.3534	0.3249	0.3000	0.3510
Choix de consommation						
$\gamma = 1$	0.5025	0.5025	0.5025	0.5025	0.5025	0.5025
$\gamma = 2$	0.5021	0.5037	0.5060	0.5104	0.5198	0.5458
$\gamma = 5$	0.5012	0.5023	0.5040	0.5073	0.5139	0.5292
$\gamma = 10$	0.5007	0.5014	0.5025	0.5044	0.5079	0.5148

Les résultats du Tableau III de [Brandt \(1999\)](#) sont présentés à la figure 2.6 suivante :

FIGURE 2.6 – (source Brandt (1999))

Table III
Single-Period Portfolio and Consumption Choice
as a Function of the Horizon

This table shows method of moments estimates of the unconditional single-period portfolio choice α and consumption choice c of an investor with constant relative risk aversion γ . Both the horizon and rebalancing period range from one month to four years. The intertemporal discount factor is $\beta = 0.99$. The estimates in Panel A represent the fraction of savings allocated to equities. The estimates in Panel B represent the fraction of wealth consumed. The method of moments estimator $\{\hat{\alpha}_T, \hat{c}_T\}$ of $\{\alpha, c\}$ is:

$$\begin{bmatrix} \hat{\alpha}_T \\ \hat{c}_T \end{bmatrix} = \left\{ \begin{bmatrix} \alpha \\ c \end{bmatrix} : \frac{1}{T} \sum_{t=1}^T m_{t+1}(\alpha, c, 1) = 0 \right\},$$

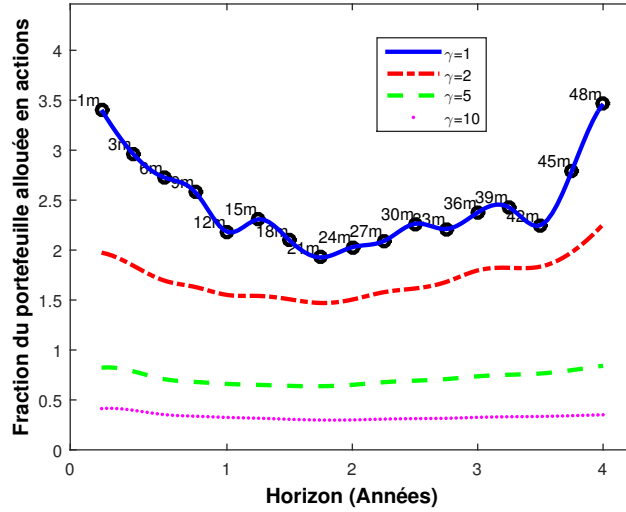
where m_{t+1} is the investor's marginal utility, given a portfolio choice α , a consumption choice c , and a horizon of one rebalancing period. Serial-correlation corrected asymptotic standard errors are in parentheses. Note that the consumption choice of an investor with logarithmic utility is deterministic: $c = 1/(1 + \beta)$.

Risk Aversion	Horizon					
	1-Month	3-Month	6-Month	1-Year	2-Year	4-Year
Panel A: Portfolio Choice						
$\gamma = 1$	3.45 (0.83)	2.88 (0.72)	2.73 (0.64)	2.21 (0.49)	2.08 (0.45)	3.09 (0.63)
$\gamma = 2$	1.93 (0.45)	1.70 (0.39)	1.69 (0.36)	1.57 (0.29)	1.56 (0.25)	2.41 (0.36)
$\gamma = 5$	0.79 (0.20)	0.72 (0.17)	0.71 (0.16)	0.69 (0.15)	0.73 (0.17)	1.02 (0.24)
$\gamma = 10$	0.40 (0.10)	0.37 (0.09)	0.36 (0.08)	0.35 (0.07)	0.37 (0.09)	0.49 (0.12)
Panel B: Consumption Choice						
$\gamma = 1$	0.50	0.51	0.52	0.53	0.56	0.62
$\gamma = 2$	na (0.09)	na (0.08)	na (0.08)	na (0.07)	na (0.06)	na (0.07)
$\gamma = 5$	0.50 (0.11)	0.50 (0.10)	0.51 (0.10)	0.51 (0.09)	0.53 (0.08)	0.56 (0.09)
$\gamma = 10$	0.50 (0.11)	0.50 (0.10)	0.50 (0.10)	0.51 (0.09)	0.52 (0.08)	0.54 (0.09)

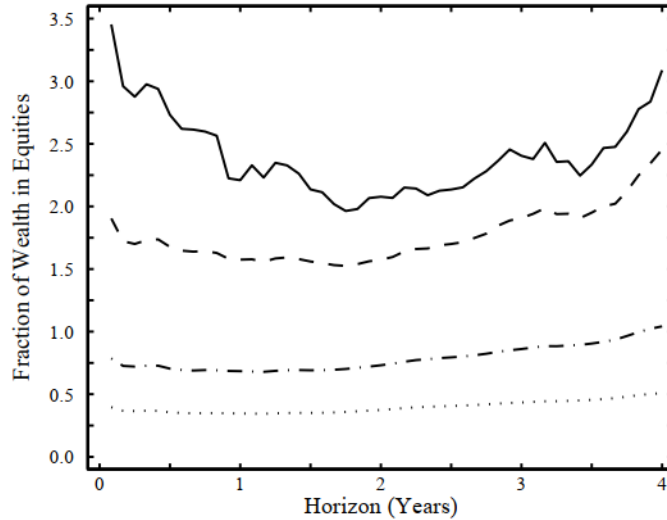
On remarque que les résultats du Tableau 2.3 sont similaires aux résultats du Tableau III de Brandt (1999), notamment pour la partie relative au choix du portefeuille (voir aussi Figure 2.7). Cependant, on note que pour le choix de consommation, notre estimateur améliore les résultats de la table III de Brandt (1999), puisque pour $\gamma = 1$, il permet d'obtenir la valeur exacte du choix de consommation, soit $c = \frac{1}{1 + \beta} = 0.5025$, avec $\beta = 0.99$.

FIGURE 2.7 – Estimateur GMM vs Figure 1 Brandt (1999). Illustrations de l'évolution du choix du portefeuille en fonction de l'horizon et l'aversion au risque

(a) Illustrations par l'estimateur GMM



(b) Illustrations Figure 1 Brandt (1999)



Les graphiques (a) et (b) de la Figure 2.7 présentent nos résultats et ceux de Brandt (1999) respectivement. Ils illustrent l'évolution du choix du portefeuille en fonction de l'horizon et l'aversion au risque. On remarque que les trajectoires du choix du portefeuille obtenues par notre estimateur GMM sont similaires aux trajectoires de la Figure 1 de Brandt (1999).

Toutefois, la Figure 2.7 montre que l'évolution du choix de portefeuille dépend de l'horizon. Par exemple, un investisseur avec une aversion au risque $\gamma = 5$ choisit un portefeuille de 82% d'actions si l'horizon est de 1-mois. Le même investisseur choisit un portefeuille de 66% d'actions si son horizon est de 1-an et un portefeuille de 84% d'actions si l'horizon est de 4-an. Par ailleurs, on remarque qu'après une décroissance initiale, la trajectoire du choix du portefeuille croît avec l'horizon. Ceci est consistant avec les régressions de Fama et French (1988b) et les ratios de variance de Poterba et Summers (1988). Les deux études ont montré que les rendements long-terme sont sériellement négativement corrélés, ce qui les rend moins risqués que les rendements mensuels (Brandt (1999)).

2.3.1.2 Résultats empiriques sur les données post-développement

Le Tableau 2.4 présente les résultats sur la période post-développement (i.e., janvier 1997 à décembre 2021) du modèle de Brandt (1999).

TABLE 2.4 – Estimateur GMM : Choix du portefeuille et consommation en fonction de l'horizon et l'aversion au risque (post-développement)

Aversion au risque	Horizon					
	1 ^{er} -mois	3 ^e -mois	6 ^e -mois	1 ^{ère} -année	2 ^e -année	4 ^e -année
Choix du portefeuille						
$\gamma = 1$	0.2543	2.4980	1.9972	1.9402	1.9131	2.8848
$\gamma = 2$	0.2439	1.5442	1.3018	1.2338	1.2296	1.7817
$\gamma = 5$	0.2486	0.6937	0.6079	0.5566	0.5468	0.7385
$\gamma = 10$	0.2491	0.3752	0.3386	0.3003	0.2970	0.3990
Choix de consommation						
$\gamma = 1$	0.5025	0.5025	0.5025	0.5025	0.5025	0.5025
$\gamma = 2$	0.1637	0.5029	0.5044	0.5076	0.5137	0.5335
$\gamma = 5$	0.0016	0.5016	0.5026	0.5047	0.5088	0.5215
$\gamma = 10$	0.8499	0.5008	0.5013	0.5024	0.5047	0.5117

La figure 2.8 illustre l'évolution de la fraction de portefeuille à allouer en actions en fonction de l'horizon pour la période post-développement.

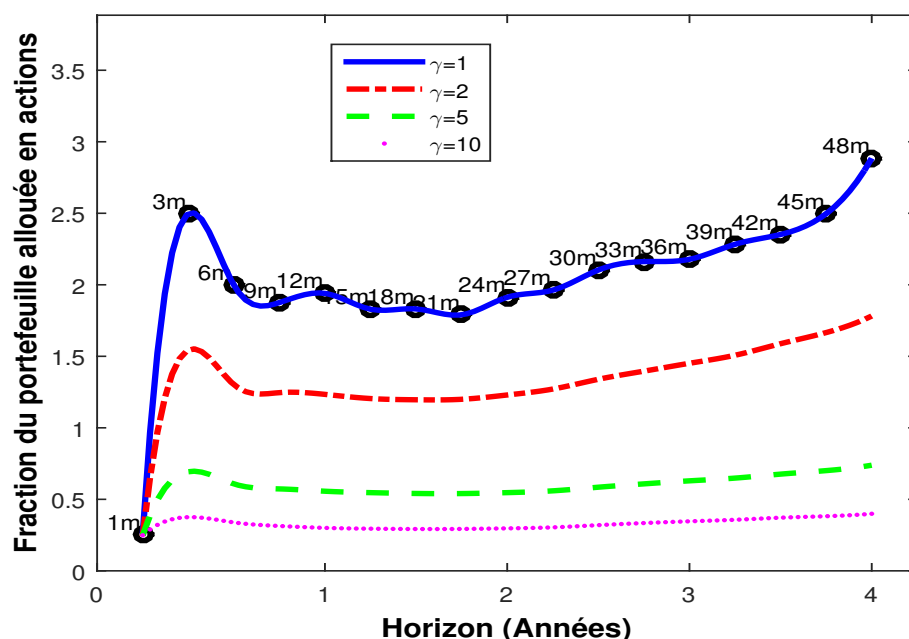


FIGURE 2.8 – Évolution du poids optimal du portefeuille à allouer en actions en fonction de l'horizon sur la période post-développement du modèle de Brandt (1999).

En comparant la figure 2.6 (échantillon de développement allant de janvier 1947 à décembre 1996) à la figure 2.8 (échantillon post-développement allant de janvier 1997 à décembre 2021), on remarque que pour les horizons supérieurs à 3 mois, les trajectoires des deux figures sont presque similaires, tandis que pour les horizons court-termes (inférieur à 3 mois) les trajectoires changent passant de décroissante à croissante respectivement. Cependant, on note que sur les données post-développement, les valeurs du choix optimal du portefeuille à allouer en actions sont en général inférieures à celles des données de développement. Par exemple, le même investisseur ayant un horizon d'un mois avec une aversion au risque $\gamma = 5$, choisit un portefeuille de 82% d'actions avec l'échantillon de développement et un portefeuille de 25% d'actions avec l'échantillon post-développement. De plus, avec l'échantillon post-développement, quelque soit son aversion au risque, l'investisseur choisit un portefeuille d'environ 25% d'actions lorsque l'horizon est d'un mois. Ce qui montre un changement de règles de décision par rapport aux données de développement, car en considérant l'échantillon post-développement, tout investisseur ayant un

horizon d'un mois reste prudent en investissant uniquement $\frac{1}{4}$ du portefeuille en actions. La politique prudentielle observée chez l'investisseur pour l'horizon 1 mois est probablement due la volatilité des rendements mensuels comparativement aux rendements long-terme, marquée également par les évènements inhabituels de ces 50 dernières années, notamment la crise financière mondiale de 2008-2009 et tout récemment la pandémie de Covid-19 et ses effets collatéraux sur le tissu économique mondial.

2.3.2 Illustrations : Moments conditionnels (Brandt (1999))

2.3.2.1 Résultats sur les données de développement

Le Tableau 2.5 donne les estimations conditionnelles du choix de portefeuille d'un investisseur ayant une aversion au risque constant. Ici l'horizon est 1 mois ainsi que la période de re-balancement du portefeuille. Les estimations représentent la fraction de l'épargne allouée aux actions lorsque la variable de prévision z est égale à ses 25^e, 50^e et 75^e percentiles. L'estimateur $\hat{\alpha}_T$ de α par la méthode des moments conditionnels est donné par :

$$\hat{\alpha}_T = \left\{ \alpha : \frac{1}{Th_T} \sum_{t=1}^T \omega\left(\frac{z - z_t}{h_T}\right) m_{t+1}(\alpha, 0.5, 1) = 0 \right\}. \quad (2.38)$$

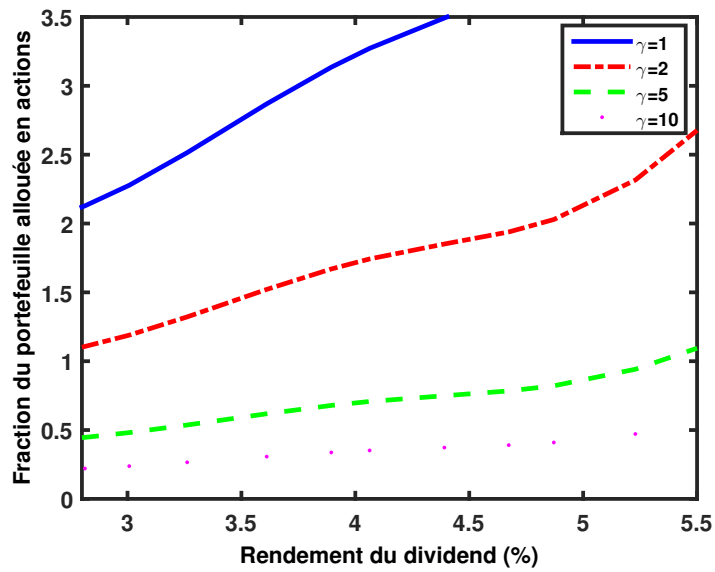
TABLE 2.5 – Estimateur conditionnel GMM : Choix du portefeuille en fonction des variables prédictives (Réplication Table V de [Brandt \(1999\)](#))

Aversion au risque	Percentile de z				Percentile de z			
	25%	50%	75%	\bar{z}	25%	50%	75%	\bar{z}
	z : Dividend yield				z : Default premium			
$\gamma = 1$	2.28	3.00	3.68	3.18	3.34	2.97	2.62	2.68
$\gamma = 2$	1.19	1.60	1.94	1.70	1.77	1.57	1.39	1.41
$\gamma = 5$	0.48	0.65	0.78	0.69	0.72	0.64	0.57	0.58
$\gamma = 10$	0.24	0.32	0.39	0.34	0.36	0.32	0.29	0.29
	z : Term premium				z : Lagged Excess return			
$\gamma = 1$	3.73	4.00	3.99	4.02	3.24	3.56	3.82	3.53
$\gamma = 2$	2.00	2.33	1.89	2.39	1.77	1.89	2.19	1.97
$\gamma = 5$	0.81	0.97	1.09	1.00	0.73	0.82	0.91	0.81
$\gamma = 10$	0.41	0.49	0.55	0.5	0.36	0.41	0.46	0.41

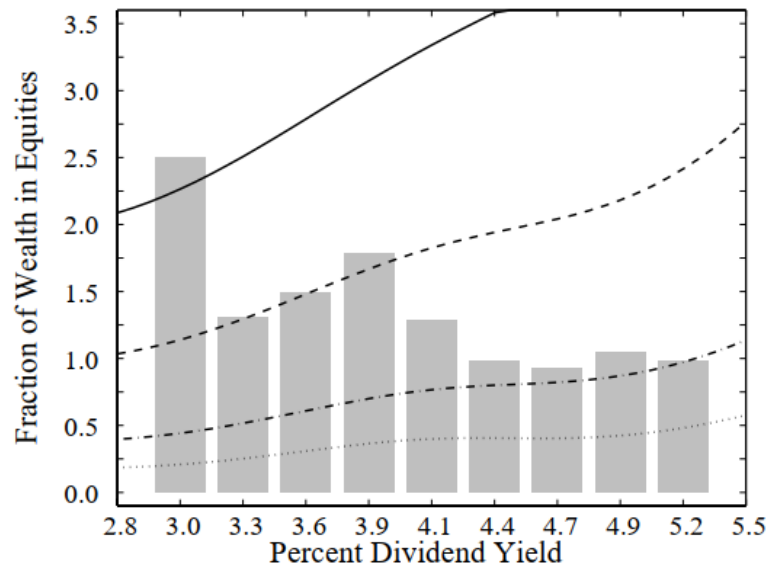
Le Tableau 2.5 résume les résultats des Figures 2.9–2.12. Il donne le choix de portefeuille au 25^e, 50^e et 75^e percentiles de chaque variable prédictive. Il donne aussi le choix du portefeuille moyen sur la période de l'échantillon. Les résultats confirment la conclusion visuelle des Figures 2.9–2.12. L'investisseur choisit des portefeuilles sensiblement différents en fonction de la réalisation de chaque variable de prévision.

Comme à la sous-section 2.3.1, considérons un investisseur avec une aversion au risque $\gamma = 5$. Pour un rendement du dividende de 4.7% au 75e percentile, l'investisseur détient un portefeuille composé de 78% d'actions, contre seulement 48% avec un rendement du dividende de 3.0% au 25e percentile. Typiquement, on remarque que le choix du portefeuille est aussi sensible aux autres variables ([Brandt \(1999\)](#)). Les différences entre le pourcentage d'actions détenues aux 75e et 25e percentiles de la prime de défaut, de la prime de durée, et du rendement excédentaire sont respectivement de 15, 28 et 18% de l'épargne. Les Figures 2.9–2.12 permettent d'observer que les décisions de choix de portefeuille sont toujours approximativement linéaires pour le rendement du dividende et le rendement excédentaire retardé que pour la prime de défaut et la prime de durée ([Brandt \(1999\)](#)). Cependant, pour la prime de défaut et la prime de durée, le degré de non-linéarité

des courbes diminue lorsque l'aversion au risque γ croit.

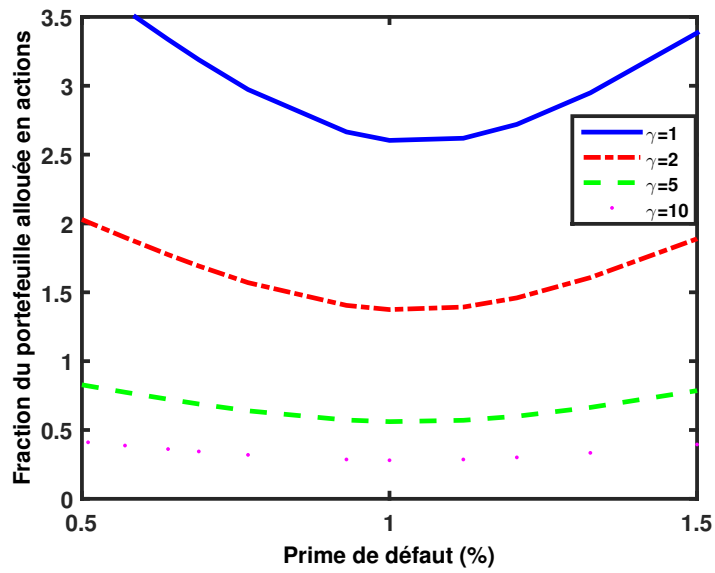


(a) Estimateur GMM conditionnel

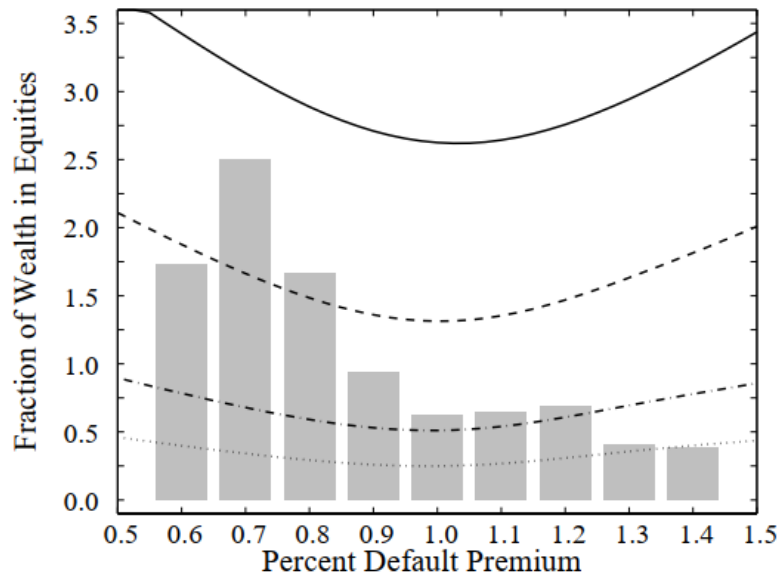


(b) Figure 2a Brandt (1999)

FIGURE 2.9 – Réplication de la figure 2a Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est le rendement du dividende (dividend yield). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).

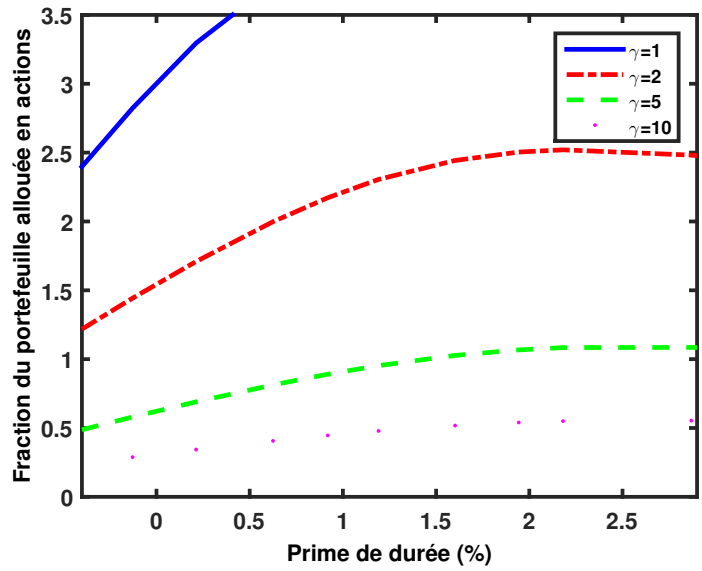


(a) Estimateur GMM conditionnel

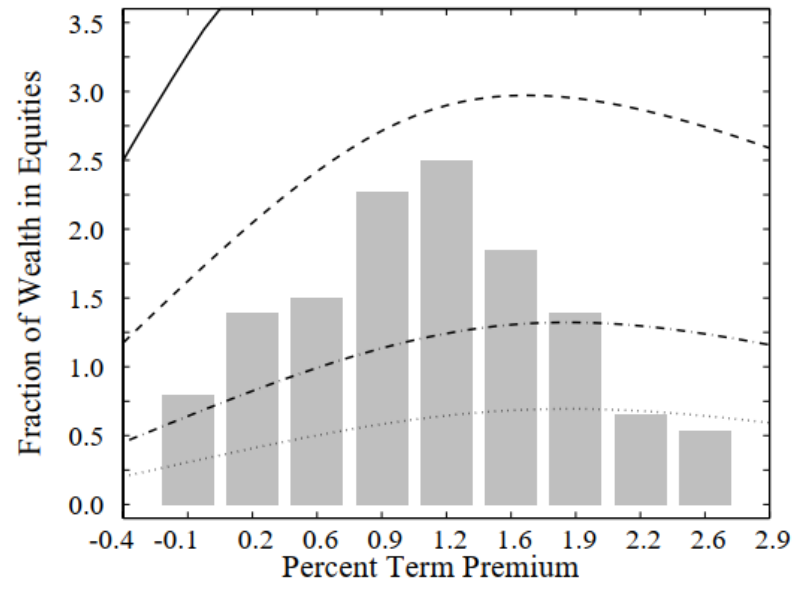


(b) Figure 2b Brandt (1999)

FIGURE 2.10 – Réplication de la figure 2b Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est la prime de défaut (default premium). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).

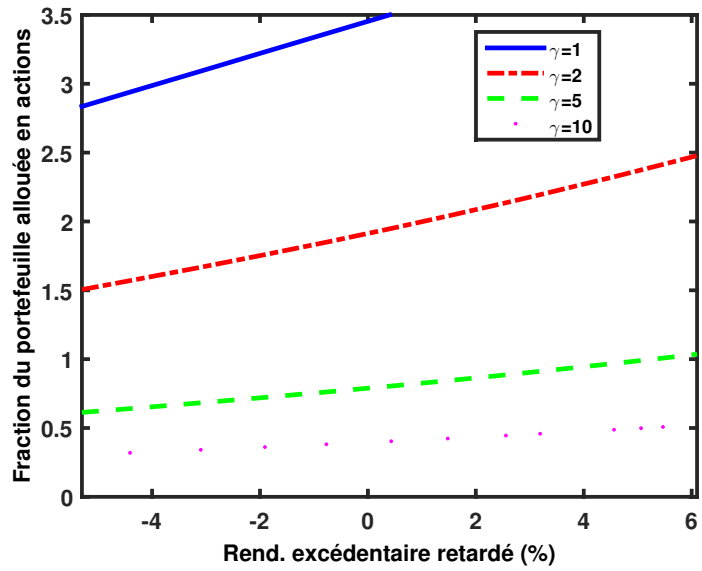


(a) Estimateur GMM conditionnel

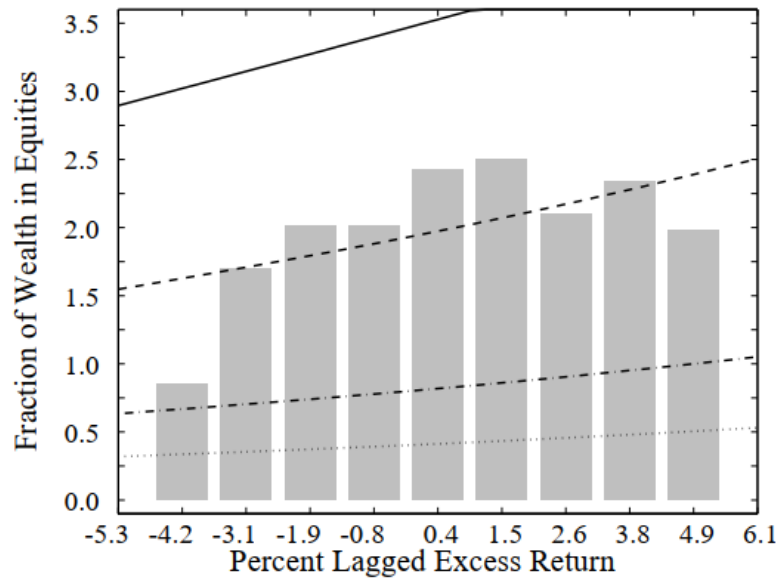


(b) Figure 2c Brandt (1999)

FIGURE 2.11 – Réplication de la figure 2c Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est la prime de durée (term premium). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).



(a) Estimateur GMM conditionnel



(b) Figure 2d Brandt (1999)

FIGURE 2.12 – Réplication de la figure 2d Brandt (1999). Cette figure illustre les estimations du choix du portefeuille lorsque la variable prédictive est le rendement excédentaire retardé (lagged excess return). La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de Brandt (1999).

En résumé, sur les données de développement, les Figures 2.9–2.12 montrent que nos

résultats empiriques sont similaires à ceux de Brandt (1999). Par conséquent, les conclusions sont les mêmes que celles de Brandt (1999). Notamment, que les règles de décision ne sont pas constantes, mais qu'elles changent en fonction des états des variables prédictives. L'investisseur choisit des portefeuilles substantiellement différents en fonction de la réalisation de chaque variable de prévision.

2.3.2.2 Résultats empiriques sur les données post-développement

Le tableau 2.6 présente les résultats des choix optimaux du portefeuille en fonctions des variables prédictives : le rendement du dividende, la prime de défaut, l'écart de taux d'intérêt, et le rendement excédentaire retardé. La période utilisée va de janvier 1997 - décembre 2021 (soit 300 observations mensuelles). L'objectif est de voir si les résultats obtenus lors du développement du modèle (voir tableau 2.5) fonctionnent sur la période post-développement.

TABLE 2.6 – Estimateur conditionnel GMM : Choix du portefeuille en fonction des variables prédictives (période d'observation : janvier 1997 - décembre 2021)

Aversion au risque	Percentile de z				Percentile de z			
	25%	50%	75%	\bar{z}	25%	50%	75%	\bar{z}
	z : Dividend yield				z : Default premium			
$\gamma = 1$	3.0769	3.2190	3.2735	3.2236	3.9027	3.7966	3.5958	3.7173
$\gamma = 2$	1.6365	1.7358	1.7862	1.7393	1.5858	2.0598	1.9361	2.0110
$\gamma = 5$	0.6759	0.7220	0.7482	0.7242	0.8829	0.8537	0.8010	0.8326
$\gamma = 10$	0.3437	0.3688	0.3838	0.3699	0.4479	0.4332	0.4062	0.4223
	z : Term premium				z : Lagged Excess return			
$\gamma = 1$	3.1282	3.0709	2.9383	3.0621	2.7316	3.2184	3.6330	3.1465
$\gamma = 2$	1.6726	1.6373	1.5858	1.6323	1.4359	1.7033	1.9363	1.6635
$\gamma = 5$	0.6928	0.6779	0.6476	0.6760	0.5905	0.7008	0.7972	0.6841
$\gamma = 10$	0.3528	0.3455	0.3312	0.3446	0.3003	0.3553	0.4037	0.3471

En comparant les tableaux 2.5 et 2.6 on remarque que lorsque les variables prédictives "dividend yield" et "default spread" sont utilisées, les poids optimaux des actions dans le

portfeuille augmentent comparativement aux résultats obtenus pendant le développement du modèle. Toutefois, pour les variables "term spread" et "lagged excess return", les valeurs des poids alloués en actions diminuent sur la période post-publication comparativement à la période de développement. Ainsi, en utilisant les données post-publication, l'investisseur mettrait plus de poids en actions si les variables dividend yield et default spread sont considérées comme variables prédictives. Par ailleurs, l'inverse est constaté lorsque ce sont les variables "term spread" et "lagged excess return" qui sont utilisées.

Les figures 2.13–2.16 illustrent les résultats du choix optimal du portefeuille en fonction des états des variables prédictives utilisées. Comparativement aux figures 2.9–2.12, on remarque que sur toutes ces figures, seul le pattern du choix de portefeuille lorsque la variable rendement excédentaire retardé est utilisé reste le même, c'est-à-dire linéairement croissant. Cela confirme d'une part que le choix du portefeuille semble sensible aux variables de prévision, et d'autre part que l'information sur les rendements retardés ne change presque pas le portrait visuel de l'évolution des décisions du portefeuille.

Pour la variable rendement du dividende (figure 2.13), comparativement à la figure 2.9a la forme des courbes est passée de quasiment linéaire à presque aplatie. Cela révèle que les valeurs des poids du choix de portefeuille ont augmenté et ont probablement atteint un seuil maximal. Toutefois, comme pour la figure 2.9, les résultats sont raisonnables et les valeurs comparables à celles observées dans d'autres analyses similaires dans la littérature (Barberis (2000) et Campbell et Viceira (1999)). Par exemple, les figures de Barberis (2000), montrent lorsque la variable prédictive considérée est rendement du dividende, un investisseur d'horizon 1 an avec une aversion au risque $\gamma = 5$ détient entre 10% et 100% d'actions lorsque la variable prédictive considérée est le rendement du dividende. Aussi, les résultats de Campbell et Viceira (1999) impliquent que lorsque la variable prédictive considérée est rendement du dividende, un investisseur de durée de vie infinie avec une aversion au risque $\gamma = 4$ détient un portefeuille avec plus de 100% d'actions.

Pour la variable "default spread" les courbes sont passées de la forme convexe à approximativement concave, tandis que pour la variable "term spread", les courbes sont passées de non-linéaire à approximativement linéaire. Par ailleurs, notons que les courbes

des figures 2.13 et 2.15 sont décroissantes.

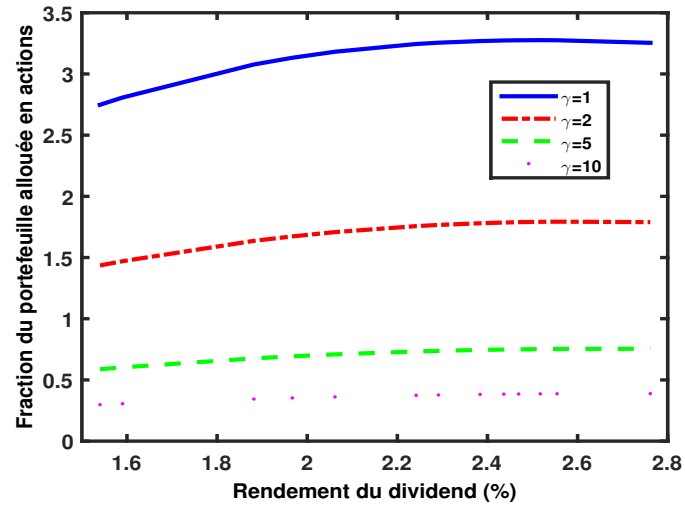


FIGURE 2.13 – Estimation du choix du portefeuille lorsque la variable prédictive est le rendement du dividende. Les données vont de janvier 1997 à décembre 2021 (période post-développement).

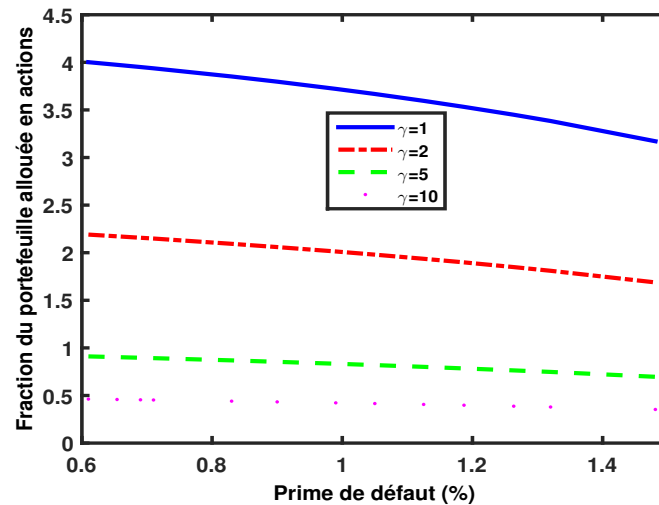


FIGURE 2.14 – Estimation du choix du portefeuille lorsque la variable prédictive est la prime de défaut. Les données vont de janvier 1997 à décembre 2021.

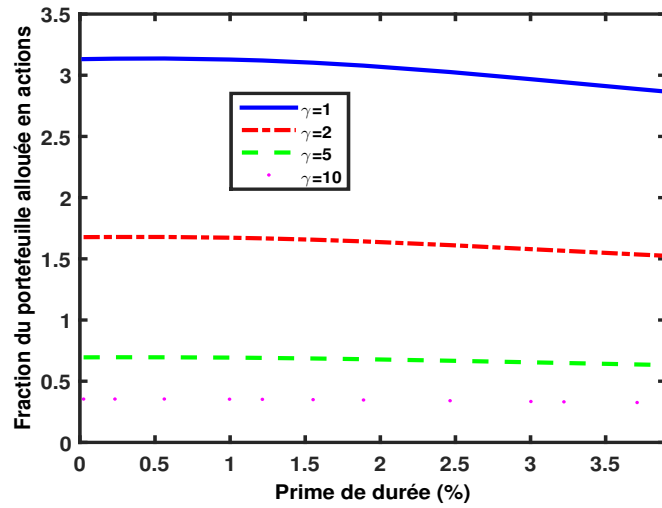


FIGURE 2.15 – Estimation du choix du portefeuille lorsque la variable prédictive est l'écart de taux. Les données vont de janvier 1997 à décembre 2021.

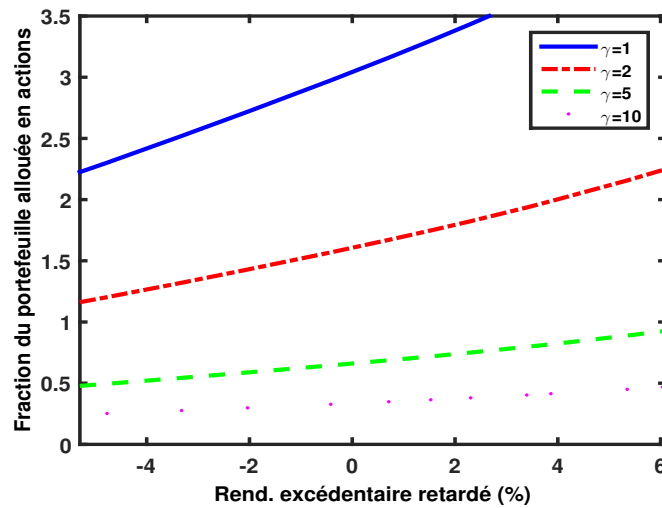
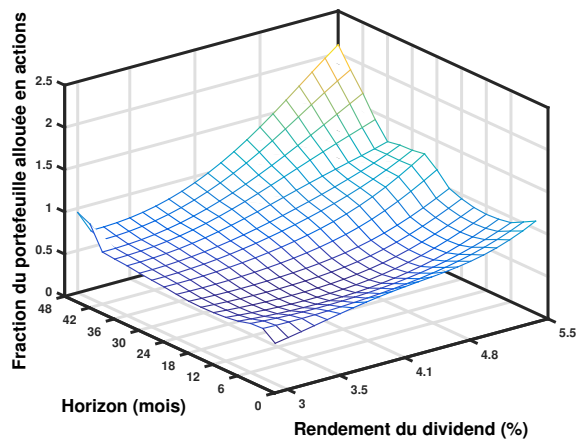


FIGURE 2.16 – Estimation du choix du portefeuille lorsque la variable prédictive est le rendement excédentaire retardé. Les données vont de janvier 1997 à décembre 2021.

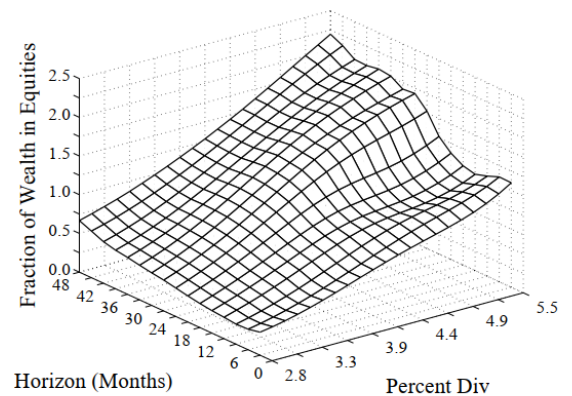
Depuis des décennies, la relation entre la variable 'dividend yield' et le rendement espéré intéresse les économistes en finance (Blume (1980), Fama et French (1988a), Hodrick (1992), Brandt (1999), et Barberis (2000)). Ces études ont documenté l'évidence selon laquelle les opportunités d'investissement disponibles évoluent dans le temps en fonction de l'évolution du rendement du dividende. Par exemple, suivant Barberis (2000),

lorsque le rendement du dividende diminue, les rendements attendus diminuent. En outre, [Fama et French \(1988a\)](#) montrent que le coefficient de pente et le R^2 d'une régression des rendements sur les rendements en dividendes augmentent de manière significative avec l'horizon de rendement. Les résultats de [Brandt \(1999\)](#) sont consistants avec ces derniers, puisqu'il montre que l'augmentation du choix de portefeuille pour un rendement de dividende donné est non-linéaire pour certains rendements du dividende, en fonction de l'horizon (voir figure 2.17b).

Les figures 2.17a et 2.18 présentent les résultats de l'évolution des choix de portefeuilles en fonction de l'horizon et des états de la variable 'dividend yield' tant pour les données de développement du modèle de [Brandt \(1999\)](#) que pour les données post-développement pour un investisseur ayant une aversion au risque $\gamma = 5$. Les résultats de la figure 2.17 montrent que l'augmentation du choix de portefeuille pour un rendement de dividende donné est non linéaire, voire non monotone pour certains rendements de dividende, à l'horizon. Ce résultat diffère de la relation linéaire rapportée par [Barberis \(2000\)](#).



(a) Estimateur GMM conditionnel



(b) Figure 4a [Brandt \(1999\)](#)

FIGURE 2.17 – Données de développement : Illustration de l'évolution du choix du portefeuille en fonction de l'horizon et du rendement du dividende pour un investisseur d'aversion au risque $\gamma = 5$. La figure (a) est obtenue avec notre estimateur GMM conditionnel, et la figure (b) est celle de [Brandt \(1999\)](#).

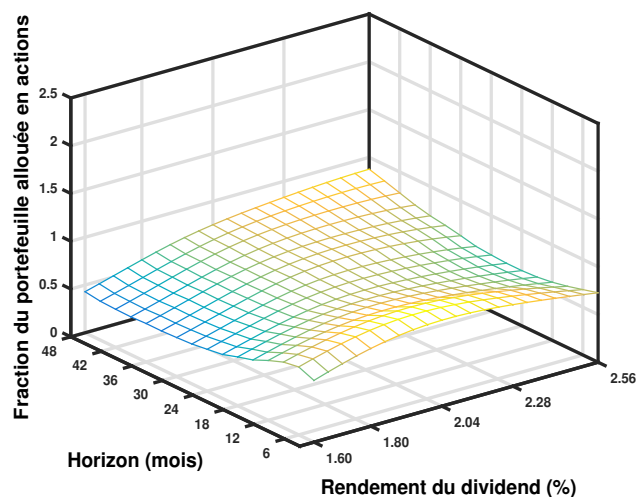


FIGURE 2.18 – Données de post-développement : Illustration de l'évolution du choix du portefeuille en fonction de l'horizon et du rendement du dividende pour un investisseur d'aversion au risque $\gamma = 5$. La figure est obtenue avec notre estimateur GMM conditionnel et les données sur la période allant de janvier 1997 à décembre 2021.

En résumé, les figures 2.17 et 2.18 présentent des résultats consistants avec l'évidence selon laquelle le choix de portefeuille dépend fortement des états du rendement du dividende et de l'horizon. Toutefois, en utilisant les données post-développement au lieu des données de développement, nous observons que la dynamique du choix de portefeuille change, même si elle demeure non-linéaire. Nous remarquons que les proportions de portefeuille allouées en actions sur la période de développement sont supérieures à celles de la période post-développement. Ce qui pourrait s'expliquer par la baisse du rendement du dividende sur la période post-développement (voir figure 2.1).

Discussions et Conclusion

Dans le présent mémoire, nous effectuons une étude comparative des résultats sur la période des données d'analyse du modèle de [Brandt \(1999\)](#) ainsi que par rapport aux données post-publication. Nous reconstituons les données et les variables de [Brandt \(1999\)](#) et répliquons ses résultats sur la période de développement. Ensuite, les données post-développement sont utilisés pour analyser les résultats du modèle pour le cas des moments inconditionnels et conditionnels respectivement.

Pour les moments inconditionnels, les résultats d'analyse empirique sur les données post-développement montrent que pour les horizons supérieurs à 3 mois, les trajectoires des choix optimaux du portefeuille ne changent pas comparativement à ceux des données de développement du modèle de [Brandt \(1999\)](#). Cependant, les valeurs obtenues post-développement sont en général inférieures aux valeurs obtenues au développement. Aussi, on note qu'avec un horizon d'un mois, l'investisseur choisit un portefeuille d'environ 25% d'actions, quelque soit son aversion au risque. Ce qui montre que lorsque les données post-développement sont utilisées, l'investisseur avec une aversion au risque $\gamma = 5$ et ayant un horizon d'un mois aurait été prudent en investissant 25% en actions au lieu de 82% en considérant les données de développement (voir les tableaux [2.4](#) et [2.3](#) respectivement).

Pour le volet conditionnel, nous avons noté quelques changements sur l'évolution des patterns des choix optimaux de portefeuille, notamment lorsque les variables prédictives sont le rendement du dividende, la prime de défaut, et la prime de durée. Par ailleurs, conditionnellement au rendement excédentaire retardé, l'évolution du pattern du choix optimal du portefeuille est sensiblement la même. Ceci implique que l'information sur les

rendements excédentaires retardés ne change pas l'évolution des choix du portefeuille. Cependant, on a remarqué aussi que comparativement aux données de développement, quelque soit l'aversion au risque, les valeurs des poids alloués en actions diminuent sur la période post-développement lorsque les variables prédictives sont la prime de durée et le rendement excédentaire retardé. Pour les variables rendement du dividende et la prime de défaut, on note plutôt une augmentation des valeurs des choix optimaux du portefeuille.

En résumé, il y'a quelques changements sur les résultats lorsqu'on utilise les données post-publication au lieu des données de développement, même si la conclusion générale est la même, à savoir que l'investisseur choisit des portefeuilles optimaux différents en fonction de la réalisation de chaque variable de prévision.

En outre, dans ce mémoire, la méthode ITGMM est couplée avec le LOOCV. Notre estimateur est efficace, consistant et fiable. Nous montrons par exemple, que notre estimateur trouve la valeur exacte du résultat théorique du choix de consommation lorsque $\gamma = 1$ (i.e., cas ayant une utilité logarithmique), ce qui n'est pas chaque fois le cas dans [Brandt \(1999\)](#). De plus, un des grands avantages de la méthode ITGMM est sa flexibilité puisqu'elle permet d'estimer les paramètres d'un modèle sans avoir besoin de faire d'hypothèses sur les distributions des variables afin de résoudre analytiquement le modèle. Cependant, bien que notre estimateur a aussi un avantage d'être précis, il est cependant coûteux en temps de calcul du fait de l'utilisation de la méthode d'apprentissage automatique cross-validation de type "leave-one-out". Pour gagner en temps de calcul, une piste pourrait-être d'utiliser une méthode d'apprentissage automatique, comme par exemple le "k-fold" cross validation.

Bibliographie

- Aït-Sahalia, Y. et M. W. Brandt. 2001, «Variable selection for portfolio choice», *The Journal of Finance*, vol. 56, n° 4, p. 1297–1351.
- Ait-Sahalia, Y. et M. W. Brandt. 2007, «Consumption and portfolio choice with option-implied state prices», cahier de recherche, National Bureau of Economic Research.
- Balduzzi, P. et A. W. Lynch. 1999, «Transaction costs and predictability : Some utility cost calculations», *Journal of Financial Economics*, vol. 52, n° 1, p. 47–78.
- Barberis, N. 2000, «Investing for the long run when returns are predictable», *The Journal of Finance*, vol. 55, n° 1, p. 225–264.
- Blume, M. E. 1980, «Stock returns and dividend yields : Some more evidence», *The Review of Economics and Statistics*, p. 567–577.
- Bollerslev, T. 1986, «Generalized autoregressive conditional heteroskedasticity», *Journal of econometrics*, vol. 31, n° 3, p. 307–327.
- Brandt, M. W. 1999, «Estimating portfolio and consumption choice : A conditional euler equations approach», *The Journal of Finance*, vol. 54, n° 5, p. 1609–1645.
- Brandt, M. W., A. Goyal, P. Santa-Clara et J. R. Stroud. 2005a, «A simulation approach to dynamic portfolio choice with an application to learning about return predictability», *The Review of Financial Studies*, vol. 18, n° 3, p. 831–873.
- Brandt, M. W. et P. Santa-Clara. 2006, «Dynamic portfolio selection by augmenting the asset space», *The journal of Finance*, vol. 61, n° 5, p. 2187–2217.

- Brandt, M. W., P. Santa-Clara et R. Valkanov. 2005b, «Parametric portfolio policies : Exploiting characteristics in the cross-section of equity returns», *The Review of Financial Studies*, vol. 22, n° 9, p. 3411–3447.
- Brennan, M. J., E. S. Schwartz et R. Lagnado. 1997, «Strategic asset allocation», *Journal of Economic dynamics and Control*, vol. 21, n° 8-9, p. 1377–1403.
- Campbell, J. Y. 1987, «Stock returns and the term structure», *Journal of financial economics*, vol. 18, n° 2, p. 373–399.
- Campbell, J. Y., J. Cocco, F. Gomes, P. J. Maenhout et L. M. Viceira. 2001, «Stock market mean reversion and the optimal equity allocation of a long-lived investor», *Review of Finance*, vol. 5, n° 3, p. 269–292.
- Campbell, J. Y. et R. J. Shiller. 1988a, «The dividend-price ratio and expectations of future dividends and discount factors», *The Review of Financial Studies*, vol. 1, n° 3, p. 195–228.
- Campbell, J. Y. et L. M. Viceira. 1999, «Consumption and portfolio decisions when expected returns are time varying», *The Quarterly Journal of Economics*, vol. 114, n° 2, p. 433–495.
- Cenesizoglu, T. et B. Essid. 2011, «The effect of monetary policy on credit spreads», *Journal of Financial Research*, vol. 35, n° 4, p. 581–613.
- Cenesizoglu, T., F. d. O. F. Ribeiro et J. J. Reeves. 2017, «Beta forecasting at long horizons», *International journal of forecasting*, vol. 33, n° 4, p. 936–957.
- Cenesizoglu, T. et A. Timmermann. 2012, «Do return prediction models add economic value?», *Journal of Banking & Finance*, vol. 36, n° 11, p. 2974–2987.
- Chaussé, P. 2010, «Computing generalized method of moments and generalized empirical likelihood with r », *Journal of Statistical Software*, vol. 34, p. 1–35.

- Cochrane, J. H. 1991, «Production-based asset pricing and the link between stock returns and economic fluctuations», *The Journal of Finance*, vol. 46, n° 1, p. 209–237.
- Connor, G., S. Li et O. B. Linton. 2020, «A dynamic semiparametric characteristics-based model for optimal portfolio selection», *Cambridge Working Papers in Economics*, vol. 21-1.
- Dammon, R. M., C. S. Spatt et H. H. Zhang. 2001, «Optimal consumption and investment with capital gains taxes», *The Review of Financial Studies*, vol. 14, n° 3, p. 583–616.
- Di Maggio, M., A. Kermani et K. Majlesi. 2020, «Stock market returns and consumption», *The Journal of Finance*, vol. 75, n° 6, p. 3175–3219.
- Engle, R. F. 1982, «Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation», *Econometrica : Journal of the Econometric Society*, p. 987–1007.
- Fama, E. F. 1970, «Multiperiod consumption-investment decisions», *American Economic Review*, p. 163–174.
- Fama, E. F. et K. R. French. 1988a, «Dividend yields and expected stock returns», *Journal of financial economics*, vol. 22, n° 1, p. 3–25.
- Fama, E. F. et K. R. French. 1988b, «Permanent and temporary components of stock prices», *Journal of political Economy*, vol. 96, n° 2, p. 246–273.
- Fama, E. F. et K. R. French. 1989, «Business conditions and expected returns on stocks and bonds», *Journal of financial economics*, vol. 25, n° 1, p. 23–49.
- Fama, E. F. et G. W. Schwert. 1977, «Asset returns and inflation», *Journal of financial economics*, vol. 5, n° 2, p. 115–146.
- Ferson, W. E. et C. R. Harvey. 1991, «The variation of economic risk premiums», *Journal of political economy*, vol. 99, n° 2, p. 385–415.

- French, K. R., G. W. Schwert et R. F. Stambaugh. 1987, «Expected stock returns and volatility», *Journal of financial Economics*, vol. 19, n° 1, p. 3–29.
- Gourieroux, C., A. Monfort, C. Tenreiro et collab.. 1994, «Kernel m-estimators : Non parametric diagnostics for structural models», cahier de recherche, CEPREMAP.
- Hall, A. R. 2005, *Generalized method of moments*, Oxford university press Oxford.
- Hall, P., S. N. Lahiri et Y. K. Truong. 1995, «On bandwidth choice for density estimation with dependent data», *The Annals of Statistics*, vol. 23, n° 6, p. 2241–2263.
- Hamilton, J. D. 1994, *Time series analysis*, Princeton university press, Princeton, NJ, U.S.A.
- Hansen, L. P. 1982, «Large sample properties of generalized method of moments estimators», *Econometrica : Journal of the Econometric Society*, p. 1029–1054.
- Hansen, L. P., J. Heaton et A. Yaron. 1996, «Finite-sample properties of some alternative gmm estimators», *Journal of Business & Economic Statistics*, vol. 14, n° 3, p. 262–280.
- Härdle, W. 1984, «Robust regression function estimation», *Journal of Multivariate Analysis*, vol. 14, n° 2, p. 169–180.
- Härdle, W. 1990, *Applied nonparametric regression*, 19, Cambridge university press.
- Härdle, W. et J. S. Marron. 1985, «Optimal bandwidth selection in nonparametric regression function estimation», *The Annals of Statistics*, p. 1465–1481.
- Harvey, C. R. 2001, «The specification of conditional expectations», *Journal of Empirical Finance*, vol. 8, n° 5, p. 573–637.
- Hayashi, F. 2000, «Econometrics», *Princeton University Press, New Jersey*.
- Heaton, J. et D. Lucas. 2000, «Portfolio choice and asset prices : The importance of entrepreneurial risk», *The journal of finance*, vol. 55, n° 3, p. 1163–1198.

- Hodrick, R. J. 1992, «Dividend yields and expected stock returns : Alternative procedures for inference and measurement», *The Review of Financial Studies*, vol. 5, n° 3, p. 357–386.
- Keim, D. B. et R. F. Stambaugh. 1986, «Predicting returns in the stock and bond markets», *Journal of financial Economics*, vol. 17, n° 2, p. 357–390.
- Kim, T. S. et E. Omberg. 1996, «Dynamic nonmyopic portfolio behavior», *The Review of Financial Studies*, vol. 9, n° 1, p. 141–161.
- Kogan, L. et R. Uppal. 2000, «Risk aversion and optimal portfolio policies in partial and general equilibrium economies», *National Bureau of Economic Research Cambridge, Mass., USA*.
- Lamont, O. 1998, «Earnings and expected returns», *The journal of Finance*, vol. 53, n° 5, p. 1563–1587.
- Lee, D. S. et T. Lemieux. 2010, «Regression discontinuity designs in economics», *Journal of economic literature*, vol. 48, n° 2, p. 281–355.
- Lettau, M. et S. Ludvigson. 2001, «Consumption, aggregate wealth, and expected stock returns», *the Journal of Finance*, vol. 56, n° 3, p. 815–849.
- Liu, J. 2007, «Portfolio selection in stochastic environments», *The Review of Financial Studies*, vol. 20, n° 1, p. 1–39.
- Markowitz, H. M. 1952, «Portfolio selection», *The Review of Financial Studies*, vol. 7, p. 77–91.
- Merton, R. C. 1969, «Lifetime portfolio selection under uncertainty : The continuous-time case», *The review of Economics and Statistics*, p. 247–257.
- Merton, R. C. 1971, «Optimum consumption and portfolio rules in a continuous-time model», *Journal of Economic Theory*, vol. 3, p. 373–413.

- Newey, W. K. et D. McFadden. 1994, «Estimation in large samples», *The Handbook of Econometrics*, vol. 4.
- Newey, W. K. et R. J. Smith. 2004, «Higher order properties of gmm and generalized empirical likelihood estimators», *Econometrica*, vol. 72, n° 1, p. 219–255.
- Pontiff, J. et L. D. Schall. 1998, «Book-to-market ratios as predictors of market returns», *Journal of Financial Economics*, vol. 49, n° 2, p. 141–160.
- Poterba, J. M. et L. H. Summers. 1988, «Mean reversion in stock prices : Evidence and implications», *Journal of financial economics*, vol. 22, n° 1, p. 27–59.
- Samuelson, P. A. 1969, «Lifetime portfolio selection by dynamic stochastic programming», *Review of Economics and Statistics*, vol. 51, p. 239–246.
- Schwert, G. W. 1989, «Why does stock market volatility change over time?», *The journal of finance*, vol. 44, n° 5, p. 1115–1153.
- Silverman, B. W. 1986, *Density estimation for statistics and data analysis*, Chapman and Hall, New York.
- Viceira, L. M. 2001, «Optimal portfolio choice for long-horizon investors with nontradable labor income», *The Journal of Finance*, vol. 56, n° 2, p. 433–470.
- Wachter, J. A. 2002, «Portfolio and consumption decisions under mean-reverting returns : An exact solution for complete markets», *Journal of financial and quantitative analysis*, vol. 37, n° 1, p. 63–91.
- Welch, I. et A. Goyal. 2008, «A comprehensive look at the empirical performance of equity premium prediction», *The Review of Financial Studies*, vol. 21, n° 4, p. 1455–1508.
- Whitelaw, R. F. 1994, «Time variations and covariations in the expectation and volatility of stock market returns», *The Journal of Finance*, vol. 49, n° 2, p. 515–541.

Wu, C. O. 1997, «A cross-validation bandwidth choice for kernel density estimates with selection biased data», *Journal of multivariate analysis*, vol. 61, n° 1, p. 38–60.

Annexes

Annexe A – Preuve du lemme 1

On a :

$$V(z_t, W_t, T - t) = \max_{\{\alpha_\tau, q_\tau\}_{\tau=t}^T} E \left[\sum_{\tau=t}^T \beta^{\tau-t} u(C_\tau) / z_t \right]$$

(par la linéarité de l'espérance)

$$= \max_{\alpha_t, q_t} E [\beta^0 u(C_t) / z_t] + \max_{\{\alpha_\tau, q_\tau\}_{\tau=t+1}^T} E \left[\sum_{\tau=t+1}^T \beta^{\tau-t-1+1} u(C_\tau) / z_t \right]$$

($E[u(C_t) / z_t] = E[u(C_t)] = u(C_t)$ puisque $C_t = q_t W_t$ et z_t sont observables)

$$= \max_{\alpha_t, q_t} u(q_t W_t) + \max_{\{\alpha_\tau, q_\tau\}_{\tau=t+1}^T} \beta E \left[\sum_{\tau=t+1}^T \beta^{\tau-(t+1)} u(C_\tau) / z_t \right]$$

(par la loi des espérances itérées $E[E[X/\mathcal{G}_2]/\mathcal{G}_1] = E[X/\mathcal{G}_1] \quad \forall \mathcal{G}_1 \subseteq \mathcal{G}_2$)

$$= \max_{\alpha_t, q_t} u(q_t W_t) + \beta E \left[\max_{\{\alpha_\tau, q_\tau\}_{\tau=t+1}^T} E \left[\sum_{\tau=t+1}^T \beta^{\tau-(t+1)} u(C_\tau) / z_{t+1} \right] / z_t \right]$$

$$= \max_{\alpha_t, q_t} u(q_t W_t) + \beta E [V(z_{t+1}, W_{t+1}, T - (t + 1)) / z_t]. \quad \#$$

Annexe B – Preuve du théorème 1

La preuve se fera en deux étapes.

1. Pour le choix de consommation. La condition de premier ordre est donnée par :

$$\begin{aligned}
0 &= \frac{\partial V}{\partial C} = u'(qW_t) + \beta E \left[\frac{\partial}{\partial C} (\psi(z_{t+1}, T-t-1)u(\tilde{W}_{t+1})) / z_t \right] \\
&= u'(qW_t) + \beta E \left[\psi(z_{t+1}, T-t-1) * \frac{\partial \tilde{W}_{t+1}}{\partial C} * u'(\tilde{W}_{t+1}) / z_t \right] \\
&= E[u'(qW_t) / z_t] - \beta E \left[\psi(z_{t+1}, T-t-1) * \tilde{R}_{t+1}^p * u'(\tilde{W}_{t+1}) / z_t \right] \\
&= E \left[u'(qW_t) - \beta \psi(z_{t+1}, T-t-1) * \tilde{R}_{t+1}^p * u'(\tilde{W}_{t+1}) / z_t \right] \\
&\text{diviser chaque membre par } -u'(qW_t) \\
&= E \left[-1 + \beta \psi(z_{t+1}, T-t-1) * \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} * \tilde{R}_{t+1}^p / z_t \right]
\end{aligned}$$

D'où,

$$E \left[-1 + \beta \psi(z_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p / z_t \right] = 0. \quad (39)$$

2. Pour le portefeuille. La condition de premier ordre est donnée par :

$$\begin{aligned}
0 &= \frac{\partial V}{\partial \alpha'} = \frac{\partial}{\partial \alpha'} (u(qW_t)) + \beta E \left[\frac{\partial}{\partial \alpha'} (\psi(z_{t+1}, T-t-1)u(\tilde{W}_{t+1})) / z_t \right] \\
&= 0 + \beta E \left[\psi(z_{t+1}, T-t-1) * \frac{\partial \tilde{W}_{t+1}}{\partial \alpha'} * u'(\tilde{W}_{t+1}) / z_t \right] \\
&= \beta E \left[\psi(z_{t+1}, T-t-1) * \frac{\partial}{\partial \alpha'} \left[(1-q)W_t(R^f + \alpha' \tilde{R}_{t+1}^e) \right] * u'(\tilde{W}_{t+1}) / z_t \right] \\
&= E \left[\beta \psi(z_{t+1}, T-t-1) * (1-q)W_t \tilde{R}_{t+1}^e * u'(\tilde{W}_{t+1}) / z_t \right] \\
&= \beta(1-q)W_t E \left[\psi(z_{t+1}, T-t-1) * \tilde{R}_{t+1}^e * u'(\tilde{W}_{t+1}) / z_t \right] \\
&\text{diviser chaque membre par } \beta(1-q)W_t \\
&= E \left[\psi(z_{t+1}, T-t-1) * \tilde{R}_{t+1}^e * u'(\tilde{W}_{t+1}) / z_t \right]
\end{aligned}$$

D'où,

$$E \left[\psi(z_{t+1}, T-t-1) * u'(\tilde{W}_{t+1}) * \tilde{R}_{t+1}^e / z_t \right] = 0. \quad (40)$$

Ainsi, d'après les équations (40) et (39) respectivement, on a :

$$m_{t+1}(\alpha, q, T-t) = \begin{bmatrix} \psi(z_{t+1}, T-t-1) u'(\tilde{W}_{t+1}) \tilde{R}_{t+1}^e \\ \beta \psi(z_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p - 1 \end{bmatrix} \cdot \# \quad (41)$$

Annexe C – Preuve du théorème 2

Les résultats (2.35) et (2.15) sont similaires exceptée qu'au dénominateur de la première équation de (2.35), on a le terme $u'(W_t)$. Donc la preuve de (2.35) découle directement de la démonstration du Théorème 1. Il suffit de diviser la condition de premier ordre $\frac{\partial V}{\partial \alpha'}$ par le terme $u'(W_t)$. On a :

$$m_{t+1}(\alpha, q, T-t) = \begin{bmatrix} \psi(\tilde{z}_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(W_t)} \tilde{R}_{t+1}^e \\ \beta \psi(\tilde{z}_{t+1}, T-t-1) \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p - 1 \end{bmatrix} \cdot \quad (42)$$

Posons :

$$m_{t+1}(\alpha, q, T-t) = \begin{bmatrix} m_{t+1}^1(\alpha, q, T-t) \\ m_{t+1}^2(\alpha, q, T-t) \end{bmatrix} \cdot \quad (43)$$

Puisque $\psi(\tilde{z}_{t+1}, 0) = 1$, alors

$$m_{t+1}(\alpha, q, 1) = \begin{bmatrix} m_{t+1}^1(\alpha, q, 1) \\ m_{t+1}^2(\alpha, q, 1) \end{bmatrix} = \begin{bmatrix} \frac{u'(\tilde{W}_{t+1})}{u'(W_t)} \tilde{R}_{t+1}^e \\ \beta \frac{u'(\tilde{W}_{t+1})}{u'(qW_t)} \tilde{R}_{t+1}^p - 1 \end{bmatrix} \cdot \quad (44)$$

D'après l'équation (2.25), on a :

$$\left\{ \begin{array}{l} \hat{D}_1 = E \left[N(z_t, \tau) \frac{\partial m_{t+1}^1}{\partial \alpha'} |z \right], \\ \hat{D}_2 = E \left[N(z_t, \tau) \frac{\partial m_{t+1}^1}{\partial q} |z \right], \\ \hat{D}_3 = E \left[N(z_t, \tau) \frac{\partial m_{t+1}^2}{\partial \alpha'} |z \right], \\ \hat{D}_4 = E \left[N(z_t, \tau) \frac{\partial m_{t+1}^2}{\partial q} |z \right]. \end{array} \right. \quad (45)$$

IL est facile de montre que :

$$\frac{\partial m_{t+1}}{\partial q} = \left\{ \begin{array}{l} \gamma(1-q)^{-\gamma-1} \times \tilde{R}_{t+1}^e \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma} \\ \frac{\beta\gamma}{q(1-q)} \times \left(\frac{1-q}{q} \right)^{-\gamma} \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{1-\gamma}, \end{array} \right. \quad (46)$$

et que

$$\frac{\partial m_{t+1}}{\partial \alpha'} = \left\{ \begin{array}{l} -\gamma(1-q)^{-\gamma} \times (\tilde{R}_{t+1}^e)^2 \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma-1} \\ \beta(1-\gamma) \times \left(\frac{1-q}{q} \right)^{-\gamma} \times \tilde{R}_{t+1}^e \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma}. \end{array} \right. \quad (47)$$

D'où,

$$\left\{ \begin{array}{l} \hat{D}_1 = E \left[-\gamma(1-q)^{-\gamma} \times (\tilde{R}_{t+1}^e)^2 \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma-1} \times N(z_t, \tau) |z \right], \\ \hat{D}_2 = E \left[\gamma(1-q)^{-\gamma-1} \times \tilde{R}_{t+1}^e \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma} \times N(z_t, \tau) |z \right], \\ \hat{D}_3 = E \left[\beta(1-\gamma) \times \left(\frac{1-q}{q} \right)^{-\gamma} \times \tilde{R}_{t+1}^e \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{-\gamma} \times N(z_t, \tau) |z \right], \\ \hat{D}_4 = E \left[\frac{\beta\gamma}{q(1-q)} \times \left(\frac{1-q}{q} \right)^{-\gamma} \times (R^f + \alpha' \tilde{R}_{t+1}^e)^{1-\gamma} \times N(z_t, \tau) |z \right]. \end{array} \right. \quad \# \quad (48)$$

Annexe D – Rappels sommaires sur les approches

1SGMM et 2SGMM

Dans cette partie nous faisons une brève description de 1SGMM et 2SGMM qui sont les étapes 1 et 2 de l'ITGMM utilisée dans ce mémoire.

GMM à une étape (1SGMM)

Soit $\{x_i, i = 1, \dots, n\}$ iid, on s'intéresse à estimer θ^* . Supposons que la loi forte des grands nombres s'applique à $\psi(\theta, x)$ pour tout θ admissible. On a :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \psi(\theta, x_i) = E[\psi(\theta, x)] = 0, \quad (49)$$

presque sûrement (i.e., avec la probabilité 1).

En général, (1.1) est un vecteur de fonctions non-linéaires de θ^* et le nombre d'équations en (1.1) n'est pas limité par la dimension de θ^* . Puisque l'efficacité augmente avec le nombre d'instruments, q est souvent supérieur à p , ce qui implique qu'il n'existe pas de solution de

$$\frac{1}{n} \sum_{i=1}^n \psi(\theta, x_i) = 0. \quad (50)$$

Suivant l'idée de base de l'estimateur GMM (i.e., Hansen (1982)), le meilleur moyen est de prendre l'expression à gauche de (50) le plus proche possible de zéro en minimisant la fonction quadratique

$$J_{\widehat{W}, n}(\theta) = \left[\frac{1}{n} \sum_{i=1}^n \psi(\theta, x_i) \right]' \widehat{W} \left[\frac{1}{n} \sum_{i=1}^n \psi(\theta, x_i) \right], \quad (51)$$

par rapport à θ , où $\widehat{W} \in \mathcal{M}_{q \times q}$ est une matrice symétrique et définie positive qui satisfait $\lim_{n \rightarrow \infty} \widehat{W} = W$ presque sûrement, où W est symétrique et définie positive. L'estimateur GMM (noté $\widehat{\theta}$) est solution de (51). Ce qui signifie que

$$\hat{\theta} = \arg \min_{\theta} J_{\hat{W},n}(\theta). \quad (52)$$

Un estimateur est consistant s'il converge vers la vraie valeur.

Proposition 1 (*Consistence*)

Sous certaines conditions de régularité appropriées, l'estimateur GMM $\hat{\theta}$ est consistant, i.e.,

$$\hat{\theta} \xrightarrow{P} \theta^*. \quad (53)$$

Formellement cela signifie que

$$Pr \left[\|\hat{\theta} - \theta^*\| > \varepsilon \right] = 0, \forall \varepsilon > 0. \quad (54)$$

Pour les détails de preuve de la proposition 1 voir [Hayashi \(2000\)](#) chapitre 7.

Proposition 2 (*Covariance asymptotique*)

(a) *Sous certaines conditions de régularité appropriées (voir [Hayashi \(2000\)](#), chapitre 3), le théorème de la limite centrale permet de montrer que*

$$\sqrt{n}(\hat{\theta} - \theta^*) \longrightarrow \mathcal{N}(0, V), \quad (55)$$

où

$$V = (\Gamma' W \Gamma)^{-1} \Gamma' W \Delta W \Gamma (\Gamma' W \Gamma)^{-1}, \quad (56)$$

avec $\Gamma = E \left[\frac{\partial \psi}{\partial \theta'}(\theta^*, x) \right] \in \mathcal{M}_{q \times p}$ le gradient de la fonction moment par rapport aux paramètres et $\Delta = E[\psi(\theta^*, x) \psi(\theta^*, x)'] \in \mathcal{M}_{q \times q}$ le produit extérieur des moments.

(b) *Sous certaines conditions de régularité appropriées, un estimateur consistant de la covariance asymptotique est*

$$\hat{V} = (\hat{\Gamma}' \hat{W} \hat{\Gamma})^{-1} \hat{\Gamma}' \hat{W} \hat{\Delta} \hat{W} \hat{\Gamma} (\hat{\Gamma}' \hat{W} \hat{\Gamma})^{-1}, \quad (57)$$

$$\text{où } \hat{\Gamma} = \frac{1}{n} \sum_{i=1}^n \frac{\partial \psi}{\partial \theta'}(\theta^*, x_i) \text{ et } \hat{\Delta} = \frac{1}{n} \sum_{i=1}^n \psi(\theta^*, x_i) \psi(\theta^*, x_i)'$$

Pour les détails de preuve de la proposition 2 voir Hayashi (2000) chapitre 3.

L'efficacité permet de clarifier si l'estimateur est précis ou pas. Notons que la covariance asymptotique (56) dépend du choix de la matrice de poids \widehat{W} . L'estimateur GMM efficace (optimal) est construit avec la matrice de poids qui produit la plus petite valeur de la covariance asymptotique (56).

Pour $\widehat{W} = \widehat{\Delta}^{-1}$, on a :

$$\widehat{V} = (\widehat{\Gamma}' \widehat{\Delta}^{-1} \widehat{\Gamma})^{-1} \widehat{\Gamma}' \widehat{\Delta}^{-1} \widehat{\Delta} \widehat{\Delta}^{-1} \widehat{\Gamma} (\widehat{\Gamma}' \widehat{\Delta}^{-1} \widehat{\Gamma})^{-1} = (\widehat{\Gamma}' \widehat{\Delta}^{-1} \widehat{\Gamma})^{-1}. \quad (58)$$

Proposition 3 (*Efficacité*)

L'estimateur GMM efficace est :

$$\widehat{\theta} = \arg \min_{\theta} J_{\widehat{\Delta}^{-1}, n}(\theta). \quad (59)$$

Pour les détails de preuve de la proposition 3 voir Hansen (1982) et Hayashi (2000) chapitre 3.

Remarque 10 i) Lorsque le système est bien défini (*just identified*), c'est-à-dire que $p = q$, l'estimateur GMM ne dépend pas du choix de la matrice de poids \widehat{W} . En effet, lorsque $p = q$, Γ est inversible car est une matrice carrée de rang maximal. La matrice de covariance asymptotique devient :

$$\begin{aligned} V &= (\Gamma' W \Gamma)^{-1} \Gamma' W \Delta W \Gamma (\Gamma' W \Gamma)^{-1} \\ &= \Gamma^{-1} W^{-1} \Gamma'^{-1} \Gamma' W \Delta W \Gamma \Gamma^{-1} W^{-1} \Gamma'^{-1} \\ &= \Gamma^{-1} \Delta \Gamma'^{-1} \\ &= (\Gamma' \Delta^{-1} \Gamma)^{-1}. \end{aligned}$$

ii) Lorsque le système est sur-identifié (*overidentified*), c'est-à-dire que $q > p$, le choix de la matrice de poids W est important. Le choix optimal de W est Δ^{-1} (voir la Proposition 3).

Pour trouver l'estimateur GMM optimal (i.e., efficient), on a considéré que la matrice de poids $W = \Delta^{-1}$. Cependant, Δ est l'espérance du produit extérieur des moments en θ^* . Toutefois, on ne connaît pas θ^* . Une solution a été proposé dans la version originale de la GMM par Hansen (1982) et est appelée GMM à deux étapes.

GMM à deux étape (2SGMM)

L'estimateur efficient GMM à deux étapes utilise un estimateur initial θ_0 consistant, mais inefficient, par minimisation de $J_{\widehat{W},n}(\theta)$ en utilisant une matrice de poids arbitraire \widehat{W} symétrique et définie positive; e.g., la matrice identité de dimension q (i.e., I_q). On a donc l'algorithme 2SGMM suivant :

1. Calculer $\theta_0 = \arg \min_{\theta} J_{I_q,n}(\theta)$.
2. Calculer $\widehat{\Delta}^{-1} = \left[\frac{1}{n} \sum_{i=1}^n \psi(\theta_0, x_i) \psi(\theta_0, x_i)' \right]^{-1}$.
3. Calculer $\widehat{\theta}_{2sgmm} = \arg \min_{\theta} J_{\widehat{\Delta}^{-1},n}(\theta)$.

Hansen et collab. (1996) proposent une version améliorée de 2SGMM. C'est la méthode GMM itérée.

