

Université de Sherbrooke

Faculté d'administration

Prévision de la prime de marché canadienne et américaine

PAR

Maxime Lemay-Crilly

MÉMOIRE

Soumis à la réalisation partielle des exigences pour le diplôme de Maîtrise en administration
incluant un cheminement de type recherche du domaine de la finance, 2015

DIRECTEURS

Frank Coggins

Guy Bellemare

16 décembre 2015

RÉSUMÉ

Dans le cadre de cette étude, il est question de prédire les primes de risque de marché pour les États-Unis et le Canada sur un horizon d'un mois en se basant sur les données économiques des 20 dernières années. En se basant sur les modèles élaborés précédemment dans la littérature, ce mémoire a pour but d'effectuer des prévisions plus précises que celles générées précédemment.

Ainsi, on observe que l'ensemble des modèles retenus, tant univariés, multivariés ou par agrégation sectorielle obtiennent un pouvoir explicatif supérieur au modèle naïf, et ce tant pour le marché américain que le marché canadien.

Pour le marché américain, le modèle multivarié *Stepwise Backward* obtient la meilleure performance du groupe étudié avec un R^2 de 0.10714 dans un contexte *In-Sample* et un R^2 de 0.22284 dans un contexte *Out-Of-Sample* selon le test de McCracken (2007). Ce modèle permet donc d'expliquer 22.28% de la variation mensuelle de la prime de risque américaine dans le contexte de l'échantillon observé dans cette étude. Le modèle en question est composé des variables économiques représentant les variations mensuelles au niveau de l'inflation, de la masse monétaire M2, ainsi que du dernier taux journalier du mois observé pour les obligations gouvernementales ayant une échéance de deux ans, cinq ans et dix ans.

Pour le marché canadien, le modèle multivarié *Stepwise Forward* obtient la meilleure performance du groupe étudié dans un contexte *In-Sample* avec un R^2 de 0.07760 selon le test de McCracken (2007). Cependant, dans un contexte *Out-Of-Sample*, le modèle de prévision par agrégation sectorielle élaboré à la section 4.4 obtient de loin la meilleure performance avec un R^2 de 0.17773 selon le test de McCracken (2007), permettant donc d'expliquer 17.77% de la variation mensuelle de la prime de risque canadienne.

La performance accrue des modèles de prévision dans un contexte *Out-Of-Sample* semble provenir d'une meilleure performance notamment dans les premières années d'observation, (2001 à 2007) grâce à l'exclusion des grandes variations affectant les dernières années de la période d'observation (2008 à 2011).

REMERCIEMENTS

Ce projet n'aurait pas été possible sans le support de mon directeur Frank Coggins, qui a corrigé mes nombreuses révisions et répondu à mes multiples questions. Merci aussi à Guy Bellemare pour son important apport. J'aimerais également remercier ma conjointe Patricia pour son support pendant cette longue période de rédaction, ainsi qu'Évelyne Girouard pour sa correction diligente de l'orthographe et de la grammaire.

Table des matières

Table des matières	iii
Liste des tableaux	v
Liste des figures	vii
1 Introduction	1
2 Revue de la littérature	1
2.1 Fondements théoriques des modèles de prévision	1
2.2 Premières études sur la prévision des rendements financiers	2
2.3 Études modernes sur la prévision des rendements financiers	7
2.4 Récentes études de prévision de rendements financiers	15
3 Objectifs	22
4 Méthodologie	24
4.1 Modèles de prévision de la prime de risque basés sur la moyenne historique	26
4.2 Modèles univariés de prévision de la prime de risque agrégée	27
4.3 Modèles multivariés de prévision de la prime de risque agrégée	28
4.4 Agrégation de modèles multivariés de prévision des primes sectorielles	28
4.4.1 Regroupements industriels	29
4.4.2 Modèles de prévision univariés	37

4.4.3	Modèles de prévision multivariés	38
4.4.4	Agrégation des prévisions sectorielles	39
4.5	Analyse comparative	40
4.6	Bootstrap	40
4.7	Prévision Out-of-Sample (OOS)	41
5	Données	44
5.1	Primes de risque	44
5.2	Regroupements industriels	45
5.3	Variables économiques	48
5.3.1	Données américaines	49
5.3.2	Données canadiennes	58
6	Résultats	61
6.1	Marché américain	61
6.1.1	Modèle de prévision de la prime de risque basés sur la moyenne historique	61
6.1.2	Modèles univariés de prévision de la prime de risque agrégée	61
6.1.3	Modèles multivariés de prévision de la prime de risque agrégée	63
6.1.4	Agrégation de modèles multivariés de prévision des primes sectorielles	66
6.1.5	Analyse comparative	70
6.1.6	Analyse Comparative - Out-Of-Sample	71
6.2	Marché canadien	75

6.2.1	Modèle de prévision de la prime de risque basés sur la moyenne historique	75
6.2.2	Modèles univariés de prévision de la prime de risque agrégée	75
6.2.3	Modèles multivariés de prévision de la prime de risque agrégée	77
6.2.4	Agrégation de modèles multivariés de prévision des primes sectorielles	79
6.2.5	Analyse comparative	82
6.2.6	Analyse Comparative - Out-Of-Sample	83
7	Conclusion	86
	Références	88
A	Description des industries par code SIC	92
B	Statistiques descriptives des industries de départ	93

Liste des tableaux

1	Variables étudiées par Welch et Goyal (2008)	17
2	Résultats abrégés de Rapach, Wohar et Rangvid (2005) (horizon de 1 mois)	19
3	Variables significatives IS et OOS dans Chrétien et Coggins (2011) ($\alpha = 0.05$)	20
4	Regroupements industriels du marché des États-Unis	33
5	Regroupements industriels du marché du Canada	34
6	Distribution des entreprises du marché canadien	35
7	Distribution des entreprises du marché canadien	36

8	Statistiques descriptives des primes de risques étudiées	44
9	Corrélation entre prime de risques observées et estimées	45
10	Statistiques descriptives des regroupements industriels américains	47
11	Statistiques descriptives des regroupements industriels canadiens	48
12	Composition du panier de monnaies étrangères	56
13	Statistiques descriptives des variables économiques américaines testées	57
14	Statistiques descriptives des variables économiques canadiennes testées	60
15	Résultats des modèles univariés de prévision de la prime de risque agrégée (US) . . .	62
16	Résultats du modèle multivarié <i>backward</i> de prévision de la prime de risque agrégée (US)	63
17	Résultats du test VIF sur le modèle multivarié <i>backward</i> de prévision de la prime de risque agrégée	64
18	Résultats du modèle multivarié <i>forward</i> de prévision de la prime de risque agrégée (US)	65
19	Résultats des modèles de prévision de la prime de risque de chaque regroupement industriel (US)	66
20	Sélection des modèles de prévisions de la prime de risque de chaque regroupement industriel (US)	69
21	Analyse comparative US - In-Sample	70
22	Analyse comparative meilleur modèle US - In-Sample	71
23	Analyse comparative US - Out-Of-Sample	72
24	Analyse comparative meilleur modèle US - Out-Of-Sample	73
25	Résultats des modèles univariés de prévision de la prime de risque agrégée (CA) . . .	76

26	Résultats du modèle multivarié <i>backward</i> de prévision de la prime de risque agrégée (CA)	77
27	Résultats du modèle multivarié <i>forward</i> de prévision de la prime de risque agrégée (CA)	78
28	Résultats des modèles de prévision de la prime de risque de chaque regroupement industriel (CA)	79
29	Résultats des prévisions de rendements sectoriels canadiens	81
30	Analyse comparative CA - In-Sample	82
31	Analyse comparative meilleur modèle CA - In-Sample	82
32	Analyse comparative CA - Out-Of-Sample	83
33	Analyse comparative meilleur modèle CA - Out-Of-Sample	84
34	Description des industries par code SIC	92
35	Statistiques descriptives des industries de départ (É-U)	94
36	Statistiques descriptives des industries de départ (Canada)	95

Liste des figures

1	Comparaison du ratio Book-to-Market de six entreprises	14
2	Vue d'ensemble de la méthodologie	25
3	Vue d'ensemble de la méthodologie	25
4	Vue d'ensemble de la méthodologie des modèles par agrégation de prévisions sectorielles	29
5	Vue d'ensemble de la méthodologie d'estimation des prévisions <i>Out-of-Sample</i>	43

6	Comparaison des résultats du modèle <i>Stepwise Backward</i>	74
7	Comparaison des résultats du modèle de prévision par agrégation sectorielle	85

1 Introduction

Plusieurs études au fil des ans ont obtenus des résultats favorables quant à la prévisibilité de la prime de risque des États-Unis. Ces études sont généralement basées sur un assortiment de variables macroéconomiques. La prévision de la prime de risque a longtemps été un sujet chaud de la littérature en raison des nombreuses décisions financières en découlant. La perception de la prime de risque future d'un individu affecte notamment ses décisions d'investissement et l'allocation du capital dans son portefeuille.

Dans le cadre de cette étude, il est question de prédire les primes de risque de marché pour les États-Unis et le Canada sur un horizon d'un mois en se basant sur les données économiques des 20 dernières années. En se basant sur les modèles élaborés précédemment dans la littérature, ce mémoire a pour but d'effectuer des prévisions plus précises que celles générées précédemment. Notamment, il est question de tester les prévisions générées par des modèles univariés, multivariés et, surtout, d'introduire et tester un modèle de prévision de la prime de risque de marché basé sur l'agrégation de prévisions de primes de risque sectorielles.

2 Revue de la littérature

2.1 Fondements théoriques des modèles de prévision

Le sujet de la prédictibilité des rendements financiers a toujours été d'actualité en finance. Parmi les premières études on retrouve l'article de Dow (1920) qui explore le rôle des ratios de dividende sur les rendements futurs. Cependant, ce n'est que dans les années 1970 que la littérature en finance a commencé à se pencher sur la prévision de rendements financiers futurs. Les premières études liées à cette vague de recherche empirique sur la prédictibilité des rendements espérés se concentrent en grande partie sur leur relation avec le taux d'inflation. Ces modèles de prédictibilité en finance, comme tous les modèles à caractère temporel, ont une base commune au sein de la littérature. Leur venue est dérivée de la contribution fondamentale à la littérature d'établissement du prix des actifs financiers, soit l'*intertemporal asset pricing model* de Merton (1973). Dans cet ouvrage, Robert C. Merton développe un modèle d'équilibre intertemporel basé sur le comportement des investisseurs. Il démontre que les demandes des investisseurs pour des actifs sont affectées par la possibilité de changements incertains au niveau de l'ensemble des opportunités d'investissement dans le temps. Ces investisseurs doivent être indemnisés pour les variations

défavorables de ces opportunités sous forme de rendements espérés, et ce, en excédent du risque systématique du marché. Selon Merton, la variable d'état représentant le mieux l'ensemble des opportunités d'investissement est le taux d'intérêt :

It (the interest rate) is observable, satisfies the condition of being stochastic over time, and while it is surely not the sole determinant of yields of other assets, it is an important factor. Hence, one should interpret the effects of a changing interest rate ... in the way economists have generally done in the past : namely, as a single (instrumental) variable representation of shifts in the investment opportunity set.

Ces variations de l'ensemble des opportunités d'investissement sont à la base des modèles à caractère intertemporel, tels que les modèles de prévision et les modèles d'évaluation conditionnelle du prix des titres. En effet, cette notion de base permet de justifier théoriquement que les rendements espérés d'un titre soient affectés par un élément provenant d'une période temporelle passée ou présente. Cette notion est très importante quant à la conceptualisation de modèles de prévision et à la justification de leur utilisation d'un point de vue théorique. Ce type de modèle étant par nature de caractère empirique, une forte base théorique sur la notion de temporalité est nécessaire pour obtenir une méthodologie empirique qui est académiquement justifiable et théoriquement valable.

Cet aspect théorique permet de donner un sens aux relations prévisionnelles entre les variables macroéconomiques et la prime de risque du marché. Cette recherche, comme plusieurs autres antérieurement, tente de démontrer un lien empirique entre ces deux éléments. Cependant, il est important de noter que, théoriquement, cette relation est démontrée par l'influence des variations incertaines de l'ensemble des opportunités d'investissement sur les rendements espérés du marché. Ceci dit, les résultats de cet ouvrage de recherche ne seront absolument pas de nature théorique, mais auront comme effet de refléter les conclusions de Merton (1973).

2.2 Premières études sur la prévision des rendements financiers

Le sujet de la prédictibilité des rendements financiers a toujours été d'actualité en finance. Même dans les débuts de la finance, Dow (1920) a exploré le rôle des ratios de dividende sur les rendements futurs. En effet, Charles Dow décrit dans son ouvrage deux méthodes générales de négociation de titres, soit une approche technique se basant sur les rendements passés et sur l'emploi d'ordres d'arrêt, puis une approche fondamentale. Cette dernière est basée sur la valeur

du titre, qui elle est basée sur l'analyse fondamentale. Dow stipule que la valeur des titres est liée à la portion de leurs bénéfices versés en tant que dividendes. Selon ses observations, les firmes versant une plus grande portion de leurs bénéfices sous forme de dividendes auraient de meilleurs rendements futurs à l'époque. Il s'agit donc de l'une des premières études faisant référence à un lien de causalité entre un facteur présent et des rendements financiers futurs.

Par la suite, l'opinion dominante dans la littérature académique jusqu'aux années 1970 est que les prix des actifs financiers pouvaient être fidèlement décrits par une marche aléatoire et qu'aucune tendance exploitable économiquement n'existait. Notamment, Kendall (1953) observe que les prix des actifs financiers ont tendance à sillonner de façon aléatoire au fil du temps lors de tests empiriques sur la prévision de prix futurs à partir de rendements passés dans le cadre d'une étude sur l'efficacité des marchés.

Cependant, à partir des années 1970, plusieurs études empiriques ont testé une variété de variables macroéconomiques dans le but de prédire des rendements financiers futurs. En effet, plusieurs d'entre elles démontrent que ces rendements sont prévisibles jusqu'à un certain niveau, soit via leurs propres historiques ou par le biais d'information publique, tels que les rendements de dividende, le ratio price-earnings, les taux d'intérêt à court et long terme, les variations de la production industrielle ou des variables calendrier (e.g. January Effect). La majorité de la littérature se concentre sur l'évidence statistique de la variation temporelle prévisible des rendements espérés, et ce, selon diverses approches, variables d'information et tests statistiques.

Les premières études liées à cette vague de recherche empirique sur la prédictibilité des rendements espérés se concentrent en grande partie sur leur relation avec le taux d'inflation. D'abord, Bodie (1976) examine la possibilité d'utiliser des actions ordinaires comme couverture contre l'inflation dans le but de réduire le risque des rendements réels de l'investisseur. Ce risque provient de l'incertitude associée au futur niveau des prix des biens de consommation. Selon l'auteur, cet élément serait représenté par la variance des rendements réels d'obligations nominales unipériodiques sans risque de défaut. Le rendement réel de ces titres est certain, sauf pour l'inflation monétaire de son environnement économique. Le but de cet ouvrage est donc de déterminer l'efficacité d'employer des actions ordinaires qui viendraient réduire la variance en question en combinant de façon représentative un portefeuille diversifié d'actions ordinaires et l'obligation sans risque. Déjà à cette époque, on observe un lien évident entre cet article et le modèle intertemporel de Merton (1973). Ce lien se retrouve au niveau du risque étudié par chacun des auteurs. Merton (1973) fait référence à une variation de l'ensemble des opportunités d'investissement d'un investisseur dans le temps, ce qui est directement lié à l'incertitude envers le prix des biens de consommation. Cette incertitude affecte nécessairement les opportunités d'investissement d'un investisseur, car son rendement réel

est en jeu . Bodie (1976) ne fait pas référence aux travaux de Merton dans son article, mais leurs motivations quant aux perspectives futures de l'investisseur suivent la même logique. C'est-à-dire que tous deux font référence à la dimension temporelle sous forme d'opportunités d'investissement futures affectant l'investisseur au moment présent.

Selon ses résultats, l'efficacité d'employer des actions ordinaires pour se prémunir contre l'inflation dépend de deux variables. La première de celles-ci serait le ratio de la variance de la composante stochastique sans inflation des rendements réels des actions ordinaires sur la variance de l'inflation non anticipée. Plus ce ratio est élevé, moins les actions ordinaires seraient efficaces comme couverture contre l'inflation. La deuxième variable est la différence entre le rendement nominal de l'obligation sans risque de défaut et le coefficient de l'inflation non anticipée en tant que variable explicative du rendement réel sur les actions ordinaires (voir équation 2.1). Plus grande est la valeur absolue de cette différence, plus sont efficaces les actions ordinaires en tant que couverture contre l'inflation. La conclusion de Bodie (1976) est que, contrairement à l'opinion des économistes à l'époque, le rendement réel sur les actions ordinaires est négativement lié à l'inflation anticipée et non anticipée à court terme.

$$1 + R_e(t) = \alpha_0 + \alpha_1 d(t) + \mu(t) \quad (2.1)$$

Où :

$R_e(t)$: Rendement réel sur les actions ordinaires par dollar investi au temps t .

α_0 : Constante du modèle.

α_1 : Coefficient de l'inflation non anticipée.

$d(t)$: Valeur de l'indice des prix à la consommation du début de la période t divisé par sa valeur à la fin de la période. Représente l'inflation non anticipée.

μ_t : Résidus du modèle au temps t .

Dans le même ordre d'idée, Fama (1981) évalue également le lien entre l'inflation et les rendements d'actions ordinaires. Dans son article, il tente d'expliquer la relation négative découverte entre ces deux éléments dans la période post-1953 par plusieurs auteurs, dont Bodie (1976) en particulier. Eugene F. Fama émet comme hypothèse que cette relation est une conséquence d'effets "proxy". Selon lui, les rendements d'actions ordinaires seraient déterminés par des prévisions

de variables réelles appropriées. La relation négative entre les rendements d'actions et l'inflation serait induite par les relations négatives entre l'inflation et l'activité réelle. Les résultats de ses tests empiriques démontrent une relation positive et significative entre les rendements d'actions ordinaires et des mesures reflétant l'activité réelle, telle que les dépenses en capital, le taux réel de rendement sur capital et sur la productivité. Parallèlement, Fama démontre également avec constance une relation négative entre l'inflation et l'activité réelle, tel que le stipule la théorie de la demande de la monnaie et la théorie quantitative de la monnaie.

Conforme à l'hypothèse d'effet *proxy*, les taux de croissance monétaires et l'activité réelle éliminent les relations négatives entre les rendements d'actions réels et les taux d'inflation anticipés, et ce, de façon annuelle, trimestrielle et mensuelle. De plus, les rendements d'actions ordinaires annuels perdent leur lien explicatif avec l'inflation non anticipée en comparaison avec l'activité réelle future. Ce phénomène ne se produit pas lorsqu'il s'agit d'observations mensuelles et trimestrielles. L'auteur explique ceci par les déficiences de chevauchement de taux de croissance annuels de l'activité réelle en tant que mesures des nouvelles informations mensuelles et trimestrielles utilisées par les marchés financiers pour fixer les prix. De façon générale, on retire de cet article une explication et une confirmation de la relation négative entre les rendements des actions ordinaires et l'inflation mise en évidence par Bodie (1976). L'élément important à retirer de cette conclusion est que l'auteur démontre empiriquement un lien significatif entre les rendements d'actions et des éléments réels du marché. Il s'agit d'un début vers les méthodes de prévision modernes, surtout en lien avec le fait que l'auteur emploie des prévisions des variables macroéconomiques.

Toujours dans l'étude de la relation entre l'inflation et les rendements d'actions, Geske et Roll (1983) ont tenté d'approfondir les explications pour la relation négative entre ces deux éléments qui ont été emmenés par Bodie (1976), Fama et Schwert (1977), et Fama (1981). De plus, leurs résultats montrent une relation négative entre le rendement des actions ordinaires, à la fois les taux d'intérêt à court terme de début de période, les variations contemporaines des taux d'intérêt à court terme, et les taux d'inflation non anticipés. Selon les auteurs, ces deux derniers éléments représenteraient deux méthodes pour mesurer le même aspect économique. Dans la littérature antérieure, seulement l'explication proposée par Fama (1981) en lien avec la demande monétaire serait logiquement constante. Cependant, comme le mentionne Fama lui-même, cette explication ne semble pas décrire la totalité du phénomène empirique. L'explication supplémentaire apportée par Geske et Roll est qu'un choc réel aléatoire négatif affectant les rendements d'actions émettrait comme signal une augmentation du taux de chômage et une diminution des bénéfices corporatifs. Ensuite, ceci mène empiriquement à des revenus imposables personnels et corporatifs moins élevés. Historiquement, les dépenses gouvernementales ne s'ajustent pas pour accommoder le changement

au niveau des revenus provenant de l'impôt, ce qui mène à une augmentation du déficit fédéral. La trésorerie réagit en accroissant les emprunts qu'elle effectue auprès du public. Une partie de cette dette est acquise par la banque centrale qui génère des fonds en amplifiant le taux de croissance de la base monétaire, ce qui mène inévitablement à une augmentation de l'inflation.

La conclusion logique serait donc que les investisseurs rationnels vont réaliser que la variation aléatoire du rendement des actions va entraîner cette chaîne d'événements. Donc, le prix de plusieurs actifs financiers à court terme est modifié pour tenir compte du signal envoyé par le rendement des actions ordinaires. Cette réaction explique le lien entre la variation des taux d'intérêt à court terme et l'inflation non anticipée. En effet, si la dette émise par la trésorerie n'est pas acquise par la banque centrale, on devrait observer une augmentation des taux d'intérêt réels à court terme. Donc, du même signal provenant des rendements d'actions ordinaires, on peut obtenir deux résultats différents représentant le même phénomène.

Pour ce qui est de la relation négative entre les rendements d'actions ordinaires et les taux d'intérêt à court terme de début de période, ce phénomène empirique serait partiellement en lien avec un renversement d'un modèle d'attentes adaptatif. Selon ce dernier, la différence entre un rendement d'actions et le taux d'intérêt à court terme de début de période causerait un changement négatif des taux des bons du Trésor. En renversant l'équation, on observe que le taux d'intérêt à court terme de début de période et la variation des taux des bons du Trésor causent la variation des prix d'actions. Les résultats empiriques sur trois décennies d'observations suggèrent que les rendements d'actions signalent un changement des taux d'intérêt nominaux et de l'inflation anticipée.

Suite à ces publications, plusieurs auteurs ont entrepris l'étude du pouvoir explicatif de plusieurs variables, tant dans l'objectif de fixer des prix d'actifs financiers que pour prévoir les rendements financiers futurs. Notamment, Chen, Roll et Ross (1986) ont effectué l'analyse de 17 variables macroéconomiques pour déterminer leur influence sur les rendements des marchés financiers. Cet article ne traite pas de prévision de rendements du marché, mais vise simplement à déterminer l'influence de variables macroéconomiques sur ces rendements. Les motivations des auteurs sont concentrées sur la fixation des prix, mais les relations observées par les auteurs ont été utilisées par plusieurs articles postérieurs traitant la prévision des rendements du marché. En effet, Chen, Roll et Ross (1986) ont évalué la relation d'un grand nombre de variables macroéconomiques avec les rendements du marché. Ils concluent que plusieurs variables peuvent potentiellement prédire les rendements du marché d'une période future.

Ainsi, en se basant sur les conclusions de Merton (1973), Cox et coll. (1985), et Ross (1976),

Chen, Roll et Ross (1986) ont, en effet, été en mesure de déterminer des variables d'état pouvant décrire l'économie et qui sont donc des sources de risque systématique. Plusieurs des variables macroéconomiques testées expliquaient significativement les rendements des actions. Notamment, on y retrouve la production industrielle, les variations de la prime de risque, les changements au niveau de pente de la courbe de taux, et l'inflation. Aux fins de cette recherche, l'élément le plus important des conclusions des auteurs est la confirmation que les rendements d'actions sont exposés aux nouvelles macroéconomiques et que leur prix dépend de l'ampleur de cette exposition.

Pendant la même période, Keim et Stambaugh (1986) effectuent l'analyse de plusieurs variables ex ante pour prédire des primes de risque ex post. Il s'agit d'un article se rapprochant considérablement de la méthodologie employée aujourd'hui pour tenter de prévoir cette prime. Il est question également d'un des premiers articles s'intéressant à la prévision de la prime de risque et, par conséquent, de la prévision des rendements d'actions. Le but des auteurs est de construire des variables pouvant refléter les niveaux de prix des actifs financiers et de déterminer si ces variables ont un pouvoir explicatif pour prévoir la prime de risque d'une variété d'actifs financiers. De plus, les auteurs vérifient si la saisonnalité des rendements financiers moyens d'actifs est associée avec des changements des espérances conditionnelles selon le niveau de prix de ces actifs. Notamment, les auteurs observent un coefficient significativement supérieur en janvier lors de la prévision de primes de risque de firmes de petite taille et d'obligations à faible grade. Aujourd'hui, ce phénomène est connu sous le nom de "l'effet janvier" et est habituellement inclus en tant que variable instrumentale dans les travaux de prévision de la prime de risque. Plusieurs des variables utilisées dans cet article sont toujours pertinentes aujourd'hui et sont employées dans cette recherche. La conclusion fondamentale de cet article est que les primes de risque espérées de plusieurs actifs fluctuent dans le temps et que ce phénomène est partiellement décrit par des variables décrivant le niveau de prix de ces actifs. Subséquemment à la publication de cet article, plusieurs biais ont été découverts par d'autres auteurs, mais il demeure que la relation avec le niveau des prix persiste et qu'elle doit être tenue en compte en effectuant la prévision de rendements financiers.

2.3 Études modernes sur la prévision des rendements financiers

Les travaux de Fama (1981), Geske et Roll (1983), Chen, Roll et Ross (1986), et Keim et Stambaugh (1986) décrits précédemment sont les plus importants sur le sujet de la prévision des rendements jusqu'au début des années 1990. Ce regroupement d'articles constitue les débuts des études de prévision de rendements financiers effectués de façon relativement semblable aux publications récentes de la littérature. Ils sont cités par la plupart des auteurs traitant du sujet

aujourd'hui et le réexamen des conclusions de ces articles fondamentaux est généralement inclus dans l'analyse empirique des relations présentes pendant la période observée. Ainsi, à partir de la fin des années 1980, on observe une tendance globale au niveau des études sur le sujet de la prévision des rendements, de sorte qu'une structure générale est employée par la majorité des auteurs jusqu'à ce jour. Cette structure, démontrée par l'équation (2.2), met en relation un rendement financier avec un facteur macroéconomique d'une période précédente. Généralement, le rendement financier étudié dans les études traitant de prévisions est la prime de risque d'un indice de marché général pour des raisons de simplification et de généralisation des résultats. Cette méthode est employée, afin de permettre une plus vaste possibilité d'utilisation pour des fins pratiques.

$$E(R_{mt}|Z_{t-1}) = \gamma_0 + \gamma_1 Z_{t-1} + \epsilon_t \quad (2.2)$$

Où :

R_{mt} : Prime de risque du marché m au temps t .

γ_0 : Constante du modèle.

γ_1 : Coefficient de la variable économique Z_{t-1} .

Z_{t-1} : Variable économique à caractère prévisionnel.

ϵ_t : Résidus du modèle au temps t .

Parmi les études modernes sur le sujet de la prévision de rendements financiers, on retrouve plusieurs articles notables publiés au cours des 20 dernières années. Tout d'abord, on note un ouvrage important de Fama et French (1988) qui ont démontré à l'aide de régressions que le rendement des dividende (dividend yield) prévoit une portion significative des rendements financiers du NYSE sur des intervalles de deux et quatre ans. Selon leurs résultats, pour ces intervalles, le rendement de dividende expliquerait 25% de la variance des rendements du marché en moyenne pour la période de 1927 à 1986. De plus, ils observent que le pouvoir explicatif de ceux-ci augmenterait avec l'horizon temporel des rendements financiers. En effet, au cours de la période observée, pour l'horizon de quatre ans, ils observent des proportions inférieures et supérieures de 19% et 64%. Ces résultats sont similaires aux études de Flood, Hodrick et Kaplan (1987), et Campbell et Shiller (1988) qui traitent du même sujet. Ces trois ouvrages utilisent des approches semblables au niveau de la définition des rendements de dividende et de leur méthodologie. Le rendement de

dividende est déterminé selon l'approche de Gordon (1962) :

$$P(t-1) = \frac{D(t)}{1+r} \left(1 + \frac{1+g}{1+r} + \frac{(1+g)^2}{(1+r)^2} + \dots \right) = \frac{D(t)}{r-g} \quad (2.3)$$

$$\frac{D(t)}{P(t-1)} = r - g \quad (2.4)$$

Où :

$P(t-1)$: Prix au temps $t-1$.

$D(t)$: Dividende au temps t .

g : Taux de croissance des dividendes.

r : Taux de rendement exigé (ajusté pour le risque).

Ainsi, la relation entre le taux de dividende $\left(\frac{D(t)}{P(t-1)}\right)$ et le taux de rendement exigé (r) suggère que le taux de dividende pourrait capturer des variations au niveau des rendements financiers espérés. Sur la base de cette fondation théorique, Fama et French (1988) ont été en mesure de montrer une relation entre le rendement de dividende et le rendement du marché en employant la régression suivante :

$$r(t, t+T) = \alpha(T) + \beta(T)Y(t) + \epsilon(t, t+T) \quad (2.5)$$

Où :

$r(t, t+T)$: Rendement futur du marché pour la période t pour un horizon temporel de T d'un mois, un trimestre, un an et quatre ans.

$\alpha(T)$: Constante du modèle pour un horizon temporel de T d'un mois, un trimestre, un an et quatre ans.

$\beta(T)$: Paramètre du taux de dividende sur le rendement futur du marché pour un horizon temporel de T d'un mois, un trimestre, un an et quatre ans.

$Y(t)$: Taux de dividende observé à la période t .

$\epsilon(t, T)$: Résidus du modèle au temps t pour l'horizon temporel de T d'un mois, un trimestre, un an et quatre ans.

Quelques années plus tard, Goetzmann et Jorion (1993) réfutent les résultats obtenus par Fama et French (1988) sur la base que leurs résultats sont affectés par plusieurs biais majeurs liés au fait que la nature persistante des paiements de dividendes fait en sorte que les mouvements des taux de dividende sont essentiellement dominés par les mouvements des prix. Ceci fait en sorte que les variables indépendantes de leur étude sont corrélées avec des variables dépendantes de périodes précédentes. De plus, l'utilisation des corrections GMM au niveau des erreurs types est seulement valide asymptotiquement et la périodicité utilisée pourrait affecter cette condition préalable. La principale solution employée par Goetzman et Jorion (1993) fut d'utiliser la méthodologie bootstrap de Efron (1979) pour modéliser les distributions des statistiques de la régression sous l'hypothèse que les rendements financiers sont indépendamment et identiquement distribués, et qu'ils ne sont pas liés à des dividendes passés. Ainsi, leurs résultats démontrent un important biais causé par les ajustements GMM et les simulations VAR de sorte que des résultats considérés fortement significatifs selon ces approches indiquent le contraire en employant un bootstrap. Donc, selon Goetzman et Jorion (1993), les taux de dividende ne prévoient pas significativement les rendements financiers et les ajustements GMM et les simulations VAR ne seraient pas appropriés dans certaines situations de vérification d'hypothèse malgré leur habileté à éliminer certains biais de petit échantillon et de chevauchement de données.

Les résultats de Kothari et Shanken (1997) indiquent que le taux de rendement de dividende prévoit de façon significative les rendements financiers du marché pour la période de 1926 à 1991 et pour la sous-période de 1941 à 1991. De plus, ils ont démontré que la mesure book-to-market (B/M) prévoit également de façon significative les rendements financiers du marché pour ces mêmes périodes. Il s'agit de l'un des premiers articles de la littérature à considérer la variable book-to-market comme variable instrumentale dans le contexte de prévision de rendements financiers. De plus, cet ouvrage est le premier à mettre en relation le pouvoir prévisionnel du taux de rendement de dividende et le ratio book-to-market. L'emploi de cette dernière variable dans ce type d'étude provient de son émergence en tant que concurrent solide comme déterminant de rendements espérés depuis Fama et French (1992, 1993). En effet, il a été démontré précédemment par Davis (1994), et Chan et coll. (1995) parmi d'autres que le ratio book-to-market explique significativement les variations en coupes transversales des rendements moyens. Le but de Kothari et Shanken (1997) est donc d'utiliser des ratios financiers, afin de prévoir des rendements financiers futurs en se basant sur deux points de vue de la littérature à l'égard du pouvoir prévisionnel des ratios financiers,

soit en séries temporelles et en coupes transversales. Le premier point de vue est que ces ratios sont affectés par le taux de rendement exigé auquel les flux monétaires sont actualisés et que tout élément faisant accroître ce taux (tel que la liquidité du titre) fera une pression réductrice sur la valeur au marché du titre. Ainsi, on s'attend à une corrélation positive entre le rendement exigé et les ratios financiers. Le point de vue alternatif sur le sujet est que les ratios financiers reflètent à quel point le marché est surévalué (ratios inférieurs) ou sous-évalué (ratios supérieurs) à un moment dans le temps. Kothari et Shanken (1997) ont donc employé une méthode de vecteur autorégressif (VAR) pour déterminer la signification statistique du pouvoir prévisionnel des ratios financiers (dividend yield et B/M) de la période $t-1$ à l'égard des rendements financiers agrégés du marché de la période t dans le cadre d'une régression OLS. La procédure de bootstrap a ensuite été employée pour confirmer ce pouvoir prévisionnel à l'extérieur de l'échantillon (OOS). L'innovation additionnelle des auteurs en comparaison avec Goetzman et Jorion (1993) est l'utilisation d'un bootstrap bayésien où la croyance de l'investisseur envers la probabilité de rendements négatifs du marché peut être ajustée de façon à influencer la régression. Leurs conclusions furent que la relation entre les rendements financiers futurs et le ratio B/M fut plus forte que le taux de rendement de dividende pour la période de 1926 à 1991. Cependant, le contraire est vrai pour la période de 1941 à 1991. Cette conclusion semble indiquer que la relation entre ces variables est portée à évoluer dans le temps avec une augmentation de l'influence du taux de rendement de dividende.

Kothari et Shanken (1997) rencontrent plusieurs biais dans leur étude qui méritent une attention particulière aux fins de cette recherche. En effet, ces derniers notent une possibilité de "data mining" étant donné la grande volatilité de la relation entre les deux variables instrumentales de l'étude et les rendements financiers futurs. Le fait que leur pouvoir prévisionnel varie considérablement d'une sous-période à l'autre devrait modérer la valeur associée aux prévisions générées par le modèle et devrait être tenu en compte lors de l'élaboration de décisions d'investissement. Les auteurs font également mention d'un "survivorship bias", de sorte que la pente et l'origine du modèle seraient biaisées à la hausse. Cependant, ils démontrent qu'un ajustement pour le "survivorship bias" viendrait réduire le niveau de signification de leur résultat de non-négativité des rendements espérés pendant la période observée. D'autre part, plus grand est l'influence de ce biais, plus fort est le rejet de non-négativité des rendements espérés pendant la période observée. Il est donc nécessaire d'effectuer les ajustements nécessaires pour corriger ce biais, afin d'accroître la fidélité des résultats. Par ailleurs, ils observent que le choix des variables instrumentales est nécessairement un biais affectant l'étude, étant donné l'absence de recherche pour déterminer les meilleures variables d'état à caractère prévisionnel. Le taux de rendement du dividende a été choisi purement pour sa prééminence dans la littérature financière et le ratio book-to-market a été considéré principalement en raison de son statut de célébrité en tant qu'élément explicatif de la variation en coupes

transversales des rendements financiers, tel que démontré par Fama et French (1992, 1993).

D'autre part, Lamont (1998) propose un modèle de prévision basé sur le modèle de croissance de dividende dynamique de Campbell et Shiller (1988) de façon à effectuer un lien entre le niveau de prix, les rendements financiers, les bénéfices et le taux de rendement de dividende. Il est question de confirmer que la variabilité des bénéfices mis en évidence par Fama et French (1988) n'est pas du bruit blanc, mais démontre un lien avec les rendements espérés. Cette étude utilise une approximation loglinéaire d'une identité comptable,

$$p_t = E \left(\sum_{j=1}^{\infty} \rho^j [(1 - \rho)d_{t+1+j} - r_{t+1+j}] + \kappa_t \right) \quad (2.6)$$

Où :

p_t : Logarithme naturel du S&P Composite Index au temps $t + 1 + j$.

d_{t+1+j} : Logarithme naturel de la somme des dividendes par action des quatre derniers trimestres au temps $t + 1 + j$.

r_{t+1+j} : Taux de rendement excédentaire du marché au temps $t + 1 + j$.

j : Nombre de périodes entre le prix actuel et la période de prévision moins un.

κ_t : Résidus du modèle au temps t .

ρ^j : Facteur d'actualisation basé sur le taux sans risque (Treasury-Bill Rate).

On définit ensuite $p_{r,t}$ comme la somme actualisée des rendements futurs à partir de la prochaine période et $p_{d,t}$ comme la somme actualisée des dividendes futurs à partir de la prochaine période, tel que :

$$p_{r,t} = E \left[\sum_{j=1}^{\infty} \rho^j r_{t+1+j} \right] \quad (2.7)$$

$$p_{d,t} = E \left[\sum_{j=0}^{\infty} \rho^j (1 - \rho) d_{t+1+j} \right] \quad (2.8)$$

Ainsi, il est possible d'exprimer le rendement espéré d'une période en combinant les équations (2.7) et (2.8) sous la forme suivante :

$$E[r_{t+1}] = -p_t + p_{d,t} - p_{r,t} + \kappa \quad (2.9)$$

Cependant, il n'est pas possible d'évaluer l'équation (2.9) étant donné qu'il s'agit d'une identité comptable contenant des espérances non observables. De plus, étant donné que les flux monétaires d'une firme croissent avec le temps, p_d et p ne sont pas stationnaires. Cette équation peut plutôt être utilisée pour interpréter une régression prévisionnelle de rendements par rapport à des prix retardés et des proxy des variables p_r et p_d retardées. En excluant le taux de rendement de dividende (p_d) du prix p il serait possible d'effectuer une régression basée sur cette variable, étant donné la nature relativement constante des dividendes dans le temps, décrite par Lintner (1956), basée sur une fraction des bénéfices de l'entreprise. Cependant, Miller et Modigliani (1961) démontrent que la politique de dividende des entreprises pourrait être établie de n'importe quelle façon arbitraire. Ainsi, la politique de dividende et la capacité de prédiction des dividendes pourraient de ne pas être stables dans le temps. La solution de l'auteur à cette problématique est d'introduire une variable n au modèle (étant le taux de dividende) qui fera en sorte que $p - n$ sera stationnaire. Ceci fera en sorte que les trois variables explicatives du modèle représenté par l'équation (2.10) seront stationnaires.

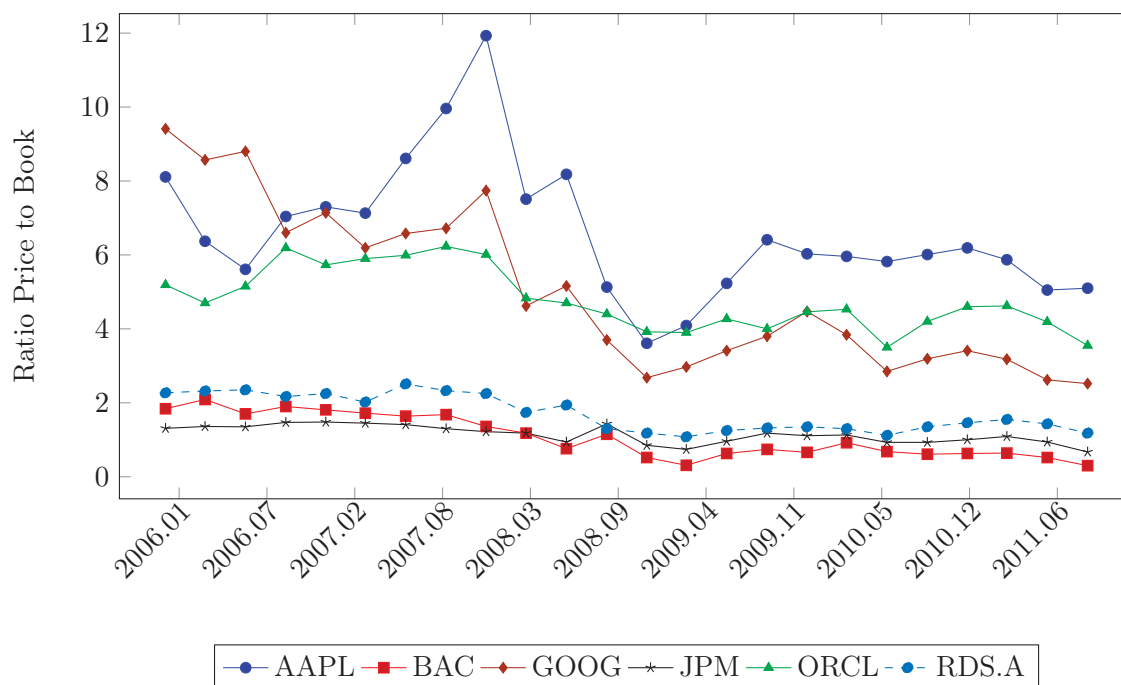
$$E_t[r_{t+1}] = -(p_t - n_t) + (p_{d,t} - n_t) - p_{r,t} + \kappa \quad (2.10)$$

Les conclusions de cette étude sont que la variable à ajouter au modèle serait le ratio de distribution de dividendes parce que les dividendes et les bénéfices ont tous deux une capacité de prédiction individuelle. De plus, ils constatent que le bruit au niveau des bénéfices trimestriels contient de l'information concernant les mouvements à court terme des rendements espérés. Des prix actuels élevés et des bénéfices actuels élevés prévoient de faibles rendements futurs. Ainsi, l'utilisation de rendements de bénéfices en tant que variable à caractère prévisionnel de rendements futurs serait biaisée à cause de ce contenu informatif. Cependant, un taux de dividende élevé prévoit des rendements financiers futurs élevés, ce qui confirme la supériorité de cette variable au niveau de la prévision. Ces résultats sont valides à court terme, mais perdent de leur efficacité à long terme, de sorte qu'ils ne sont pas importants ni significatifs dans le contexte de prévision à long terme. Selon ces résultats, les prévisions inférieures à long terme des rendements du marché américain au milieu des années 1990 seraient dues majoritairement au niveau des prix qui furent relativement

élevés en comparaison avec les ratios financiers de référence habituels (P/S, P/E, P/B). Il est donc crucial de retenir de cette étude que le niveau des prix affecte significativement la prévision à long terme et qu'il est nécessaire d'en tenir compte dans ce type d'analyse, et ce, surtout dans le contexte d'industries où le niveau des prix est généralement élevé en comparaison avec les ratios comptables de référence. Un exemple flagrant de ce type d'industrie est le secteur des technologies de l'information où le niveau des prix est extrêmement éloigné de plusieurs références comptables. La figure 1 ci-dessous met en évidence l'écart entre trois entreprises de cette industrie par rapport à des comparables ayant un prix plus rapproché de leur valeur comptable. En effet, les ratios *Price-to-Book* moyens historiques de *Apple Inc.* (AAPL), *Google Inc.* (GOOG) et *Oracle Corporation* (ORCL) sont significativement plus élevés que ceux de *Bank of America Corporation* (BAC), *JP Morgan Chase & Co.* (JPM) et *Royal Dutch Shell* (RDS.A). L'ex-président de la réserve fédérale des États-Unis Alan Greenspan fait référence à cet écart significatif entre la valeur intrinsèque de ces titres technologiques et leur prix sur le marché. Il surnomme ce phénomène une "exubérance irrationnelle" des investisseurs qui gonfle le prix des titres loin au-delà de leur valeur comptable.

FIGURE 1 – Comparaison du ratio Book-to-Market de six entreprises

Cette figure présente le ratio *Price-to-Book* moyens historiques de *Apple Inc.* (AAPL), *Google Inc.* (GOOG) et *Oracle Corporation* (ORCL) sont significativement plus élevés que ceux de *Bank of America Corporation* (BAC), *JP Morgan Chase & Co.* (JPM) et *Royal Dutch Shell* (RDS.A).



2.4 Récentes études de prévision de rendements financiers

Au cours des dix dernières années, plusieurs études furent publiées sur la prévision des rendements financiers. Certaines confirment la présence de variables prévoyant significativement les rendements du marché, tandis que d'autres réfutent cette relation. Cette section se concentre sur les récents articles de la littérature qui ont un impact significatif sur le processus méthodologique de cet ouvrage. Certains sont mis en évidence pour leur innovation au niveau des nouvelles variables démontrant une prévisibilité significative, d'autres pour l'ampleur du répertoire de variables instrumentales employées et, enfin, certaines pour leurs innovations méthodologiques. L'ensemble des publications citées dans cette section fait l'objet de la base de la méthodologie employée dans cette étude empirique.

Baker et Wurgler (2000) mettent en évidence dans leur étude que les firmes du marché ont tendance à préférer le financement par actions avant des périodes à faible rendement financier et de préférer le financement par la dette avant des périodes à fort rendement financier. Les auteurs n'ont pas été en mesure d'associer cette relation à un quelconque mécanisme de façon empirique. Leur conclusion est que les marchés financiers seraient inefficients à ce niveau et que les gestionnaires exploiteraient cette inefficience lorsqu'ils effectuent leurs décisions de financement. Selon leurs résultats, la capacité à anticiper le marché serait un déterminant significatif des actions managériales, de sorte que les rendements financiers eux-mêmes seraient affectés par ces actions. En effet, on observe dorénavant qu'elle est une composante fréquente dans la majorité des études en matière de prévision de rendements financiers.

Welch et Goyal (2008), dans une étude défavorable à l'égard de la performance empirique de la prévision de la prime de risque, font usage des variables concernant l'émission d'actions et de dette de Baker et Wurgler (2000). Malgré le biais probable de "self fulfilling prophecy" dans cet ouvrage, les variables en question démontrent les meilleurs résultats à l'égard de la prévision de la prime de risque de marché américaine. Celles-ci se démarquent parmi un regroupement de 20 variables testées par les auteurs. Ces dernières sont présentées dans la table 1 ci-dessous. Selon leurs conclusions, ces modèles de prévision individuels souffriraient d'incertitude et d'instabilité, de sorte que leur pouvoir de prévision ne serait pas utile à un gestionnaire. L'implication de cette étude dans le contexte de ce mémoire se résume à la confirmation de la relation démontrée par Baker et Wurgler (2000) et au choix de variables instrumentales employées par les auteurs (à l'exception de la variable "kitchen sink" qui n'est d'aucune utilité pratique aux fins de cette étude). La méthodologie employée par Welch et Goyal (2008) a été réfutée par plusieurs auteurs prenant la défense de la capacité de prévision des rendements. Notamment, Campbell et Thompson (2008) réfutent

les résultats de Welch et Goyal (2008) en affirmant qu'il est possible de surpasser la moyenne historique avec plusieurs modèles de prévision univariés en imposant une restriction au niveau du signe des rendements prévus. En imposant une restriction sur le signe des coefficients et des prévisions de rendement, Campbell et Thompson (2007) ont été en mesure d'obtenir des prévisions de rendements significatives à l'intérieur et à l'extérieur de l'échantillon (IS et OOS). D'autre part, Cochrane (2008) obtient des résultats empiriques qui montrent que si les rendements du marché ne sont pas prévisibles, la croissance des taux de dividende doit l'être. Selon ses résultats, cette relation doit être vraie pour que les ratios dividende-prix observés soient possibles. Il observe ensuite que la croissance des taux de dividende n'est pas prévisible au cours des 25 dernières années en fonction du ratio dividende-prix. D'autre part, il obtient comme résultat que les rendements excédentaires du marché sont prévisibles sur cette même période. Il en conclut donc que, contrairement à ce que la théorie de portefeuille enseigne, les marchés peuvent être pronostiqués (*Market Timing*). Cette conclusion découle du fait que ses résultats montrent que les rendements financiers ne sont pas liés au taux de croissance de dividende et qu'un tiers élément serait responsable des variations de prix. L'auteur fait un lien avec les variables d'état de Merton (1973) en stipulant que la présence de ce tiers élément affecte l'hypothèse de stabilité des opportunités d'investissement présente dans la littérature. La capacité de prévision découle de la variation de ces opportunités d'investissement, donc le lien entre le modèle intertemporel de Merton (1973) et la prévision de la prime de risque est très important.

TABLEAU 1 – Variables étudiées par Welch et Goyal (2008)

Ce tableau énumère les variables économiques testées par Welch et Goyal (2008) pour prédire la prime de risque du marché américain. La variable *Percent Equity Issuing* est la plus significativement performante pendant la période testée.

Variable	Description
Stock Returns	S&P 500 index returns from 1926 to 2005
Risk-free Rate	Treasury-bill rate from 1920 to 2005
Dividends	12-month moving sum of div. paid on S&P 500
Dividend Price Ratio	Difference between log of div. and log of prices
Dividend yield	Difference between log of div. and log of lagged prices
Earnings	12-month moving sum of earnings on S&P 500
Stock Variance	Sum of squared daily returns on S&P 500
Cross-Sectional Premium	Relative valuations of high and low-beta stocks
Book-to-Market Ratio	Ratio of book value to market value
Net Equity Expansion	Ratio of issued stocks on total market cap of NYSE
Percent Equity Issuing	Ratio of equity issues on total issuing activity
Treasury Bills	T-Bill rates
Long Term Rate of Returns	U.S. Yield on Long-Term United States Bonds series
Term Spread	Difference between long term yield on gov bonds and T-Bill
Default Yield Spread	Difference between BAA and AAA yields
Default Return Spread	Difference between long-term corp. bond and l-t gov. bond returns
Inflation	Consumer Price Index from 1919 to 2005
Investment to Capital Ratio	Ratio of aggregate investment to aggregate capital for whole economy
Kitchen Sink	Includes all above variables

Campbell et Yogo (2007) évaluent la validité des tests conventionnels utilisés pour évaluer la prévision de rendements financiers. L'hypothèse nulle de prédictibilité serait rejetée trop fréquemment lorsque la variable prédictive est persistante et fortement corrélée avec les rendements financiers. Ils observent que le test conventionnel (t-test) serait invalide pour les ratios dividende/prix et bénéfice net/prix. Cependant, ils observent des cas valides de prédictibilité pour les taux d'intérêt à court terme et l'écart de rendement *long-court*.

D'autre part, Rapach, Strauss et Zhou (2010) étudient le pouvoir prévisionnel des rendements aux États-Unis en tant que facteur de prévision pour les rendements futurs de onze pays industrialisés. Ils observent que les rendements des États-Unis ont une performance supérieure aux variables

économiques testées par les auteurs en tant que variable instrumentale. Ils concluent que les États-Unis jouent un rôle de leader sur le plan international au niveau des rendements financiers et que les variations de ceux-ci ont un impact significatif sur les rendements futurs des autres pays industrialisés étudiés.

Les deux articles centraux à cette étude au niveau de leur méthodologie sont les articles de Rapach, Wohar et Rangvid (2005), et Chrétien et Coggins (2011). Leurs approches sont très similaires et représentent les méthodes de prévision les plus modernes de la littérature financière. Les tests statistiques employés sont les plus dépourvus de biais pour ce type d'étude, que ce soit par l'utilisation du test statistique de McCracken (2007) ou des erreurs types de Newey et West (1987). La procédure de bootstrap utilisée est identique à celle employée par Kothari et Shanken (1997) et reflète la méthode qui est mise en application pour la prévision de rendements à l'extérieur de l'échantillon observé. Cette procédure est décrite dans la section 3.6 du présent document. Les deux articles sont innovateurs en leur propre sens et tous deux confirment la présence de prédictibilité significative pour certaines variables économiques.

Dans le cas de Rapach, Wohar et Rangvid (2005), leur innovation provient de la couverture internationale de leur ouvrage. Ces derniers examinent la prédictibilité des rendements financiers dans douze pays industrialisés. Dans la majorité des cas, les taux d'intérêt sont les variables à caractère prévisionnel les plus consistantes et fiables dans ce regroupement de pays. Les variables prévoyant significativement la prime de risque de marché pour chaque pays étudié sur l'horizon d'un mois sont énumérées dans la table 2 ci-dessous. Les descriptions relatives à ces variables économiques se trouvent dans la section *Données* du présent document.

TABLEAU 2 – Résultats abrégés de Rapach, Wohar et Rangvid (2005) (horizon de 1 mois)

Ce tableau énumère les variables les plus significativement performantes pour prédire les primes de risque de marché de chaque pays étudié par Rapach, Wohar et Rangvid (2005).

Pays	Variable(s) significative(s)
Belgium	Relative government bond yield
Canada	Relative government bond yield
Denmark	Relative money market rate Relative government bond yield Inflation Rate
France	Relative government bond yield
Germany	Relative government bond yield
Italy	Relative Treasury bill rate Relative government bond yield
Japan	Broad money growth
Netherlands	Relative government bond yield Inflation Rate
Norway	Relative government bond yield
Sweden	Relative money market rate Relative Treasury bill rate
U.K.	Relative government bond yield
U.S.	Relative government bond yield Inflation Rate

Pour ce qui est de Chrétien et Coggins (2011), leur innovation se trouve au niveau du marché étudié et du répertoire de variables d'information employées. Il s'agit de la première étude effectuant un examen approfondi de variables d'information et de la prévision de la prime de risque au Canada. Cet article fait l'usage de 36 variables d'information, dont plusieurs spécifiques au Canada. Les variables prévoyant significativement la prime de risque canadienne au seuil $\alpha = 0.05$ à l'intérieur et à l'extérieur de l'échantillon sont résumés dans la table 3 ci-dessous. Ce tableau inclut seulement les résultats pour la période complète étudiée par les auteurs. Les descriptions relatives aux variables économiques se trouvent dans la section 5 du présent document.

TABLEAU 3 – Variables significatives IS et OOS dans Chrétien et Coggins (2011) ($\alpha = 0.05$)

Ce tableau énumère les variables significatives à l'intérieur et à l'extérieur de l'échantillon au seuil $\alpha = 0.05$ dans l'étude de Chrétien et Coggins (2011).

Variable	In-Sample	Out-Of-Sample
Previous Equity Premium	✓	✓
Januray Dummy	✓	
T-Bill Yield	✓	
T-Bill Yield Variation	✓	✓
T-Bill Yield Relative	✓	
Long Gov Bond Yield	✓	
Long Gov Bond Yield Variation	✓	✓
Long Gov Bond Yield Relative	✓	✓
Term Premium	✓	
GDP Growth	✓	✓
Prime Rate	✓	
Leading Indicator Growth	✓	✓

En résumé, la littérature académique sur le sujet de la prévision de rendements financiers est très diverse. Plusieurs innovations ont été effectuées depuis les premiers articles sur le sujet, afin d'améliorer la précision, la validité et la stabilité des prévisions effectuées. Certains innove au niveau du nombre de variables économiques testées et sur la façon de les mesurer, tel que Chen, Roll et Ross (1986), Keim et Stambaugh (1986) et Baker et Wurgler (2000). D'autres améliorent la méthodologie entourant le modèle de prévision utilisé, tel que Goetzmann et Jorion (1993), Kothari et Shanken (1997) et Lamont (1998). Les articles de la littérature récente mettent l'emphase, généralement, sur la découverte de nouvelles variables prédictives et sur l'amélioration du modèle de prévision, afin de réduire le nombre de facteurs pouvant affecter la précision, la validité et la stabilité des prévisions effectuées.

Les résultats obtenus par ces auteurs sont également très divers. Certains observent la présence de prévision significative de rendements financiers, tandis que d'autres observent le contraire. Parmi ceux du premier groupe on observe plusieurs variables économiques effectuant une prévision significativement supérieure aux autres. La littérature financière sur ce sujet est très loin de s'approcher d'un consensus et il s'agit d'une motivation additionnelle pour découvrir de nouvelles approches méthodologiques. Un élément notable de la littérature observant la présence de prévision est le très faible pouvoir explicatif de la prévision à l'égard des rendements financiers futurs. Très souvent,

le coefficient de détermination des variables jugées supérieures parmi l'échantillon testé ne franchit pas 0.10. Donc, la prévision ne permet d'obtenir de l'information que sur moins de 10% des rendements futurs. Il y a donc un potentiel évident d'accroître le pouvoir explicatif des prévisions générées par le modèle en améliorant sa précision.

3 Objectifs

Ce mémoire comprend un objectif principal et six objectifs secondaires. Chacun d'entre eux est énuméré ci-dessous.

Objectif principal

Prédire les primes de risque de marché pour les États-Unis et le Canada sur un horizon d'un mois en se basant sur les données économiques des 20 dernières années. Tenter d'effectuer des prévisions plus précises que celles générées par les modèles précédemment utilisés dans la littérature, afin d'atteindre un pouvoir explicatif plus élevé.

Objectifs secondaires

1. Évaluer la performance des modèles de prévision de la prime de risque basés sur la moyenne historique, afin d'utiliser ce modèle comme élément de comparaison pour tous les modèles subséquemment testés dans cette étude.
2. Évaluer la performance des modèles univariés de prévision de la prime de risque agrégée. Cet objectif vise à prévoir la prime de risque de marché de chaque pays dans sa totalité avec une seule variable économique. Ce modèle reflète la méthodologie la plus récente de la littérature, afin de pouvoir comparer le modèle univarié par agrégation de prévisions sectorielles de cette étude avec la méthodologie existante.
3. Évaluer la performance des modèles multivariés de prévision de la prime de risque agrégée. Cet objectif vise à prévoir la prime de risque de marché de chaque pays dans sa totalité en combinant plusieurs variables économiques. Ce modèle est une version multivariée de la méthodologie du modèle de prévision univarié employé par la littérature financière récente. Ce modèle multivarié, obtenu à l'aide d'une régression *Stepwise*, est utilisé afin de pouvoir comparer le modèle de prévision multivarié par agrégation de prévisions sectorielles de cette étude avec la méthodologie existante.
4. Évaluer la performance des modèles multivariés de prévision des primes sectorielles qui seront cumulées pour prévoir la prime de risque agrégée. Chacun des rendements sectoriels est prévu via un modèle de prévision optimal selon son pouvoir explicatif. Le modèle effectuant les meilleures prévisions pour chaque regroupement industriel est choisi, peu importe s'il est univarié ou multivarié. Ensuite, une somme pondérée de toutes les prévisions sectorielles est effectuée, afin de prédire la prime de risque de marché.
5. Effectuer une analyse comparative de ces quatre méthodes de prévision de la prime de risque.

Cette comparaison est effectuée selon l'erreur de prévision quadratique moyenne de chaque modèle. Celui ayant la valeur la plus faible est considéré comme ayant le meilleur pouvoir explicatif à l'égard des rendements financiers à prédire. Il s'agit donc du modèle effectuant les meilleures prévisions de la prime de risque de marché du pays à l'étude. Ce modèle est ensuite comparé à chaque modèle obtenu lors des étapes précédentes afin de confirmer qu'il est bien significativement supérieur à ceux-ci. Ces deux étapes sont représentées graphiquement par les figures 2 et 3 respectivement.

4 Méthodologie

La base méthodologique de cette recherche s'inspire des approches employées par Rapach, Wohar et Rangvid (2010), et Chrétien et Coggins (2011). Les méthodes utilisées par ces études pour prédire la prime de risque du marché sont très semblables. Il s'agit des articles de la littérature proposant la méthodologie la plus récente sur le sujet. Donc, les éléments concernant le modèle de prévision univarié et l'ensemble des tests statistiques s'y référant sont tirés de ces deux publications de référence. De plus, la méthode d'élargissement de la taille de l'échantillon, soit le Bootstrap de Efron (1979), provient également de ces références fondamentales.

Le cadre de cette recherche se limite à la prévision de primes de risque agrégées et spécifiques à des regroupements d'industries des marchés canadiens et américains. La prévision de ces primes de risque s'effectue de façon mensuelle et, pour le marché américain, comprend une période d'observation de janvier 1991 à octobre 2011, soit 20 ans ou, plus précisément, 250 observations. Pour le marché canadien, en raison d'accès plus limité à l'information historique, le nombre d'observations est plutôt de 238. Il est question d'évaluer et de comparer les prévisions obtenues selon les diverses méthodes suivantes :

1. Modèles de prévision de la prime de risque basés sur la moyenne historique.
2. Modèles univariés de prévision de la prime de risque agrégée.
3. Modèles multivariés de prévision de la prime de risque agrégée (*backward* et *forward*).
4. Agrégation de modèles multivariés de prévision des primes sectorielles pour prévision de la prime de risque agrégée.

Le processus méthodologique de cette étude est résumé graphiquement par les figures 2 et 3 de la présente section.

FIGURE 2 – Vue d'ensemble de la méthodologie

Cette figure est une représentation graphique de la comparaison initiale de chaque modèle avec le modèle de prévision de la prime de risque basé sur la moyenne historique, tel que présenté à la section 4.

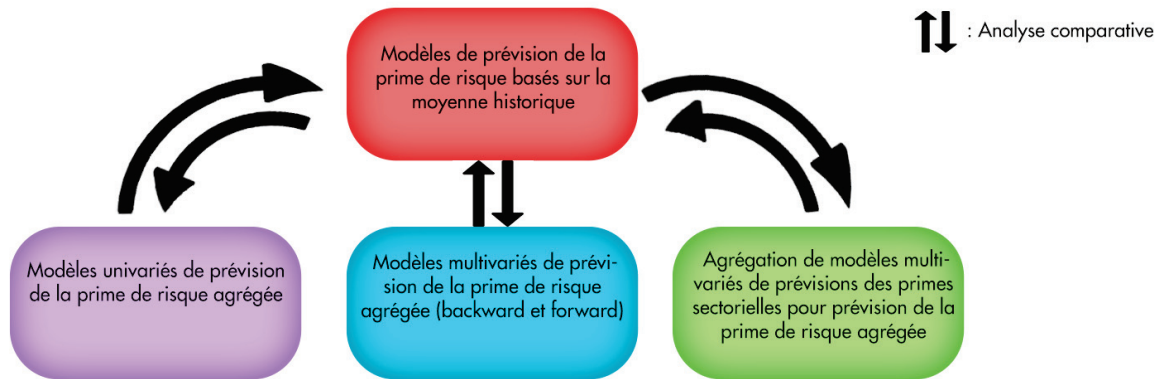
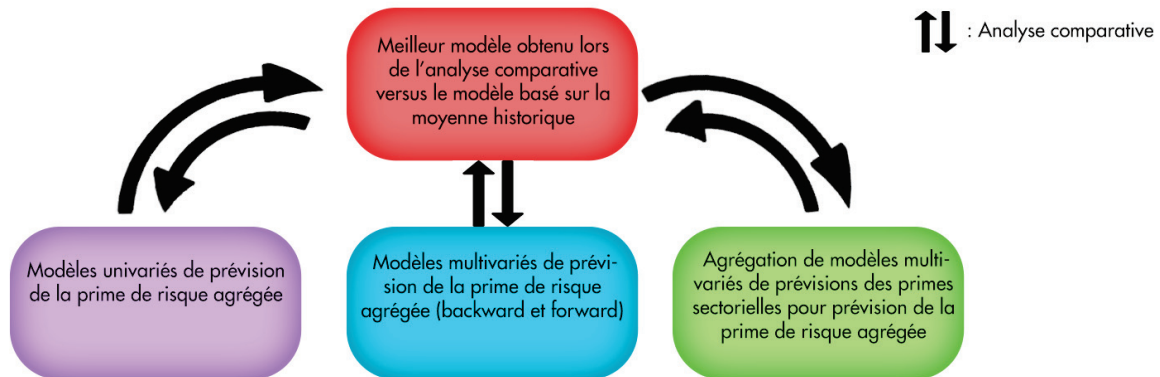


FIGURE 3 – Vue d'ensemble de la méthodologie

Cette figure est une représentation graphique de la comparaison initiale de chaque modèle avec le modèle de prévision de la prime de risque basé sur la moyenne historique, tel que présenté à la section 4.



4.1 Modèles de prévision de la prime de risque basés sur la moyenne historique

Le modèle de prévision employant la moyenne historique, aussi appelé le modèle naïf est la méthode la plus simple pour prédire des rendements financiers. Dans le cadre de cette étude, ce modèle sert de référence pour situer la performance des autres modèles évalués. Évidemment, un modèle qui n'est pas en mesure d'effectuer de meilleures prévisions qu'une moyenne historique des rendements passés n'est pas d'une très grande utilité pratique. Donc, l'inclusion de ce *benchmark* est essentielle pour déterminer la véritable valeur ajoutée de chacune des autres méthodes proposées. La formule (4.1) illustre cette méthode mathématiquement dans un contexte In-Sample. La formule (4.2) illustre cette même méthode dans un contexte Out-of-Sample.

$$EQP_{m,t} = \sqrt[N]{\prod_{x=1}^N EQP_{m,t-x}} \quad (4.1)$$

$$EQP_{m,t} = \sqrt[N-t]{\prod_{x=1}^{t-1} EQP_{m,t-x}} \quad (4.2)$$

Où :

$EQP_{m,t}$: Prime de risque espérée à la période t .

$EQP_{m,t-x}$: Prime de risque du marché m observée au temps $t - x$.

x : Nombre de mois de recul. (pour $x = 1 \dots N$)

N : Nombre total d'observations. 250 pour le marché américain et 238 pour le marché canadien.

La moyenne géométrique est employée pour réduire l'impact des observations s'éloignant de la moyenne. Ainsi, la tendance centrale des primes de risques observées sera plus constante.

4.2 Modèles univariés de prévision de la prime de risque agrégée

Les modèles univariés de prévision de la prime de risque agrégée étudiés se réfèrent à la méthode employée par la majorité des articles précédents de la littérature financière. Il s'agit du modèle prévoyant la prime de risque de marché directement via une variable économique observée pendant la période précédent la prévision. En effet, on peut, notamment, retrouver ce modèle dans les études de Chrétien et Coggins (2011), et Welch et Goyal (2008). La formule (4.3) ci-dessous, tirée de Chrétien et Coggins (2011), représente l'expression générale de cette méthode. L'ensemble des variables économiques énumérées dans la section données sont testées pour prédire la prime de risque des États-Unis et du Canada avec cette méthode.

$$EQP_t = \beta_0 + \beta_{1,i} Z_{i,t-1} + \epsilon_t \quad (4.3)$$

Où :

EQP_t : Prime de risque observée à la période t . Pour le marché américain, chaque observation représente le facteur mensuel $Rm - Rf$ du modèle Fama French, tel qu'obtenu sur le site web de Kenneth R. French. Pour le marché canadien, chaque observation représente la variation mensuelle de l'indice composite du TSE additionné au rendement mensuel de dividende des titres du TSE. À cette somme est soustrait le taux mensuel des bons du Trésor ayant une échéance d'un mois représentant le taux sans risque. Les données canadiennes proviennent des tables 176-0047 et 176-0043 de la base de données CANSIM.

$Z_{i,t-1}$: Variable économique i observée à la période $t - 1$. Les variables économiques utilisées sont énumérées et décrites à la section 5.3 du document.

β_0 : Constante du modèle.

$\beta_{1,i}$: Bêta exprimant le lien entre la prime de risque espérée et la variable économique i .

ϵ_t : Résidus du modèle au temps t .

Afin de vérifier la stabilité de chaque modèle et d'éliminer la possibilité de biais en lien avec la taille de l'échantillon et le chevauchement d'observations, chaque modèle estimé dans cette étude est soumis à la procédure de Bootstrap de Efron (1979). L'emploi de cette technique résulte en une meilleure précision et une plus grande stabilité des résultats et de leur niveau de signification.

La p-value du Wald test est utilisée pour déterminer le niveau de signification de chaque modèle.

4.3 Modèles multivariés de prévision de la prime de risque agrégée

Les modèles multivariés de prévision de la prime de risque agrégée sont une combinaison de variables économiques utilisées pour prédire la prime de risque dans sa totalité. Cette méthode consiste à sélectionner des variables utilisées dans le type de modèle précédent et de les combiner, afin d'obtenir des prévisions significatives. La formule (4.4) ci-dessous, basée sur le modèle univarié, représente l'expression générale de cette méthode de prévision.

$$EQP_t = \beta_0 + \sum_{i=1}^N \beta_{1,i} Z_{i,t-1} + \epsilon_t \quad (4.4)$$

La construction de ce modèle s'effectue selon une régression *Stepwise* de Draper & Smith (1998). Ce processus consiste à construire un modèle de régression en effectuant une sélection automatique des variables indépendantes. La sélection s'effectue en testant une variable indépendante à la fois et en incluant celles qui sont statistiquement significatives dans le modèle multivarié (*forward selection*). Il est également possible d'exécuter ce triage en incluant toutes les variables indépendantes dans le modèle multivarié, puis en excluant toutes celles qui ne sont pas statistiquement significatives (*backward selection*). Pour les fins de cette étude, les deux approches sont employées. Dans le cas où le modèle de régression généré est différent d'une méthode à l'autre, tous deux sont conservés afin de participer à l'analyse comparative des différents modèles de prévision. La p-value du Wald test est utilisée pour déterminer le niveau de signification de chaque modèle.

4.4 Agrégation de modèles multivariés de prévision des primes sectorielles

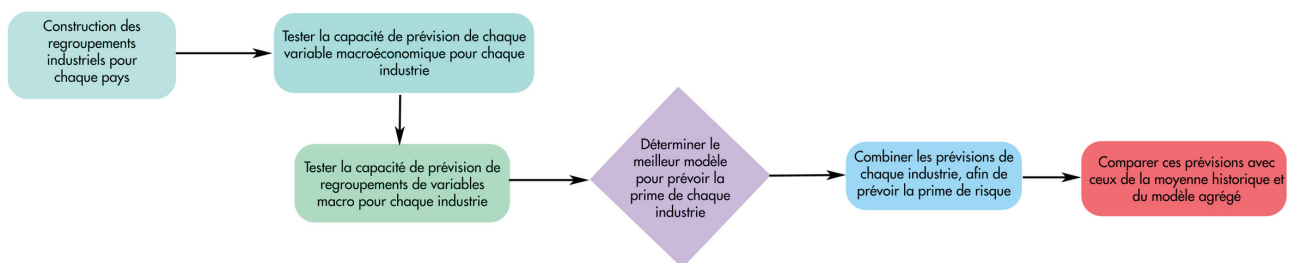
Les modèles de régression univariés par agrégation de prévisions sectorielles sont issus d'un processus tentant de prévoir la prime de risque de marché en cumulant les prévisions de l'ensemble des regroupements industriels composant le marché. Ce processus est résumé par la figure 4 ci-dessous. Les étapes de ce processus se résument comme suit :

1. Construire des regroupements industriels.

2. Tester chaque variable économique pour prédire chaque regroupement industriel de façon univariée.
3. Construire et tester des modèles multivariés de prévision pour chaque regroupement industriel.
4. Comparer les modèles générés aux étapes 2 et 3.
5. Choisir un modèle par regroupement industriel.
6. Regrouper les prévisions générées pour chaque regroupement industriel, afin de prédire la prime de risque de marché.

FIGURE 4 – Vue d’ensemble de la méthodologie des modèles par agrégation de prévisions sectorielles

Cette figure est une représentation graphique de chaque étape de la méthodologie des modèles par agrégation de prévisions sectorielles de cette étude, tel que présenté à la section 4.4.



4.4.1 Regroupements industriels

Les regroupements industriels utilisés dans cette étude sont déterminés selon les deux premiers chiffres de leur code SIC (Standard Industrial Classification). Cette approche est similaire à celles utilisées par Sweeney and Warga (1986), Carleton and Lakonishok (1986), Breeden et al. (1989), Shanken (1990) et Ferson and Harvey (1991). Cependant, pour réduire le nombre de regroupements industriels, certaines industries sont combinées en fonction de leurs similarités. Cette fusion est effectuée via un regroupement des industries ayant une corrélation de 0.7 (*Très forte*) entre elles pendant la période totale observée.

Les industries de départ étant établies par les deux premiers chiffres du code SIC de chaque entreprise, les États-Unis et le Canada débutent respectivement avec 63 et 61 industries. Une description de chaque industrie de départ est disponible à la table 34 de l'Annexe B. Les rendements de tous les titres des bourses de chacun des pays (NYSE, Amex et NASDAQ pour les É-U et TSE pour le Canada) sont d'abord calculés. Les versements de dividendes augmentent le rendement du mois de versement de l'ordre du taux de dividende ($\frac{Dividende(\$)}{Prix}$). Cette première étape est ignorée pour les Appels Publics à l'Épargne (*Initial Public Offerings*), de sorte que si le prix d'un titre est nul pour une période d'observation, le rendement de la prochaine période sera également nul, afin d'éliminer cet impact. Ceci est nécessaire afin d'assurer que les rendements d'industries soient comparables entre chaque période. Par la suite, ces rendements sont multipliés à la proportion de la valeur au marché (*Market Cap*) du titre par rapport à la valeur totale de son industrie de départ. C'est-à-dire la somme des valeurs au marché de toutes les entreprises du marché ayant le même code SIC à deux chiffres. La somme de toutes ces opérations pour chacun des codes SIC à deux chiffres représente le rendement de cette industrie pour la période. Les équations (4.5), (4.6), (4.7) et (4.8) résument ces opérations.

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1} + D_{i,t}}{P_{i,t-1}} \quad (4.5)$$

$$Q_{ind,t-1} = \sum_{i=1}^N Q_{i,t-1} \quad (4.6)$$

$$EQP_{ind,t} = \sum_{i=1}^N \left[(R_{i,t} - R_{f,t}) \times \frac{Q_{i,t-1}}{Q_{ind,t-1}} \right] \quad (4.7)$$

$$E\hat{Q}P_{m,t} = \sum_{ind=1}^{IND} \left[E\hat{Q}P_{ind,t} \times \frac{Q_{ind,t-1}}{Q_{m,t-1}} \right] \quad (4.8)$$

Où :

$R_{i,t}$: Rendement du titre i au temps t .

$R_{f,t}$: Taux sans risque au temps t .

$P_{i,t}$: Prix du titre i au temps t .

$D_{i,t}$: Dividende versé par le titre i au temps t .

$Q_{i,t-1}$: Valeur au marché (*Market Cap*) du titre i au temps $t - 1$.

$Q_{ind,t-1}$: Valeur au marché de l'industrie ind au temps $t - 1$.

$Q_{m,t-1}$: Capitalisation boursière totale du marché m au temps $t - 1$.

N : Nombre de titres maximal composant le regroupement industriel ind pendant la période d'observation.

IND : Nombre total de regroupements industriels utilisées pour effectuer la prévision de la prime de risque de marché. 48 pour le marché américain et 14 pour le marché canadien.

$EQP_{ind,t}$: Prime de risque de l'industrie ind au temps t .

$\hat{EQP}_{m,t}$: Prime de risque espérée du marché m au temps t . Diffère de EQP_t décrit à l'équation (4.3) de la section 4.2 qui fait référence à la prime de risque observée.

Par la suite, une matrice de corrélation est générée pour l'ensemble des industries et pour la totalité de la période d'observation. Les industries étant fortement corrélées entre elles (coefficient de Pearson de 0.7 ou plus) sont ensuite résumées selon un regroupement industriel. Les industries n'étant pas fortement corrélées avec les autres forment un regroupement industriel ayant une seule industrie de départ comme composante. Les tableaux 4 et 5 ci-dessous énumèrent chaque regroupement industriel et les industries de départ le composant. De plus, on y retrouve une courte description de chaque regroupement industriel. Dans le cas des États-Unis, cette opération réduit le nombre de regroupements d'entreprises de 63 industries à 48 regroupements industriels.

Pour ce qui est du Canada, cette opération réduit le nombre de regroupements d'entreprises de 61 industries à 60 industries. Ce faible taux de compression est dû au fait que la majorité des industries du marché canadien sont faiblement corrélées ($\rho < 0.7$) entre elles. Dans la majorité des cas, la corrélation est inexistante ($-0.3 < \rho < 0.3$). Ce phénomène est probablement causé par la distribution des entreprises à travers les diverses industries dans le marché. Pendant la période d'observation, trois industries regroupent 50.63 % des entreprises du marché canadien, soit les secteurs miniers (métaux seulement), de l'extraction pétrolière et de l'immobilier. Le reste des industries ont, en moyenne, 20 entreprises les composant. Plusieurs ont moins de 5 entreprises représentant leurs rendements pour la période d'observation. Ces industries sont assurément très affectées par les activités individuelles de leurs entreprises composantes, ce qui affecte nécessaire-

ment les liens de corrélation entre ces industries. La table 6 ci-dessous illustre la proportion de chaque industrie en termes de nombre d'entreprises par rapport au marché canadien.

En terme de solution, l'opération de regroupement d'industries pour le Canada est modifiée, afin de réduire davantage le nombre de regroupements industriels générés. Ainsi, le nombre d'industries sélectionnées pour former les regroupements industriels à l'usage de cette étude se voit limité selon leur proportion sur le marché canadien. Le processus de regroupement n'inclut désormais que les industries de base ayant une capitalisation boursière agrégée moyenne supérieure à 2 % du total des capitalisations boursières des industries du marché canadien pendant la période d'échantillon. Les industries retenues, mises en évidence dans le tableau 7, forment une approximation du marché canadien. De cette nouvelle composition du marché, l'opération régulière de regroupement d'industries est effectuée, afin de regrouper les industries rencontrant le critère d'application. Le résultat de cette dernière opération est présenté dans le tableau 5 ci-dessous. On observe que le nombre de regroupements industriels est réduit à quatorze, ce qui est une nette amélioration.

Suite à ces altérations, on peut maintenant observer une différence entre la prime de risque observée (EQP) et ces primes de risques estimées à partir d'une reconstruction simplifiée du marché ($E\hat{Q}P$). Pour le marché américain, $EQP_{m,t}$ fait référence au facteur mensuel $Rm - Rf$ du modèle Fama French, tel qu'obtenu sur le site web de Kenneth R. French pour la période t . Pour le marché canadien, chaque observation représente la variation mensuelle de l'indice composite du TSE additionné au rendement mensuel de dividende des titres du TSE. À cette somme est soustrait le taux mensuel des bons du Trésor ayant une échéance d'un mois représentant le taux sans risque. Les données canadiennes proviennent des tables 176-0047 et 176-0043 de la base de données CANSIM. Par ailleurs et tel que décrit à l'équation (4.8) et à la section 4.4.4, $E\hat{Q}P_{m,t}$ est une estimation de la prime de risque obtenue en cumulant les primes de risques de chaque regroupement industriel. Donc, un cumul de prévisions de primes de risques de 48 regroupements industriels pour les États-Unis et quatorze pour le Canada. Le tableau 9 de la section 5.1 présente la corrélation entre $EQP_{m,t}$ et $E\hat{Q}P_{m,t}$ sur la durée de l'échantillon pour chaque marché. Tel qu'anticipé, ces deux ensembles de données sont fortement corrélées pour chaque marché.

TABLEAU 4 – Regroupements industriels du marché des États-Unis

Ce tableau énumère et décrit l'ensemble des regroupements industriels des États-Unis obtenus suite au processus de consolidation d'industries. Les regroupements sont nommés selon un élément commun unissant les industries constituantes. Ces industries sont également énumérées selon leur code SIC à deux chiffres pour chaque regroupement industriel.

Regroupement	Description	Constituants
Metal	Entreprises minières (métaux) et produits de métaux	10, 33
Coal	Entreprises minières (charbon)	12
Oil	Extraction et raffinage de pétrole	13, 29
Quarry	Entreprises minières (non-métaux)	14
Construction	Construction générale	15
HeavyCons	Construction industrielle	16
SpecialCons	Construction spéciale	17
Food	Produits alimentaires	20
Tabacco	Produits du tabac	21
Textile	Production et vente de textile (tissus et vêtements)	22, 23, 56
Paper	Bois, Pâtes et Papiers	24, 25, 26, 27
Chemicals	Production de produits chimiques	28
Rubber	Produits du caoutchouc	30
Leather	Produits du cuir	31
Concrete	Produits d'argile, de verre et de béton	32
MetalProd	Produits dérivés du métal, équipements de transport	34, 37, 50
Computers	Produits électroniques et informatiques	35, 36, 38
MiscManuf	Autres entreprises manufacturières	39
Railroad	Transport ferroviaire	40
Trucks	Transport routier de marchandises	42
Water	Transport par bateau	44
Air	Transport par voie aérienne	45
Pipelines	Pipelines, sauf pour le gaz naturel	46
TransportLogic	Services et logistiques de transport	47
Comm	Entreprises de communication	48
PowerGasWaste	Entreprises de service (Électricité, Gaz et sanitaire)	49
TradeND	Vente de produits non durables	51
Hardware	Vente de matériaux de construction et de jardin	52
GeneralMerch	Magasins de marchandise générale	53
Grocery	Épiceries et distributeurs	54
CarDealers	Concessionnaires automobiles	55
Furniture	Magasins d'ameublement ménager	57
Restaurants	Restaurants	58
MiscRetail	Autres entreprises de vente au détail	59
Finance	Banques, institutions financières et courtiers	60, 61, 62, 63
Insurance	Courtiers et agents d'assurance	64
RealEstate	Immobilier (inclus REITS)	65, 67
Hotels	Entreprises d'hébergement	70
PersonalServ	Entreprises de services personnels	72
Software	Production de logiciels	73
CarRental	Location et réparation d'automobiles	75
Movies	Entreprises de production cinématographique	78
Casinos	Services de récréation (maj. sont des casinos)	79
Healthcare	Services de santé	80
Education	Services d'éducation	82
SocialServices	Services sociaux	83
Engineering	Entreprises de recherche en ingénierie	87
Misc	Autres entreprises (Maj. sont des conglomérats)	99

TABLEAU 5 – Regroupements industriels du marché du Canada

Ce tableau énumère et décrit l'ensemble des regroupements industriels du Canada obtenus suite au processus de consolidation d'industries. Les regroupements sont nommés selon un élément commun unissant les industries constituantes. Ces industries sont également énumérées selon leur code SIC à deux chiffres pour chaque regroupement industriel.

Regroupement	Description	Constituants
MetalMining	Entreprises minières (métaux)	10
Oil	Extraction et raffinage de pétrole	13, 29
Printing	Imprimerie et publication	27
Chemicals	Production de produits chimiques	28
Computers	Produits électroniques et informatiques	36
TransEquip	Production d'équipement de transport	37
Railroad	Transport ferroviaire	40
Comm	Entreprises de communication	48
PowerGasWaste	Entreprises de service (Électricité, Gaz et sanitaire)	49
Grocery	Épiceries et distributeurs	54
Banks	Institutions financières - Banques	60
InsuranceCarry	Assureurs	63
RealEstate	Gestion immobilière	65
Holdings	Fonds de placement immobiliers	67

TABLEAU 6 – Distribution des entreprises du marché canadien

Ce tableau démontre le **nombre** d'entreprises de chaque industrie de base (2 digit SIC) du marché canadien, ainsi que sa **proportion** dans ce marché.

Sic Code	Nombre	Proportion
10	394	16.61%
12	15	0.63%
13	361	15.22%
14	30	1.26%
15	7	0.30%
16	4	0.17%
17	2	0.08%
20	42	1.77%
21	1	0.04%
22	2	0.08%
23	3	0.13%
24	19	0.80%
25	6	0.25%
26	19	0.80%
27	23	0.97%
28	90	3.79%
29	7	0.30%
30	8	0.34%
31	1	0.04%
32	8	0.34%
33	19	0.80%
34	17	0.72%
35	38	1.60%
36	64	2.70%
37	26	1.10%
38	32	1.35%
39	7	0.30%
40	2	0.08%
41	2	0.08%
42	12	0.51%
44	6	0.25%
45	13	0.55%
46	3	0.13%
48	35	1.48%
49	56	2.36%
50	29	1.22%
51	19	0.80%
52	2	0.08%
53	6	0.25%
54	7	0.30%
55	5	0.21%
56	10	0.42%
57	8	0.34%
58	11	0.46%
59	17	0.72%
60	15	0.63%
61	143	6.03%
62	47	1.98%
63	20	0.84%
64	5	0.21%
65	41	1.73%
67	446	18.80%
70	1	0.04%
72	2	0.08%
73	115	4.85%
75	4	0.17%
78	13	0.55%
79	9	0.38%
80	8	0.34%
82	3	0.13%
87	12	0.51%
Total	2372	100.00%

TABLEAU 7 – Distribution des entreprises du marché canadien

Ce tableau démontre la **capitalisation boursière moyenne** de chaque industrie de base (2 digit SIC) du marché canadien pendant la période d'échantillon, ainsi que sa **proportion** dans ce marché.

Sic Code	Cap. Bour. Moy.	Proportion
60	137,374.53	15.34%
13	108,548.90	12.12%
10	85,328.72	9.53%
63	75,886.28	8.47%
48	59,130.22	6.60%
36	59,081.11	6.60%
49	40,178.91	4.49%
29	36,766.03	4.10%
28	31,516.63	3.52%
67	31,008.63	3.46%
27	28,429.98	3.17%
40	22,435.65	2.50%
54	22,365.32	2.50%
65	19,360.90	2.16%
37	19,094.41	2.13%
62	15,018.14	1.68%
73	11,234.97	1.25%
20	11,115.10	1.24%
61	6,308.25	0.70%
24	6,169.16	0.69%
59	5,710.30	0.64%
50	5,023.15	0.56%
26	4,741.91	0.53%
33	4,294.78	0.48%
51	4,271.86	0.48%
87	4,184.19	0.47%
12	3,606.00	0.40%
55	3,230.32	0.36%
35	3,215.58	0.36%
53	3,161.59	0.35%
34	2,599.38	0.29%
58	2,341.92	0.26%
14	2,093.95	0.23%
38	2,022.33	0.23%
46	2,008.31	0.22%
45	1,943.69	0.22%
78	1,631.61	0.18%
42	1,630.26	0.18%
56	1,439.68	0.16%
23	1,326.99	0.15%
57	1,008.00	0.11%
25	1,003.18	0.11%
32	901.27	0.10%
52	863.32	0.10%
79	809.49	0.09%
21	807.76	0.09%
16	601.34	0.07%
44	529.43	0.06%
80	518.96	0.06%
30	492.92	0.06%
15	376.25	0.04%
39	344.28	0.04%
64	191.49	0.02%
75	111.78	0.01%
41	88.52	0.01%
22	65.60	0.01%
70	63.91	0.01%
82	41.36	0.00%
17	29.31	0.00%
72	15.67	0.00%
31	3.58	0.00%
Total	895,697.02	100.00%

4.4.2 Modèles de prévision univariés

Losque les regroupements industriels sont créés, il est ensuite question de déterminer un modèle de prévision univarié pour chacun de ceux-ci. Pour ce faire, chaque variable économique de la section *Données* est testée individuellement pour prédire les rendements de chacun des regroupements industriels. Ce processus suit l'expression générale représentée par l'équation (4.9) ci-dessous. Les rendements des regroupements industriels sont prévus en fonction d'une variable économique observée pendant une période antérieure entre un et six mois. Cet écart temporel est déterminé par la date de publication de la variable économique employée.

$$EQP_{ind,t} = \beta_0 + \beta_1 Z_{t-1} + \epsilon_t \quad (4.9)$$

Où :

$EQP_{ind,t}$: Prime de risque de l'industrie *ind* au temps t .

Z_{t-1} : Variable économique observée au temps $t - 1$.

ϵ_t : Résidus du modèle au temps t .

L'exemple (4.10) ci-dessous propose un modèle de prévision des rendements de l'industrie du pétrole (Oil and gas extraction) en employant la variable économique *ECRUDEG* qui représente la croissance mensuelle du prix des contrats à terme sur la commodité *Light Sweet Crude Oil* du mois précédent.

$$EQP_{oil,t} = \beta_0 + \beta_1 ECRUDEG_{t-1} + \epsilon_t \quad (4.10)$$

Sous l'hypothèse nulle, la variable Z_t n'a pas de pouvoir prévisionnel à l'égard des rendements futurs ($\beta_1 = 0$). L'alternative présume donc un pouvoir prévisionnel chez la variable Z_t ($\beta_1 \neq 0$). Le pouvoir de prédiction de cette variable est déterminé selon sa statistique t dans le modèle, ainsi que par sa mesure R^2 ajusté. Les complications économétriques entourant la taille de l'échantillon et le chevauchement d'observations présentes dans ce type d'étude sont traités via l'utilisation des erreurs types de Newey et West (1987) et la procédure de *Bootstrap* employée notamment par Kothari et Shanken (1997). Les erreurs types de Newey et West (1987) sont robustes envers l'hétéroscédasticité et l'autocorrélation au niveau du terme d'erreur. Le *Bootstrap* est utilisé de façon

à simuler un échantillon de plus grande taille, afin d'éliminer les distorsions de taille provenant du fait que les inférences du modèle sont basées sur la théorie de distribution asymptotique standard. (voir Goetzmann et Jorion (1993)). La méthodologie entourant l'application de cette procédure est décrite extensivement à la section 3.6.

La performance du modèle est évaluée selon sa capacité à minimiser l'erreur quadratique moyenne (Mean Squared Absolute Error) des prévisions générées par ce dernier. De façon similaire à Rapach, Wohar et Rangvid (2010), et Chrétien et Coggins (2011), ce critère de décision est mesuré par le R^2 ajusté et son niveau de signification (déterminé ici par la p-value du Wald test). Le niveau de signification de cette statistique est déterminé suite à une série de simulations de variables générées par la procédure du Bootstrap. Ainsi, le niveau de signification du modèle n'est pas affecté par la taille de l'échantillon, tout en conservant les caractéristiques de l'échantillon observé.

4.4.3 Modèles de prévision multivariés

Suite à l'élaboration de modèles univariés pour prédire les rendements de chaque regroupement industriel, il est question de déterminer la capacité de prévision d'un modèle multivarié dans le même contexte. La méthode employée pour construire ce modèle de prévision est identique à celui employé pour prédire la prime de risque de marché agrégée. C'est-à-dire par l'utilisation de la méthode *Stepwise* de Draper & Smith (1998). Le même processus est répété pour chaque regroupement industriel, afin de construire un modèle multivarié pour chacun d'eux. L'exception créée par une différence entre les modèles générés par les approches *forward selection* et *backward selection* est toujours possible et les prévisions obtenues via les deux méthodes sont retenues et comparées lors de l'analyse comparative. La formule (4.11) ci-dessous représente l'expression générale des modèles multivariés de prévision des rendements de chaque regroupement industriel.

$$EQP_{ind,t} = \beta_0 + \sum_{i=1}^N \beta_{1,i} Z_{i,t-1} + \epsilon_t \quad (4.11)$$

Où :

$EQP_{ind,t}$: Prime de risque de l'industrie ind au temps t .

$Z_{n,t-1}$ Variable économique n observée au temps $t - 1$.

ϵ_t : Résidus du modèle de prévision actuel au temps t .

4.4.4 Agrégation des prévisions sectorielles

Suite à la génération de modèles de prévision univariés et multivariés pour chacun des regroupements industriels vient l'étape de la prévision de la prime de risque de marché. Pour chacun de ces regroupements, le modèle ayant le R^2 ajusté le plus élevé pour la période d'échantillon est sélectionné. Chacun de ces modèles est ensuite utilisé pour prédire les rendements de leur regroupement industriel respectif.

Donc, chacune des variables macroéconomiques aura été testée, tant en contexte univarié et multivarié, afin de déterminer la variable ou la combinaison de variables permettant d'effectuer la meilleure prévision des rendements mensuels de chacun des regroupements industriels. Ainsi, les modèles optimaux ayant été déterminés pour chacun de ceux-ci, le regroupement de leurs prévisions mensuelles devrait générer une prévision optimale de la prime de risque pour cette période. Cette combinaison de prévisions est accomplie en effectuant une somme pondérée des prévisions de chacun des regroupements industriels du modèle en fonction de leur capitalisation boursière en comparaison avec leur marché de référence. Cette opération est représentée par l'équation suivante :

$$E\hat{Q}P_{m,t} = \sum_{ind=1}^{IND} \left[E\hat{Q}P_{ind,t} \times \frac{Q_{ind,t-1}}{Q_{m,t-1}} \right] \quad (4.12)$$

Où :

$E\hat{Q}P_{m,t}$: Prévision de la prime de risque du marché m au temps t .

$E\hat{Q}P_{ind,t}$: Prévision de la prime de risque de l'industrie ind au temps t .

$Q_{ind,t-1}$: Capitalisation boursière de l'industrie ind au temps $t - 1$.

$Q_{m,t-1}$: Capitalisation boursière totale du marché m au temps $t - 1$.

IND : Nombre total de regroupements industriels utilisées pour effectuer la prévision de la prime de risque de marché. 48 pour le marché américain et 14 pour le marché canadien.

4.5 Analyse comparative

Cette section a pour but de déterminer le modèle apportant la meilleure prévision de la prime de risque de marché. Cette comparaison est effectuée pour les États-Unis et pour le Canada. Ainsi, les quatre modèles proposant une prévision de la prime de risque de marché sont comparés de façon similaire aux modèles prévoyant les rendements des différents regroupements industriels dans la section 4.4.4.

1. Modèles de prévision de la prime de risque basés sur la moyenne historique.
2. Modèles univariés de prévision de la prime de risque agrégée.
3. Modèles multivariés de prévision de la prime de risque agrégée (*backward* et *forward*).
4. Agrégation de modèles multivariés de prévision des primes sectorielles pour prévision de la prime de risque agrégée.

Le modèle générant l'erreur quadratique moyenne la plus faible est considéré comme effectuant une prévision supérieure. Donc, le modèle avec le R^2 ajusté le plus élevé est le plus performant à l'intérieur de l'échantillon.

4.6 Bootstrap

La procédure de *Bootstrap* utilisée dans cette étude est similaire à celle employée par Goetzmann et Jorion (1993). Il s'agit également de la méthodologie employée par Rapach, Wohar et Rangvid (2010), et Chrétien et Coggins (2011). Ce processus est utilisé pour contrôler l'autocorrélation entre les erreurs de prévisions. Cette technique permet d'éviter une régression fallacieuse (*Spurious Regression*) engendrée par l'utilisation de données de séries temporelles. Cette procédure est utilisée dans la génération de tous les modèles utilisés dans cette étude.

Dans le cadre de cette étude, la méthode de *Bootstrap* non paramétrique est employée suite à l'exécution de la régression OLS de chaque modèle. Le *Bootstrap* est appliqué à chaque variable du modèle, afin d'augmenter le nombre d'observations pour élever la précision des statistiques t et F (ou Wald χ^2 dans ce cas) représentant le niveau de signification du modèle. Chaque opération de *Bootstrap* est répétée 10 000 fois par le logiciel.

4.7 Prévision Out-of-Sample (OOS)

La méthode de prévision *Out-of-Sample* (OOS) employée dans cette étude est la méthode réursive utilisée par Chrétien et Coggins (2011). Cette technique de prévision estime les modèles de prévision à partir de l'ensemble des données disponibles au moment de leur estimation. Les prévisions selon cette méthode débutent à la 126^e période, soit juin 2001, jusqu'à la fin de l'échantillon, de sorte que la fenêtre d'estimation augmente avec le temps. Pour les fins de cette étude, cette technique sera dorénavant identifiée en tant que *fenêtre de longueur variable*. Ce nombre correspond à la moitié des mois observée dans l'échantillon de 250 mois (238 pour le Canada) et est utilisé comme délais initial pour permettre l'utilisation d'un échantillon de départ suffisamment étendu. Cette méthode est employée de façon à tester chaque modèle dans une situation où l'information utilisée pour effectuer la prévision des rendements financiers futurs se limite à ce qui est disponible au moment présent. Elle reflète la réalité du marché boursier et produit des résultats plus concluants que la méthode de prévision *In-Sample* (IS).

La performance du modèle OOS en comparaison avec sa performance IS est évaluée selon le test statistique MSE-F de McCracken (2007). Ce test est employé pour déterminer si les erreurs quadratiques des prévisions du modèle OOS sont significativement inférieures à celles générées par les prévisions du modèle employant la moyenne historique. Cette comparaison est également effectuée entre le modèle ayant le meilleur R^2 ajusté et ses compétiteurs. Le test et ses composantes sont exprimés par les formules 4.13, 4.14, 4.15 et 4.16 ci-dessous.

$$MSE - F = n \times \frac{MSE_N - MSE_A}{MSE_A} \quad (4.13)$$

$$R^2 = 1 - \frac{MSE_A}{MSE_N} \quad (4.14)$$

$$MSE_A = \frac{1}{n} \sum_{i=1}^n \epsilon_{A_t}^2 \quad (4.15)$$

$$MSE_N = \frac{1}{n} \sum_{i=1}^n \epsilon_{N_t}^2 \quad (4.16)$$

Où :

n : Nombre d'observations.

$\epsilon_{N_t}^2$: Erreur quadratique du modèle de la moyenne historique au temps t .

$\epsilon_{A_t}^2$: Erreur quadratique du modèle testé au temps t .

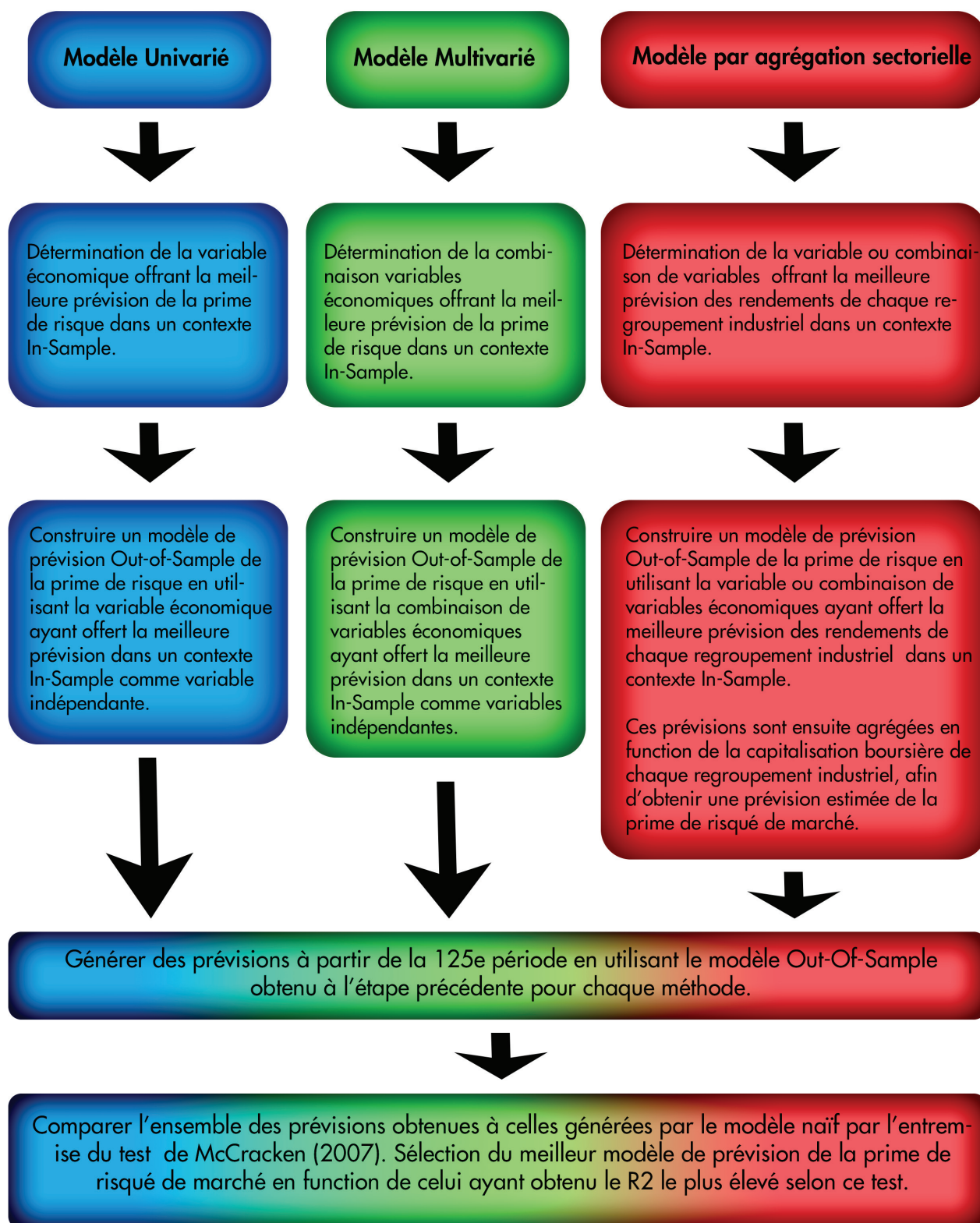
MSE_A : Erreurs quadratiques moyennes du modèle testé.

MSE_N : Erreurs quadratiques moyennes du modèle de la moyenne historique.

L'ensemble de la méthodologie employée pour effectuer ces prévisions *Out-of-Sample* est résumée à la figure 5 de cette section.

FIGURE 5 – Vue d'ensemble de la méthodologie d'estimation des prévisions *Out-of-Sample*

Cette figure est une représentation graphique de chaque étape de la méthodologie d'estimation des prévisions *Out-of-Sample* de cette étude, tel que présenté à la section 4.7.



5 Données

Cette section énumère et décrit les différentes données utilisées dans cette étude. De plus, les statistiques descriptives de chacune de celles-ci y sont également présentées.

5.1 Primes de risque

Les primes de risques étudiées sont décrites ci-dessous et leurs statistiques descriptions sont présentées dans le tableau 8.

Prime de risque américaine ($USDPrem$) : Données historiques provenant de la base de données des facteurs Fama/French disponible sur le site de Kenneth French. Cette prime de risque est calculée en fonction du rendement excédentaire mensuel de tous les titres américains affichés sur le NYSE, AMEX et NASDAQ sur le taux des bons du Trésor 1 mois.

Prime de risque canadienne ($CADPrem$) : Données historiques calculées en fonction du rendement excédentaire mensuel du S&P/TSX Composite Index sur le taux des bons du Trésor 1 mois. Les données canadiennes proviennent des tables 176-0047 et 176-0043 de la base de données CANSIM.

TABLEAU 8 – Statistiques descriptives des primes de risques étudiées

Ce tableau présente les statistiques descriptives pour chacune des primes de risques étudiées. On y retrouve, en ordre, la moyenne, l'écart-type, les valeurs minimales et maximales, et les p -value des tests de normalité *Skewness-Kurtosis*, *Shapiro-Wilk* et *Shapiro-Francia*.

Variable	Moyenne	Écart Type	Min.	Max.	sktest	swilk	sfrancia
EQP_{US}	0.006	0.045	-0.172	0.113	0.0001	0.00021	0.00029
EQP_{CA}	0.005	0.044	-0.204	0.116	0.0000	0.0000	0.0000
N	250						

Le tableau 8 démontre que selon les tests de normalités utilisés (*Skewness-Kurtosis*, *Shapiro-Wilk* et *Shapiro-Francia*) l'hypothèse que les primes de risque américaines et canadiennes sont normalement distribuées doit être rejetée avec un niveau de confiance supérieur ou égal à 95%. Il faut noter à ce stade que ceci pourrait biaiser les résultats obtenus par les tests paramétriques

utilisés dans cette étude. Ce biais pourrait faussement signaler que les modèles de prévisions sont significatifs avec un niveau de confiance supérieur ou égal à 95%.

TABLEAU 9 – Corrélation entre prime de risques observées et estimées

Ce tableau présente la très forte corrélation entre les primes de risques observées (i.e. EQP_{US} , EQP_{CA}) et les primes de risque estimées par la méthode d'agrégation de primes de risques de regroupements industriels décrite à la section 4.4 (i.e. $E\hat{Q}P_{US}$, $E\hat{Q}P_{CA}$).

Corrélation	$E\hat{Q}P_{US}$	$E\hat{Q}P_{CA}$
EQP_{US}	0.877	
EQP_{CA}		0.967

Le tableau 9 présente la corrélation entre $EQP_{m,t}$ et $E\hat{Q}P_{m,t}$ sur la durée de l'échantillon pour chaque marché. Tel que mentionné à la section 4.4.1, ces deux ensembles de données sont fortement corrélées pour chaque marché.

5.2 Regroupements industriels

Les regroupements industriels employés dans cette étude pour les États-Unis et le Canada sont décrits dans les tableaux 4 et 5 de la section 3.4.1. L'ensemble des données utilisées pour construire ces regroupements industriels proviennent de la base de donnée *S&P Research Insights*, maintenant connu sous le nom de *S&P Capital IQ*. De cette base de données ont été extrait le prix historique de chaque titre des marchés américains et canadiens, l'historique des dividendes versés par ces titres, ainsi que le code SIC associé à chacun d'eux.

Cette section rapporte les statistiques descriptives de ces regroupements industriels via les tableaux 10 et 11 ci-dessous. Les statistiques descriptives des industries de départ (via Code SIC) sont disponibles à l'Annexe B du présent document. Ces statistiques incluent la moyenne arithmétique, l'écart type, la valeur minimale et maximale de chaque variable. De plus, les résultats des tests de normalité *Skewness/Kurtosis*, *Shapiro-Wilk* et *Shapiro-Francia* sont effectués via le logiciel STATA. Le test *Skewness/Kurtosis* est une version améliorée du test de *Jarque-Bera* telle que proposée par D'Agostino, Belanger & D'Agostino (1990). Le test de *Shapiro-Wilk* effectue la vérification de normalité de l'échantillon selon la méthode proposée par Shapiro & Wilk (1965). Enfin, le test de *Shapiro-Francia* est une version plus sévère et plus précise du test de *Shapiro-Wilk* proposée par Shapiro & Francia (1972) en vertu de sa cohérence supérieure. Dans le cas des

regroupements américains, la majorité des regroupements industriels obtiennent des résultats de ces tests démontrant que leurs rendements historiques ne sont pas normalement distribués avec un niveau de confiance supérieur ou égal à 95%. L'exception à cette règle sont les variables *Chemicals*, *Railroad*, *Air*, *TransportLogic*, *Hardware*, *GeneralMerch*, *CarDealers*, *Restaurants* et *Healthcare*. Dans le cas des regroupements canadiens, la seule variable ayant des rendements historiques suivant une distribution normale avec un niveau de confiance supérieur ou égal à 95% est *Railroad*. Tout comme pour les primes de risques, il faut noter que ce rejet d'hypothèse de normalité pourrait biaiser les résultats obtenus par les tests paramétriques utilisés dans cette étude.

TABLEAU 10 – Statistiques descriptives des regroupements industriels américains

Ce tableau présente les statistiques descriptives pour chacun des regroupements industriels du marché américain générés par cette étude pour la période d'échantillon. On y retrouve, en ordre, la moyenne, l'écart-type, les valeurs minimales et maximales, et les p-value des tests de normalité Skewness-Kurtosis, Shapiro-Wilk et Shapiro-Francia. La description des regroupements industriels est présentée au tableau 4 du présent document.

Variable	Moyenne	Écart Type	Min.	Max.	sktest	swilk	sfrancia
Metal	0.02	0.078	-0.299	0.323	0.0005	0.00007	0.00006
Coal	0.02	0.104	-0.364	0.378	0.0307	0.029	0.01165
Oil	0.014	0.056	-0.157	0.174	0.0913	0.01547	0.01493
Quarry	0.011	0.065	-0.279	0.181	0	0.00003	0.00004
Construction	0.02	0.094	-0.235	0.339	0.0188	0.03681	0.03095
HeavyCons	0.017	0.082	-0.271	0.282	0.0344	0.03631	0.01676
SpecialCons	0.016	0.102	-0.304	0.391	0.0404	0.0288	0.01427
Food	0.012	0.041	-0.147	0.152	0.0003	0.00002	0.00003
Tabacco	0.014	0.072	-0.246	0.297	0.0009	0.00009	0.00009
Textile	0.02	0.07	-0.228	0.245	0.0389	0.02864	0.01376
Paper	0.011	0.05	-0.183	0.272	0	0	0.00001
Chemicals	0.014	0.04	-0.098	0.154	0.6265	0.49289	0.33002
Rubber	0.018	0.065	-0.193	0.272	0.0027	0.00101	0.00067
Leather	0.026	0.089	-0.26	0.43	0.0013	0.00048	0.00029
Concrete	0.015	0.085	-0.22	0.318	0.0756	0.00642	0.00554
MetalProd	0.015	0.054	-0.177	0.264	0.0002	0.00003	0.00003
Computers	0.02	0.073	-0.242	0.264	0.0106	0.0039	0.00228
MiscManuf	0.017	0.067	-0.187	0.275	0.0331	0.01213	0.00649
Railroad	0.015	0.06	-0.207	0.211	0.2313	0.57507	0.25276
Trucks	0.012	0.055	-0.202	0.238	0.0058	0.01047	0.0029
Water	0.017	0.08	-0.286	0.312	0.0002	0	0.00001
Air	0.015	0.07	-0.208	0.198	0.4274	0.21547	0.17254
Pipelines	0.014	0.045	-0.127	0.167	0.0073	0.00053	0.00055
TransportLogic	0.02	0.066	-0.177	0.259	0.1647	0.43828	0.21808
Comm	0.018	0.076	-0.141	0.655	0	0	0.00001
PowerGasWaste	0.013	0.074	-0.389	0.629	0	0	0.00001
TradeND	0.015	0.05	-0.208	0.16	0.0007	0.00425	0.00205
Hardware	0.016	0.077	-0.183	0.288	0.2259	0.22501	0.15302
GeneralMerch	0.013	0.06	-0.164	0.222	0.1662	0.26155	0.13922
Grocery	0.013	0.054	-0.175	0.147	0.0261	0.04758	0.03964
CarDealers	0.022	0.074	-0.184	0.236	0.6152	0.57774	0.5768
Furniture	0.024	0.096	-0.26	0.439	0.0001	0.00012	0.00012
Restaurants	0.016	0.052	-0.136	0.196	0.7407	0.84397	0.59568
MiscRetail	0.02	0.059	-0.16	0.222	0.0102	0.02262	0.01369
Finance	0.014	0.058	-0.223	0.247	0	0	0.00001
Insurance	0.014	0.061	-0.218	0.416	0	0	0.00001
RealEstate	0.013	0.041	-0.177	0.209	0	0	0.00001
Hotels	0.014	0.075	-0.252	0.418	0	0	0.00001
PersonalServ	0.014	0.067	-0.217	0.3	0.0004	0.00009	0.00008
Software	0.022	0.075	-0.187	0.37	0	0.00004	0.00005
CarRental	0.016	0.112	-0.318	0.563	0	0	0.00001
Movies	0.017	0.073	-0.277	0.35	0.0005	0.00012	0.00007
Casinos	0.027	0.102	-0.3	0.899	0	0	0.00001
Healthcare	0.017	0.066	-0.206	0.239	0.1175	0.23561	0.10583
Education	0.029	0.092	-0.205	0.478	0	0	0.00001
SocialServices	0.024	0.159	-0.466	0.857	0	0	0.00001
Engineering	0.024	0.07	-0.182	0.567	0	0	0.00001
Misc	0.013	0.059	-0.205	0.212	0.0188	0.00431	0.00254
N	250						

TABLEAU 11 – Statistiques descriptives des regroupements industriels canadiens

Ce tableau présente les statistiques descriptives pour chacun des regroupements industriels du marché canadien générés par cette étude pour la période d'échantillon. On y retrouve, en ordre, la moyenne, l'écart-type, les valeurs minimales et maximales, et les p-value des tests de normalité Skewness-Kurtosis, Shapiro-Wilk et Shapiro-Francia. La description des regroupements industriels est présentée au tableau 5 du présent document.

Variable	Moyenne	Écart Type	Min.	Max.	sktest	swilk	sfrancia
MetalMining	0.02	0.09	-0.37	0.452	0.0002	0.00009	0.00004
Oil	0.018	0.064	-0.199	0.28	0.0126	0.01614	0.00539
Printing	0.007	0.053	-0.173	0.326	0	0	0.00001
Chemicals	0.022	0.072	-0.254	0.329	0.0002	0	0.00001
Computers	0.017	0.14	-0.452	0.642	0	0	0.00001
TransEquip	0.014	0.078	-0.314	0.294	0.0005	0.00013	0.0001
Railroad	0.014	0.063	-0.165	0.208	0.1348	0.119	0.09363
Comm	0.014	0.056	-0.269	0.255	0	0	0.00001
PowerGasWaste	0.013	0.045	-0.125	0.46	0	0	0.00001
Grocery	0.013	0.05	-0.192	0.162	0.065	0.07429	0.0273
Banks	0.014	0.052	-0.283	0.156	0	0.00001	0.00001
InsuranceCarry	0.015	0.057	-0.217	0.237	0.0001	0	0.00001
RealEstate	0.015	0.056	-0.244	0.143	0	0.00003	0.00005
Holdings	0.02	0.05	-0.205	0.211	0	0	0.00001
N	238						

5.3 Variables économiques

Cette section énumère et décrit les données qui sont utilisées pour prédire les primes de risque de marché du Canada et des États-Unis, et les rendements de chaque regroupement industriel. En plus de ces variables, les rendements de la période précédente de chaque regroupement industriel sont testés, afin de prévoir les rendements courants de chaque regroupement industriel. Par exemple, il ne serait pas surprenant d'observer que les rendements d'une industrie du secteur primaire prévoient significativement les rendements d'une industrie du secteur secondaire.

5.3.1 Données américaines

Indice des Prix à la Consommation (EPCIG) : Cette variable correspond à la croissance mensuelle de l'Indice des Prix à la Consommation pour tous les consommateurs des États-Unis. L'année de base de cet indice est 1982. Cet item est désaisonnalisé. Source : U. S. Department of Labor, Bureau of Labor Statistics.

Employment Growth (EEG) : Cette variable correspond à la croissance mensuelle du nombre d'emplois comblés aux États-Unis. Le nombre d'emplois inclut les emplois à temps partiel et les emplois surnuméraires. Les emplois issus du secteur agricole sont exclus de cette mesure. Cet item est désaisonnalisé. Source : U. S. Department of Labor, Bureau of Labor Statistics.

Federal Funds Rate (EFFR) : Cette variable correspond au taux directeur mensuel de la U.S. Federal Reserve. Ceci correspond au taux d'intérêt chargé par les banques ayant un surplus de réserves à la Federal Reserve aux banques nécessitant un prêt d'un jour pour répondre aux exigences de réserve. Source : Dow Jones Telerate.

Government Notes (EGN2, EGN3, EGN5, EGN7, EGN10) : Ces variables correspondent aux derniers taux journaliers disponibles des Treasury Notes émises par la U.S. Federal Reserve pour le mois observé. Ces titres sont émis avec une échéance entre 2 et 10 ans avec une valeur nominale minimale de 1000\$. De plus, ces titres sont garantis par le gouvernement des États-Unis. Les revenus obtenus de ces obligations sont exempts d'impôt. Source : S&P Capital IQ

Gross Domestic Product Growth (EGDPG) : Cette variable correspond à la croissance mensuelle du Produit Intérieur Brut (PIB) réel des États-Unis. La publication du PIB national est trimestrielle, donc les observations des trois mois suivants le mois de publication en sont un duplicata. Cet item est désaisonnalisé. Source : U. S. Department of Commerce, Bureau of Economic Analysis.

Housing Starts Growth (EHSG) : Cette variable correspond à la croissance mensuelle des mises en chantier privées aux États-Unis. Cet item est désaisonnalisé. Source : U.S. Department of Commerce, Bureau of the Census.

Inflation rate (EINF) : Cette variable correspond au taux d'inflation annuel du mois observé en comparaison avec la même période de l'année précédente. Elle capture la différence de prix des produits et services sur le marché entre ces deux observations. Source : S&P Capital IQ

Money Supply - M1 Growth (EM1G) : Cette variable correspond à la croissance mensuelle de la masse monétaire M1 des États-Unis, tel que déterminé par la Federal Reserve Board of Governors. La masse monétaire M1 équivaut à la somme de la monnaie en circulation, des dépôts dans les institutions financières et des chèques de voyageurs. Cet item est désaisonnalisé. Source : S&P Capital IQ

Money Supply - M2 Growth (EM2G) : Cette variable correspond à la croissance mensuelle de la masse monétaire M2 des États-Unis, tel que déterminé par la Federal Reserve Board of Governors. La masse monétaire M2 équivaut à la somme de M1, des dépôts dans les comptes d'épargne et dans des comptes d'investissement, des fonds investis dans les fonds mutuels, des comptes sur le marché monétaire, des accords de rachats journaliers, et des dépôts en Eurodollar. Cet item est désaisonnalisé. Source : S&P Capital IQ

Producer Price Index Growth (EPPIG) : Cette variable correspond à la croissance mensuelle de l'Indice des Prix à la Production (IPP) des États-Unis. La période de base de cet indice est 1982, à l'époque où il se nommait le Wholesale Price Index. Cet indice mesure la variation moyenne des prix reçus par les producteurs domestiques pour des commodités à chaque étape du processus de production. Cet item est désaisonnalisé. Source : U. S. Department of Labor, Bureau of Labor Statistics.

Prime Interest Rate (EPRIME) : Cette variable correspond aux taux de base établis par la Federal Reserve Bank of New York. Elle représente la moyenne des taux de base journaliers pour le mois observé. Le taux de base est le taux d'intérêt que les banques chargent à leurs clients ayant le meilleur crédit. Ce taux est affecté par les forces du marché affectant le coût des fonds des banques et les taux que les emprunteurs vont accepter, tel que le taux directeur émis par la U.S. Federal Reserve. Source : S&P Capital IQ

Sale of Passenger Cars (ECARG) : Cette variable correspond à la croissance mensuelle des ventes au détail de nouvelles automobiles pour particuliers aux États-Unis. Cet item n'est pas désaisonnalisé. Source : U.S. Department of Commerce, Bureau of Economic Analysis.

Treasury Bills (ETBILL3, ETBILL6) : Ces items correspondent au taux mensuels des bons du Trésor émis par la U.S. Federal Reserve des États-Unis pour des échéances de 3 et 6 mois. Ces données mensuelles représentent le dernier taux journalier disponible pour le mois observé. Ces titres sont émis avec une valeur nominale minimale de \$10,000. Source : Bank of America.

Unemployment Rate (EUNEMP) : Cette variable correspond aux taux de chômage national

mensuel des États-Unis. Ce taux représente la proportion de citoyens sans emplois parmi la population active du pays. Cet item est désaisonnalisé. Source : U. S. Department of Labor, Bureau of Labor Statistics.

Industrial Production Growth (EIPG) : Cette variable correspond à la croissance mensuelle de la production industrielle des États-Unis. Source : U. S. Department of Commerce, Bureau of Economic Analysis.

Consumer Confidence Index Growth (ECONSG) : Cette variable correspond à la croissance mensuelle de l'indice de confiance des consommateurs aux États-Unis, tel que mesuré par l'Université du Michigan. Source : Bloomberg CONSENT Index.

Jet Fuel Price Growth (EJETG) : Cette variable correspond à la croissance mensuelle du prix du kérosène, tel que rapporté par l'agence d'information énergétique *Platts*. Source : Bloomberg JETINYPR Index.

Personal Income Index Growth (EPIIG) : Cette variable correspond à la croissance mensuelle de l'indice de revenu personnel des États-Unis. Cet indice, mesuré par le U.S. Bureau of Economic Analysis (BEA), indique le revenu moyen des ménages des États-Unis pour le mois courant. Source : Bloomberg PITLCHNG Index.

Durable Goods Orders Growth (EDGNOG) : Cette variable correspond à la croissance mensuelle des commandes de biens durables enregistrées par les entreprises manufacturières des États-Unis. Source : Bloomberg DGNOCHNG Index.

OPEC Monthly Production Growth (EOPECG) : Cette variable correspond à la croissance mensuelle de la production de pétrole mensuelle des pays membres de l'organisation des pays exportateurs de pétrole (OPEP). Source : Bloomberg DOEOMPOT Index.

China A Shares Index Growth (ECHINAG) : Cette variable correspond à la croissance mensuelle de l'indice des actions classées "A" du marché chinois. Ces actions reflètent généralement la situation économique de la Chine qui effectue un grand nombre d'activités commerciales avec les États-Unis. Cet indice inclut des actions du *Shanghai Stock Exchange* et du *Shenzhen Stock Exchange*. Source : Bloomberg SHASHR Index.

Natural Gas Heating Degree Day (EHEATG) : Cette variable correspond à la moyenne mensuelle de la différence entre la température des jours ayant une température moyenne inférieure à 65 degrés Fahrenheit (18°C) aux États-Unis et la température de référence (65°F). Donc, cet

indice capture l'écart mensuel moyen de température à partir de la température extérieure nécessitant l'utilisation de chauffage pour conserver une température intérieure confortable. Source : Bloomberg NOAHAUST Index.

Natural Gas Cooling Degree Day (ECOOLG) : Cette variable est l'opposée de la variable précédente. Elle correspond à la différence entre la température des jours ayant une température moyenne supérieure à 65 degrés Fahrenheit (18°C) aux États-Unis et la température de référence (65°F). Source : Bloomberg NOACAUST Index.

S&P Brent Crude Spot Index Growth (EBRENTG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Brent Crude sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du pétrole brut de Brent Blend, Forties Blend, Oseberg et Ekofisk qui sont tous extraits dans la Mer du Nord de l'océan Atlantique. Source : Bloomberg SPGSBR Index.

S&P Crude Oil Spot Index Growth (ECRUDEG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Light Sweet Crude sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du pétrole brut extrait par les pays membres de l'OPEP. Source : Bloomberg SPGSCL Index.

S&P Unleaded Gas Spot Index Growth (EUGASG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Unleaded Gas sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du pétrole raffiné vendu aux stations d'essence du pays. Source : Bloomberg SPGSHU Index.

S&P Heating Oil Spot Index Growth (EHOILG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Heating Oil sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix de l'huile de chauffage sur le marché. Source : Bloomberg SPGSHO Index.

S&P Natural Gas Spot Index Growth (ENGASG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Natural Gas sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du gaz naturel sur le marché. Source : Bloomberg SPGSNG Index.

S&P Gold Spot Index Growth (EGOLDG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Gold sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix de l'or sur le marché. Source : Bloomberg SPGSGC Index.

S&P Silver Spot Index Growth (ESILVG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Silver sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix de l'argent sur le marché. Source : Bloomberg SPGSSI Index.

S&P Palladium Spot Index Growth (EPALLG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Palladium sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du palladium sur le marché. Source : Bloomberg SPGSPA Index.

S&P Platinum Spot Index Growth (EPLATG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Platinum sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix de la platine sur le marché. Source : Bloomberg SPGSPL Index.

S&P Aluminium Spot Index Growth (EALUMG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Aluminium sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix de l'aluminium sur le marché. Source : Bloomberg SPGSIA Index.

S&P Copper Spot Index Growth (ECOPPG) : Cette variable correspond à la croissance men-

suelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Copper sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du cuivre sur le marché. Source : Bloomberg SPGSIC Index.

S&P Lead Spot Index Growth (ELEADG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Lead sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du plomb sur le marché. Source : Bloomberg SPGSIL Index.

S&P Zinc Spot Index Growth (EZINCG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Zinc sur le *New York Mercantile Exchange* (NYMEX). Cette commodité reflète le prix du zinc sur le marché. Source : Bloomberg SPGSIK Index.

S&P Wheat Spot Index Growth (EWHEATG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Wheat sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix du blé sur le marché. Source : Bloomberg SPGSWH Index.

S&P Corn Spot Index Growth (ECORNG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Corn sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix du maïs sur le marché. Source : Bloomberg SPGSCN Index.

S&P Soybeans Spot Index Growth (ESOYG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Soybeans sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix des graines de soja sur le marché. Source : Bloomberg SPGSSO Index.

S&P Cotton Spot Index Growth (ECOTTG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de

la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Cotton sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix du coton sur le marché. Source : Bloomberg SPGSCT Index.

S&P Sugar Spot Index Growth (ESUGAG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Sugar sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix du sucre sur le marché. Source : Bloomberg SPGSSB Index.

S&P Coffee Spot Index Growth (ECOFFG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Coffee sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix des grains de café sur le marché. Source : Bloomberg SPGSKC Index.

S&P Cocoa Spot Index Growth (ECOCO) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Cocoa sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix du chocolat sur le marché. Source : Bloomberg SPGSCC Index.

S&P Live Cattle Spot Index Growth (ECOWSG) : Cette variable correspond à la croissance mensuelle de l'indice S&P GSCI (Goldman Sachs Commodity Index) reflétant le prix journalier de la dernière journée du mois d'observation du contrat à terme ayant la date d'échéance la plus rapprochée pour la commodité Live Cattle sur le *Chicago Board of Trade* (CBOT). Cette commodité reflète le prix du bétail sur le marché. Source : Bloomberg SPGSLC Index.

U.S. Dollar Index Growth (EUSDOLG) : Cette variable correspond à la croissance mensuelle de l'indice mesurant la valeur du dollar américain en comparaison avec un panier comprenant plusieurs monnaies importantes. Ce panier équivaut à une moyenne pondérée de six monnaies, tel que démontré par le tableau 12 ci-dessous. Cette variable est mesurée quotidiennement. Pour les fins de cette étude, la valeur mensuelle de cette variable correspond à la valeur capturée pendant le dernier jour du mois observé. Source : Bloomberg DXY Curncy.

TABLEAU 12 – Composition du panier de monnaies étrangères

Ce tableau présente la **proportion** de chaque **devise étrangère** utilisée par le *U.S. Dollar Index (DXY)* pour déterminer la valeur du dollar américain sur le plan mondial.

Devise	Proportion
Euro (EUR)	58.6%
Japanese yen (JPY)	12.6%
Pound sterling (GBP)	11.9%
Canadian dollar (CAD)	9.1%
Swedish krona (SEK)	4.2%
Swiss franc (CHF)	3.6%

Selon le tableau 13, la majorité des variables économiques utilisées dans cette étude ne sont pas normalement distribués avec un niveau de confiance supérieur ou égal à 95%. L'exception à cette règle sont les variables *EHSG*, *ECARG*, *EALUMG*, *ECORNG* et *ESUGAG* décrites précédemment.

TABLEAU 13 – Statistiques descriptives des variables économiques américaines testées

Ce tableau présente les statistiques descriptives pour chacune des variables économiques du marché américain utilisées dans cette étude pour la période d'échantillon. On y retrouve la moyenne, l'écart-type, les valeurs minimales et maximales, et les p-value des tests de normalité Skewness-Kurtosis, Shapiro-Wilk et Shapiro-Francia. La source académique employée pour obtenir chaque variable est également présentée.

Variable	Source	Moyenne	Écart Type	Min.	Max.	sktest	swilk	sfrancia	
EPCIG	Research Insight	0.002	0.003	-0.018	0.014	0	0	0.00001	
EEG	Research Insight	0.001	0.002	-0.006	0.004	0	0	0.00001	
EFFR	Research Insight	0.037	0.022	0.001	0.082	0	0	0.00001	
EGN10	Research Insight	0.052	0.014	0.019	0.082	0.0015	0.00647	0.02068	
EGN2	Research Insight	0.04	0.02	0.002	0.077	0	0	0.00001	
EGN3	Research Insight	0.043	0.019	0.003	0.078	0	0	0.00001	
EGN5	Research Insight	0.047	0.017	0.01	0.079	0.0001	0.00028	0.00171	
EGN7	Research Insight	0.05	0.016	0.014	0.081	0.0002	0.00424	0.01542	
EGDPG	Research Insight	0.002	0.005	-0.023	0.02	0	0	0.00001	
EHSB	Research Insight	0.001	0.071	-0.186	0.209	0.3821	0.16525	0.16512	
EINF	Research Insight	0.026	0.012	-0.02	0.063	0	0	0.00001	
EM1G	Research Insight	0.004	0.009	-0.033	0.06	0	0	0.00001	
EM2G	Research Insight	0.004	0.004	-0.007	0.025	0	0	0.00001	
EPIG	Research Insight	0.002	0.006	-0.029	0.02	0	0	0.00001	
EPRIME	Research Insight	0.065	0.02	0.033	0.1	0	0	0.00001	
ECARG	Research Insight	0.008	0.136	-0.448	0.368	0.4587	0.04701	0.04107	
ETBILL3	Research Insight	0.033	0.019	0	0.064	0	0	0.00001	
ETBILL6	Research Insight	0.034	0.019	0.001	0.064	0	0	0.00001	
EUNEMP	Research Insight	0.06	0.016	0.038	0.101	0	0	0.00001	
EIPG	Bloomberg	0.002	0.007	-0.041	0.022	0	0	0.00001	
ECONSG	Bloomberg	0.001	0.052	-0.181	0.246	0.0001	0.00002	0.00002	
EJETG	Bloomberg	0.011	0.11	-0.386	0.498	0	0	0.00001	
EPIIG	Bloomberg	0.004	0.006	-0.037	0.034	0	0	0.00001	
EDGNOG	Bloomberg	0.003	0.037	-0.14	0.167	0.0003	0.00004	0.00003	
EOPECG	Bloomberg	0.001	0.01	-0.062	0.076	0	0	0.00001	
ECHINAG	Bloomberg	0.02	0.187	-0.317	1.895	0	0	0.00001	
EHEATG	Bloomberg	60.736	78.84	0	272	0	0	0.00001	
ECOOLG	Bloomberg	18.088	27.001	0	97	0	0	0.00001	
EBRENTG	Bloomberg	0.012	0.072	-0.326	0.378	0	0	0.00001	
ECRUDEG	Bloomberg	0.009	0.09	-0.326	0.366	0.016	0.01442	0.00453	
EUGASG	Bloomberg	0.01	0.102	-0.391	0.422	0.0046	0.01175	0.00335	
EHOILG	Bloomberg	0.009	0.09	-0.281	0.359	0.0085	0.00134	0.00088	
ENGASG	Bloomberg	0.013	0.155	-0.416	0.626	0	0	0.00001	
EGOLDG	Bloomberg	0.007	0.045	-0.185	0.164	0.0029	0.00045	0.00028	
ESILVG	Bloomberg	0.012	0.083	-0.28	0.283	0.034	0.0061	0.00379	
EPALLG	Bloomberg	0.01	0.097	-0.333	0.469	0	0	0.00001	
EPLATG	Bloomberg	0.007	0.059	-0.31	0.255	0	0	0.00001	
EALUMG	Bloomberg	0.003	0.055	-0.162	0.165	0.2027	0.07185	0.05851	
ECOPPG	Bloomberg	0.007	0.074	-0.356	0.303	0.0001	0.00001	0.00001	
ELEADG	Bloomberg	0.008	0.077	-0.274	0.257	0.0013	0	0.00001	
EZINCG	Bloomberg	0.01	0.098	-0.273	0.351	0.0622	0.04103	0.02528	
EWHEATG	Bloomberg	0.007	0.083	-0.23	0.377	0.0026	0.01645	0.00648	
ECORNG	Bloomberg	0.007	0.078	-0.228	0.231	0.2303	0.22454	0.17712	
ESOYG	Bloomberg	0.006	0.074	-0.299	0.208	0.0004	0.0002	0.00022	
ECOTTG	Bloomberg	0.005	0.089	-0.278	0.313	0.0119	0.00206	0.0017	
ESUGAG	Bloomberg	0.009	0.099	-0.297	0.311	0.2693	0.20727	0.13262	
ECOFFG	Bloomberg	0.01	0.113	-0.376	0.521	0	0	0.00001	
ECOCOG	Bloomberg	0.007	0.087	-0.25	0.346	0.0006	0.00146	0.00114	
ECOWSG	Bloomberg	0.003	0.047	-0.211	0.125	0.0001	0.00031	0.00028	
EUSDOLG	Bloomberg	0	0.025	-0.062	0.094	0.0032	0.00759	0.00672	
N				250					

5.3.2 Données canadiennes

Indice des Prix à la Consommation (EPCIGC) : Cette variable correspond à la croissance mensuelle de l'Indice des Prix à la Consommation pour tous les consommateurs du Canada. L'année de base de cet indice est 1982. Cet item est désaisonnalisé. Source : Statistiques Canada.

Employment Growth (EECG) : Cette variable correspond à la croissance mensuelle du nombre d'emplois comblés au Canada. Le nombre d'emplois inclut les emplois à temps partiel et les emplois surnuméraires. Les emplois issus du secteur agricole sont exclus de cette mesure. Cet item est désaisonnalisé. Source : Statistiques Canada.

Government Bonds (EGOV10C, EGOV30C) : Cette variable correspond aux derniers taux disponibles pour les obligations gouvernementales du Canada pour le mois observé. Les échéances disponibles sont de 5 (EGN5C), 10 (EGOV10C) et 30 (EGOV30C) ans avec une valeur nominale minimale de \$1,000. Source : S&P Capital IQ.

Gross Domestic Product Growth (EGDPGC) : Cette variable correspond à la croissance mensuelle du Produit Intérieur Brut (PIB) du Canada. Cet item est désaisonnalisé. Source : Statistiques Canada.

Housing Starts Growth (EHS GC) : Cette variable correspond à la croissance mensuelle des mises en chantier privées au Canada. Cet item est désaisonnalisé. Source : Statistiques Canada.

Producer Price Index Growth (EPPIGC) : Cette variable correspond à la croissance mensuelle de l'Indice des Prix à la Production (IPP) du Canada. La période de base de cet indice est 2002. Cet indice mesure la variation moyenne des prix reçus par les producteurs domestiques pour des commodités à chaque étape du processus de production. Cet item est désaisonnalisé. Source : Statistiques Canada.

Prime Interest Rate (EPRIMEC) : Cette variable correspond aux taux de base établis par la Banque Centrale du Canada. Elle représente la moyenne des taux de base journaliers pour le mois observé. Le taux de base est le taux d'intérêt que les banques chargent à leurs clients ayant le meilleur crédit. Source : Statistiques Canada.

Raw Material Price Index Growth (ERAWGC) : Cette variable correspond à la croissance mensuelle de l'indice du coût des matières premières au Canada. Cet indice reflète la moyenne mensuelle du prix payé pour ces matières premières par les entreprises du secteur secondaire au

Canada. Source : Statistiques Canada.

Retail Sales Growth (ERETAILGC) : Cette variable correspond à la croissance des ventes mensuelles des entreprises de vente au détail au Canada. Source : Statistiques Canada.

Sale of Passenger Cars (ECARGC) : Cette variable correspond à la croissance mensuelle des ventes au détail de nouvelles automobiles pour particuliers au Canada. Cet item n'est pas désaisonnalisé. Source : Statistiques Canada.

Treasury Bills (ETBILL3C, ETBILL12C) : Ces items correspondent aux taux mensuels des bons du Trésor émis par la Banque Centrale du Canada pour des échéances de 3 et 12 mois. Ces données mensuelles représentent le dernier taux journalier disponible pour le mois observé. Ces titres sont émis avec une valeur nominale minimale de \$1,000.

Unemployment Rate (EUNEMPC) : Cette variable correspond au taux de chômage national mensuel du Canada. Ce taux représente la proportion de citoyens sans emplois parmi la population active du pays. Cet item est désaisonnalisé. Source : Statistiques Canada.

Composite Leading Indicator Growth (ECOMPGC) : Cette variable correspond à la croissance de l'Indicateur Avancé Composite du Canada. Cet indicateur comprend dix composantes qui conduisent de l'activité cyclique et qui représentent les catégories majeures du PIB. Ces composantes sont l'Indice des prix des logements, l'emploi dans les industries de services commerciaux, l'indice des actions à la bourse de Toronto, la durée hebdomadaire moyenne de travail dans la fabrication, la masse monétaire (M1), l'Indicateur Avancé Composite des États-Unis, les nouvelles commandes de biens durables, les expéditions/inventaires de biens finis, et la vente de biens durables (1992=100). Source : Statistiques Canada.

Durable Goods Sales Growth (EDURSGC) : Cette variable correspond à la croissance mensuelle des ventes totales de biens durables au Canada pour le mois observé. Source : Statistiques Canada.

Money Supply Growth (EMONGC) : Cette variable correspond à la croissance mensuelle de la masse monétaire (M2) du Canada, tel que déterminé par la Banque Centrale du Canada. La masse monétaire M2 équivaut à la somme de M1, des dépôts dans les comptes d'épargne et dans des comptes d'investissement, des fonds investis dans les fonds mutuels, des comptes sur le marché monétaire, des accords de rachats journaliers, et des dépôts en Eurodollar. Source : Statistiques Canada.

Toronto Stock Exchange Growth (ETSEIGC) : Cette variable correspond à la croissance men-

suelle de l'indice S&P/TSX Composite Index. Cet indice reflète la valeur au marché des 300 titres ayant le plus d'influence sur le Toronto Stock Exchange (TSE). Ses valeurs et années de base sont 1000 et 1975, respectivement. Source : Statistiques Canada.

United States Composite Leading Index (EUSAGC) : Cette variable correspond à la croissance mensuelle de l'Indicateur Avancé Composite des États-Unis (1992=100). Source : Statistiques Canada.

TABLEAU 14 – Statistiques descriptives des variables économiques canadiennes testées

Ce tableau présente les statistiques descriptives pour chacune des variables économiques du marché canadien utilisées dans cette étude pour la période d'échantillon. On y retrouve la moyenne, l'écart-type, les valeurs minimales et maximales, et les *p-value* des tests de normalité *Skewness-Kurtosis*, *Shapiro-Wilk* et *Shapiro-Francia*. La ressource académique employée pour obtenir chaque variable est également présentée.

Variable	Source	Moyenne	Écart Type	Min.	Max.	sktest	swilk	sfrancia
ECPIGC	Research Insight	0.0016361	0.0030289	-0.0081395	0.0237203	0	0	0.00001
EECG	Research Insight	0.0013105	0.0117823	-0.0241805	0.0348025	0.2708	0.00011	0.00038
EGOV10C	Research Insight	0.059566	0.0187773	0.0274	0.1051	0	0	0.00001
EGOV30C	Research Insight	0.0478909	0.0239059	0	0.094248	0.0053	0	0.00001
EGN10C	Research Insight	0.0563556	0.0189343	0.021995	0.102589	0.0003	0	0.00001
EGN5C	Research Insight	0.0404201	0.0223938	0	0.092381	0.3469	0.00002	0.00093
EGDPGC	Research Insight	0.0020888	0.0036284	-0.0143248	0.0128422	0.0002	0.00044	0.00031
EHSGC	Research Insight	0.0066681	0.0988563	-0.27388	0.266537	0.644	0.54053	0.58848
EPIGIC	Research Insight	0.0015274	0.0074333	-0.0287338	0.0236962	0.0001	0.00011	0.00012
EPRIMEC	Research Insight	0.05678	0.0198256	0.0225	0.1275	0.0025	0.00012	0.00282
ERAWGC	Research Insight	0.0036901	0.0332549	-0.1569208	0.1255337	0	0	0.00001
ERETAILGC	Research Insight	0.012024	0.1178797	-0.3420484	0.236166	0	0	0.00001
ECARGC	Research Insight	0.0177219	0.196442	-0.3315725	0.7497916	0	0	0.00001
ETBILL12C	Research Insight	0.0332752	0.0216904	0	0.087067	0.0018	0.0001	0.03221
ETBILL3C	Research Insight	0.0395238	0.0218183	0.0017	0.1147	0.0081	0.0001	0.00033
EUNEMPC	Research Insight	0.082524	0.017109	0.053	0.125	0.0005	0	0.00001
ECOMPGC	Statistiques Canada	0.0040708	0.0054983	-0.0204871	0.0209436	0.0001	0.00005	0.00005
EDURSGC	Statistiques Canada	0.0036101	0.0137674	-0.0488563	0.0501367	0.0001	0.00013	0.00014
EMONGC	Statistiques Canada	0.0067672	0.0093552	-0.0281242	0.0340397	0.0212	0.00943	0.0062
ETSEIGC	Statistiques Canada	0.0058347	0.0291911	-0.1257577	0.0788093	0	0	0.00001
EUSAGC	Statistiques Canada	0.0014264	0.0032359	-0.0077425	0.0115429	0.0028	0.00005	0.00013
N				250				

Selon le tableau 14, la majorité des variables économiques utilisées dans cette étude ne sont pas normalement distribués avec un niveau de confiance supérieur ou égal à 95%. L'exception à cette règle sont les variables *EECG* et *EHSGC* décrites précédemment.

6 Résultats

6.1 Marché américain

6.1.1 Modèle de prévision de la prime de risque basés sur la moyenne historique

Le modèle de prévision par moyenne historique dans un contexte In-Sample consiste à effectuer une moyenne géométrique de la prime de risque de marché sur l'ensemble des périodes observées. Le résultat de cette opération est une moyenne de **0.5%**.

6.1.2 Modèles univariés de prévision de la prime de risque agrégée

Suite à l'exécution des modèles univariés de l'ensemble des variables économiques sur la prime de risque des États-Unis du mois subséquent, trois modèles obtiennent une performance significative avec un niveau de confiance supérieur ou égal à 95%, soit ceux employant la croissance mensuelle de la masse monétaire M2 des États-Unis, la variation mensuelle du taux d'inflation annuel et l'écart mensuel moyen de température des jours ayant une moyenne supérieure à 65°F aux États-Unis versus la température de référence de 65°F. Le modèle retenu est celui ayant obtenu la meilleure performance, soit celui ayant la croissance mensuelle de la masse monétaire M2 des États-Unis comme variable indépendante avec un R^2 ajusté de .0224. Ce niveau d'explication concorde avec les résultats observés dans la littérature où les modèles univariés ne présentent pas un niveau d'explication élevé.

TABLEAU 15 – Résultats des modèles univariés de prévision de la prime de risque agrégée (US)

Ce tableau présente les résultats de modèles de prévision univariés de la prime de risque américaine agrégée avec une seule variable économique de la période précédente comme variable indépendante. Les statistiques présentées sont le R^2 , le R^2 ajusté, la p -value de la régression, le coefficient B_1 ainsi que le coefficient B_0 obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.2. Les lignes en gras désignent les régressions ayant un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%. Veuillez vous référer à la section 5.3.1 du présent document pour les descriptions des variables économiques étudiées.

Variabes	R2	R2 Adj	Prob > chi2	B1	B0
EM2G	0.0263	0.0224	0.0220	(182.7292)	1.3496
EINF	0.0265	0.0225	0.0380	(61.8004)	2.1839
ECOOLG	0.0222	0.0182	0.0380	(0.0246)	1.0060
EIPG	0.0317	0.0278	0.0726	118.3111	0.3490
ECOCOG	0.0100	0.0060	0.1366	(5.1203)	0.5952
EHSG	0.0090	0.0049	0.1438	5.9847	0.5536
EUNEMP	0.0089	0.0049	0.1749	26.3202	(1.0084)
ESUGAG	0.0065	0.0024	0.1830	(3.6359)	0.5915
ECOTTG	0.0083	0.0043	0.1838	4.5795	0.5366
EGOLDG	0.0058	0.0018	0.3170	(7.6124)	0.6101
ECOWSG	0.0036	(0.0004)	0.3403	5.7153	0.5422
EM1G	0.0088	0.0048	0.3414	(45.2266)	0.7342
EJETG	0.0033	(0.0008)	0.3771	(2.3317)	0.5847
EPPIG	0.0038	(0.0002)	0.3834	44.7170	0.4761
ENGASG	0.0036	(0.0004)	0.3943	1.7367	0.5368
ECORNG	0.0042	0.0002	0.4175	(3.7409)	0.5848
EHOILG	0.0035	(0.0005)	0.4236	(2.9327)	0.5860
EPALLG	0.0043	0.0003	0.4513	3.0386	0.5288
ESOYG	0.0037	(0.0004)	0.4565	(3.6751)	0.5802
EGDPG	0.0034	(0.0007)	0.4577	53.5803	0.4478
ECOPPG	0.0033	(0.0007)	0.4698	3.5004	0.5356
EDGNOG	0.0019	(0.0021)	0.5418	5.2625	0.5451
ECARG	0.0017	(0.0023)	0.5587	1.3622	0.5483
ESILVG	0.0017	(0.0024)	0.5604	(2.2096)	0.5843
EPLATG	0.0036	(0.0004)	0.5636	4.5501	0.5273
EEG	0.0025	(0.0015)	0.5696	125.7310	0.4655
EUSDOLG	0.0017	(0.0023)	0.5725	(7.3208)	0.5599
EALUMG	0.0014	(0.0026)	0.6038	3.0561	0.5502
ECONSG	0.0010	(0.0030)	0.6536	2.7204	0.5567
ECRUDEG	0.0009	(0.0031)	0.6793	(1.5013)	0.5716
EGN10	0.0007	(0.0034)	0.7012	8.1963	0.1317
EWHEATG	0.0009	(0.0032)	0.7023	(1.5828)	0.5700
ECOFFG	0.0006	(0.0034)	0.7033	0.9948	0.5494
ETBILL3	0.0006	(0.0035)	0.7408	5.4403	0.3795
EGN7	0.0005	(0.0035)	0.7419	6.6040	0.2293
ELEADG	0.0005	(0.0036)	0.7673	1.2420	0.5500
EHEATG	0.0003	(0.0038)	0.7821	(0.0010)	0.6180
EUGASG	0.0004	(0.0036)	0.7831	(0.8840)	0.5683
EFFR	0.0004	(0.0037)	0.7866	4.1315	0.4074
ETBILL6	0.0004	(0.0037)	0.7882	4.4020	0.4092
EGN2	0.0002	(0.0038)	0.8249	3.5161	0.4173
EBRENTG	0.0003	(0.0037)	0.8280	(1.0756)	0.5715
EGN3	0.0002	(0.0038)	0.8281	3.6398	0.4037
EGN5	0.0002	(0.0038)	0.8326	3.9039	0.3761
EPIIG	0.0001	(0.0040)	0.8906	6.1037	0.5354
EPCIG	0.0001	(0.0040)	0.9021	14.0520	0.5295
EPRIME	0.0001	(0.0040)	0.9174	1.6447	0.4516
EZINCG	-	(0.0040)	0.9421	(0.2388)	0.5615
EOPECG	-	(0.0040)	0.9486	(1.4606)	0.5609
ECHINAG	-	(0.0040)	0.9618	(0.0750)	0.5608

6.1.3 Modèles multivariés de prévision de la prime de risque agrégée

TABLEAU 16 – Résultats du modèle multivarié *backward* de prévision de la prime de risque agrégée (US)

Ce tableau présente les résultats du modèle de prévision multivarié de la prime de risque américaine agrégée par régression *Stepwise* avec *backward selection*. Cette méthode exclut chaque variable économique du modèle sauf celles étant significatives avec un niveau de confiance supérieur ou égal à 95%. Les éléments présentés sont le R^2 ajusté du modèle, la p -value du modèle, ainsi que le coefficient, l'écart-type, la statistique t et la p -value de la statistique t de chaque variable indépendante obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.3. Veuillez vous référer à la section 5.3.1 du présent document pour les descriptions des variables indépendantes retenues.

Variables	Coef.	Écart Type	t	P> t
EINF	-89.58337	26.00068	-3.45	0.001
EM2G	-217.674	76.03706	-2.86	0.005
EGN5	-1390.522	368.3214	-3.78	0.000
EGN10	903.9845	245.523	3.68	0.000
EGN2	587.7156	154.3418	3.81	0.000
Constante	-1.774908	1.955245	-0.91	0.365

R^2 ajusté du modèle	0.0883
P-value du modèle	0.0000

En employant cette méthode de régression multivariée, le modèle obtient un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%. Les variables retenues par le modèle sont les suivantes :

- **EINF** : Taux d'inflation annuel de la période vs la période précédente.
- **EM2G** : Croissance mensuelle de la masse monétaire M2 des États-Unis.
- **EGN2** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **2** ans.
- **EGN5** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **5** ans.
- **EGN10** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **10** ans.

On y retrouve les variables économiques EINF et EM2G qui ont également été identifiées comme variables significatives dans la prévision de la prime de risque sur un plan univarié. À ces

variables, s'ajoutent les obligations gouvernementales 2 ans, 5 ans et 10 ans qui sont également des variables ayant un impact global sur l'économie américaine.

Cependant, il est possible que ces trois variables soient suffisamment corrélées pour causer un problème de multicollinéarité au modèle. Suite à l'exécution d'un test VIF du modèle, dont les résultats sont présentés dans le tableau 17 ci-dessous, une multicollinéarité entre ces trois variables est effectivement détectée. Ces trois variables suspectes obtiennent un VIF supérieur à 5 et une tolérance de moins de 0.10. Cependant, il faut noter que suite au retrait du modèle de l'une de ces trois variables, les coefficients de régression des variables restantes ne sont pas significativement affectés. De plus, chacune de ces trois variables est significative dans le modèle multivarié. Enfin, le coefficient de détermination de la régression multivariée est affectée négativement de façon significative suite au retrait de l'une de ces trois variables. Il est donc suggéré de conserver toutes les variables de ce modèle afin de maximiser le pouvoir explicatif de celui-ci. Il est cependant important de porter une attention particulière au comportement de ces variables indépendantes dans le cadre d'une application de ce modèle sur une période d'observation différente de celle employée dans le cadre de cette étude.

TABLEAU 17 – Résultats du test VIF sur le modèle multivarié *backward* de prévision de la prime de risque agrégée

Ce tableau énumère chacune des variables du modèle multivarié backward présenté à la section 6.1.3. Le facteur de tolérance représente $1-R_j^2$ où R_j^2 est le coefficient de détermination de la variable j sur toutes les autres variables du modèle. Le facteur VIF représente $\frac{1}{tolerance}$. Un facteur de tolérance inférieur à 0.20 ou 0.10 et/ou un facteur VIF supérieur à 5 ou 10 indiquent un problème potentiel de multicollinéarité. Les variables EGN5, EGN10 et EGN2 entraînent donc un problème de multicollinéarité.

Variable	VIF	Tolérance
EGN5	523.58	0.001910
EGN10	163.58	0.006113
EGN2	125.64	0.007959
EINF	1.28	0.784234
EM2G	1.24	0.807296

TABLEAU 18 – Résultats du modèle multivarié *forward* de prévision de la prime de risque agrégée (US)

Ce tableau présente les résultats du modèle de prévision multivarié de la prime de risque américaine agrégée par régression *Stepwise* avec *forward selection*. Cette méthode inclut chaque variable économique étant significative avec un niveau de confiance supérieur ou égal à 95% sur base univariée dans le modèle multivarié. Les éléments présentés sont le R^2 ajusté du modèle, la *p*-value du modèle, ainsi que le coefficient, l'écart-type, la statistique *t* et la *p*-value de la statistique *t* de chaque variable indépendante obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.3. Veuillez vous référer à la section 5.3.1 du présent document pour les descriptions des variables indépendantes retenues.

R^2 ajusté du modèle		0.0667		
P-value du modèle		0.0002		
Variabes	Coef.	Écart Type	t	P> t
EIPG	100.9294	41.04756	2.46	0.015
EINF	-63.44282	23.59898	-2.69	0.008
ECOOLG	-0.026395	0.102541	-2.57	0.011
Constante	2.526362	0.7353551	3.44	0.001

En employant cette méthode de régression multivariée, le modèle obtient un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%. Les variables retenues par le modèle sont les suivantes :

- **EIPG** : Croissance mensuelle de la production industrielle des États-Unis.
- **EINF** : Taux d'inflation annuel de la période.
- **ECOOLG** : Écart mensuel moyen de température des jours ayant une moyenne supérieure à 65°F aux États-Unis versus la température de référence de 65°F.

On y retrouve les variables économiques EINF et ECOOLG qui ont également été identifiées comme variables significatives dans la prévision de la prime de risque sur un plan univarié. À ces variables, s'ajoute la croissance mensuelle de production industrielle des États-Unis qui est définitivement une variable économique ayant un impact global sur l'économie américaine de par sa nature. Ce modèle obtient un pouvoir explicatif inférieur au modèle multivarié *backward*, mais offre l'avantage de ne pas être affecté par un problème de multicollinéarité.

6.1.4 Agrégation de modèles multivariés de prévision des primes sectorielles

Cette section résume les résultats obtenus dans le cadre de la prévision des rendements mensuels de chaque regroupement industriel par le biais de variables économiques. Le tableau 19 reporte les modèles univariés et multivariés ayant un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95% pour chaque regroupement industriel. La sélection du meilleur modèle est effectuée en fonction de celui obtenant le coefficient de détermination ajusté le plus élevé.

TABLEAU 19 – Résultats des modèles de prévision de la prime de risque de chaque regroupement industriel (US)

Ce tableau présente les résultats de modèles de prévision univariés et multivariés ayant un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95% pour chaque regroupement industriel. Les statistiques présentées sont le R^2 , le R^2 ajusté et la p-value de la régression obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.4. Veuillez vous référer à la section 5.3.1 du présent document pour les descriptions des variables économiques étudiées, ainsi qu'au tableau 4 pour les descriptions des regroupements industriels.

Regroupement	Variables	Chi 2	R2	Adj R2
Metal	EPCIG	0.0092	0.0226	0.0186
	EINF	0.0431	0.0264	0.0224
	Combiné	0.0015	0.0493	0.0416
Coal	EPPIG	0.0474	0.0234	0.0195
	EHEATG	0.0115	0.0198	0.014
	Combiné	0.0077	0.0424	0.0346
Oil	EIPG	0.0056	0.0477	0.0439
	ECONSG	0.0222	0.0267	0.0228
	Combiné	0.0001	0.0816	0.0741
Quarry	ECOCOG	0.0169	0.0188	0.0148
Construction	EM2G	0.0255	0.0196	0.0156
	EJETG	0.0204	0.0227	0.0188
	ECOOLG	0.0442	0.0248	0.0191
	EHOILG	0.0203	0.0287	0.0248
	EGOLDG	0.0475	0.0188	0.0148
	Combiné	0.0026	0.0684	0.0492
HeavyCons	ECOOLG	0.0281	0.0415	0.0359
	ECOCOG	0.0416	0.0173	0.0133
	Combiné	0.0077	0.0411	0.0333
SpecialCons	EHEATG	0.0478	0.0149	0.0091
	ECOOLG	0.014	0.0577	0.0522
	Combiné	0.0192	0.0263	0.0183
Food	EM2G	0.0449	0.0177	0.0137
Tabacco	EGDPG	0.0226	0.0212	0.0172
	ENGASG	0.0169	0.0269	0.0222
	Combiné	0.0035	0.0431	0.0353
Textile	EJETG	0.0433	0.021	0.017
	ECOOLG	0.0052	0.0607	0.0552
	Combiné	0.005	0.0535	0.0458
Paper	EINF	0.0246	0.0385	0.0346
	EM2G	0.0217	0.0259	0.0219
	ECOOLG	0.0334	0.019	0.0132
	Combiné	0.0015	0.0809	0.0697
Chemicals	EM2G	0.0283	0.0195	0.0155
	ETBILL3	0.0432	0.0188	0.0148
	ETBILL6	0.0482	0.0175	0.0135
	EIPG	0.0451	0.0234	0.0195
	Combiné	0.0152	0.056	0.0405

Suite à la page suivante

TABLEAU 19 – SUITE

Regroupement	Variables	Chi 2	R2	Adj R2
Rubber	EHSG	0.0484	0.017	0.013
	EM2G	0.0097	0.0257	0.0217
	ECOOLG	0.0369	0.0164	0.0106
	Combiné	0.0043	0.0507	0.0391
Leather	ECOOLG	0.001	0.062	0.0565
Concrete	ECOOLG	0.0004	0.0686	0.0632
	ESUGAG	0.0117	0.0242	0.0202
	Combiné	0.0001	0.0642	0.0566
MetalProd	ECOOLG	0.0086	0.041	0.0354
Computers	EINF	0.0154	0.0267	0.0227
MiscManuf	ECOOLG	0.0005	0.0448	0.0392
	ESUGAG	0.0366	0.0144	0.0104
	Combiné	0.0012	0.0536	0.0459
Railroad	EHSG	0.0257	0.0212	0.0172
	EPPIG	0.0375	0.0302	0.0263
	EDGNOG	0.0091	0.0329	0.0288
	ECOOLG	0.039	0.0281	0.0224
	Combiné	0.0003	0.0941	0.0792
Trucks	Aucun.			
Water	Aucun.			
Air	Aucun.			
Pipelines	EM1G	0.0355	0.0289	0.025
TransportLogic	EALUMG	0.0136	0.0246	0.0207
Comm	EINF	0.0043	0.0294	0.0255
	EIPG	0.0081	0.0317	0.0278
	ECOCOG	0.043	0.0133	0.0093
	Combiné	0.0007	0.0696	0.0582
PowerGasWaste	EIPG	0.0046	0.0288	0.0249
TradeND	Aucun.			
Hardware	EHSG	0.0353	0.0183	0.0143
	ECOOLG	0.0424	0.0129	0.0071
	ECOWSG	0.0405	0.0128	0.0088
	Combiné	0.0066	0.0474	0.0357
GeneralMerch	EOPECG	0.0389	0.0159	0.0116
Grocery	EJETG	0.0143	0.0204	0.0164
	EHOILG	0.0095	0.0243	0.0203
	EGOLDG	0.0075	0.0254	0.0214
	ESOYG	0.027	0.0241	0.0201
	ECOCOG	0.0175	0.0199	0.0159
	ECOWSG	0.0499	0.0113	0.0073
	Combiné	0.0011	0.0789	0.056
CarDealers	EEG	0.0125	0.0305	0.0266
	EUNEMP	0.0453	0.0142	0.0102
	EJETG	0.0261	0.0207	0.0167
	ECOOLG	0.0144	0.0278	0.0221
	Combiné	0.0013	0.0809	0.0659
Furniture	ECARG	0.0123	0.0273	0.0233
Restaurants	EJETG	0.0059	0.0413	0.0374
	EHOILG	0.0121	0.0312	0.0273
	Combiné	0.0199	0.0414	0.0336
MiscRetail	EHSG	0.0103	0.0245	0.0205
	EINF	0.0122	0.0302	0.0263
	ECRUDEG	0.0368	0.0255	0.0215
	EHOILG	0.0239	0.0255	0.0216
	ECOWSG	0.0468	0.0146	0.0107
	Combiné	0.0001	0.1071	0.0888
Finance	EM2G	0.0014	0.0572	0.0534
Insurance	EM1G	0.0104	0.0348	0.0309
RealEstate	EM2G	0.0131	0.0349	0.031
	ECOTTG	0.047	0.0251	0.0212
	Combiné	0.0192	0.0534	0.0457

Suite à la page suivante

TABLEAU 19 – SUITE

Regroupement	Variables	Chi 2	R2	Adj R2
Hotels	EM2G	0.0215	0.0245	0.0206
	ECOOLG	0.001	0.0523	0.0467
	Combiné	0.0003	0.059	0.0514
PersonalServ	ECOOLG	0.0157	0.0231	0.0174
	ECOTTG	0.0079	0.0298	0.0259
	Combiné	0.0011	0.0517	0.044
Software	EINF	0.0078	0.0272	0.0233
CarRental	ECOOLG	0.0004	0.0568	0.0513
Movies	EM2G	0.03	0.042	0.0381
	EUNEMP	0.0259	0.023	0.019
	ECOPPG	0.0273	0.0229	0.0189
	ELEADG	0.009	0.0299	0.025
	EZINCG	0.0318	0.0235	0.0191
	ECOTTG	0.0255	0.0255	0.0216
	Combiné	0.0028	0.0943	0.0718
Casinos	EUNEMP	0.0498	0.0235	0.0196
Healthcare	EM2G	0.0117	0.0227	0.0187
	ECOOLG	0.0264	0.0303	0.0246
	EBRENTG	0.0049	0.0603	0.054
	Combiné	0.0007	0.0695	0.0581
Education	EFFR	0.0433	0.0246	0.0206
	EPRIME	0.0463	0.0196	0.0157
	ECOOLG	0.0351	0.0149	0.0091
	EALUMG	0.005	0.0299	0.026
	ELEADG	0.0332	0.0265	0.0216
	Combiné	0.004	0.0772	0.0582
SocialServices	ECOTTG	0.0471	0.0158	0.0118
Engineering	EFFR	0.0302	0.0152	0.0112
	EGN2	0.0405	0.0177	0.0138
	EGN3	0.0466	0.0161	0.0121
	EPRIME	0.0193	0.0214	0.0174
	ETBILL3	0.0193	0.0214	0.0174
	ETBILL6	0.0227	0.021	0.017
Misc	EM2G	0.0171	0.0326	0.0287
	EIPG	0.0442	0.0393	0.0354
	EJETG	0.0415	0.0209	0.0169
	EGOLDG	0.0481	0.023	0.019
	Combiné	0.0029	0.0935	0.0786

Le tableau 20 ci-dessous présente la variable ou combinaison de variables fournissant la meilleure prévision des rendements pour chaque regroupement industriel. La sélection du meilleur modèle est effectuée en fonction de celui obtenant le coefficient de détermination ajusté le plus élevé.

TABLEAU 20 – Sélection des modèles de prévisions de la prime de risque de chaque regroupement industriel (US)

Ce tableau présente la variable ou combinaison de variables fournissant la meilleure prévision des rendements pour chaque regroupement industriel ainsi que leur χ^2 et R^2 ajusté. La sélection du meilleur modèle est effectuée en fonction de celui obtenant le coefficient de détermination ajusté le plus élevé. Veuillez vous référer à la section 5.3.1 du présent document pour les descriptions des variables économiques étudiées, ainsi qu'au tableau 4 pour les descriptions des regroupements industriels.

Regroupement	Variables	Chi 2	Adj R2
Metal	EPCIG, EINF	0.0015	0.0416
Coal	EPPIG, EHEATG	0.0077	0.0346
Oil	EIPG, ECONSG	0.0001	0.0741
Quarry	ECOCOG	0.0169	0.0148
Construction	EM2G, EJETG, ECOOLG, EHOILG, EGOLDG	0.0026	0.0492
HeavyCons	ECOOLG	0.0281	0.0359
SpecialCons	ECOOLG	0.014	0.0522
Food	EM2G	0.0449	0.0137
Tabacco	EGDPG, ENGASG	0.0035	0.0353
Textile	ECOOLG	0.0052	0.0552
Paper	EINF, EM2G, ECOOLG	0.0015	0.0697
Chemicals	EM2G, ETBILL3, ETBILL6, EIPG	0.0152	0.0405
Rubber	EHSG, EM2G, ECOOLG	0.0043	0.0391
Leather	ECOOLG	0.001	0.0565
Concrete	ECOOLG	0.0004	0.0632
MetalProd	ECOOLG	0.0086	0.0354
Computers	EINF	0.0154	0.0227
MiscManuf	ECOOLG, ESUGAG	0.0012	0.0459
Railroad	EHSG, EPPIG, EDGNOG, ECOOLG	0.0003	0.0792
Trucks	Modèle naif		
Water	Modèle naif		
Air	Modèle naif		
Pipelines	EM1G	0.0355	0.025
TransportLogic	EALUMG	0.0136	0.0207
Comm	EINF, EIPG, ECOCOG	0.0007	0.0582
PowerGasWaste	EIPG	0.0046	0.0249
TradeND	Modèle naif		
Hardware	EHSG, ECOOLG, ECOWSG	0.0066	0.0357
GeneralMerch	EOPECG	0.0389	0.0116
Grocery	EJETG, EHOILG, EGOLDG, ESOYG, ECOCOG, ECOWSG	0.0011	0.056
CarDealers	EEG, EUNEMP, EJETG, ECOOLG	0.0013	0.0659
Furniture	ECARG	0.0123	0.0233
Restaurants	EJETG	0.0059	0.0374
MiscRetail	EHSG, EINF, ECRUDEG, EHOILG, ECOWSG	0.0001	0.0888
Finance	EM2G	0.0014	0.0534
Insurance	EM1G	0.0104	0.0309
RealEstate	EM2G, ECOTTG	0.0192	0.0457
Hotels	EM2G, ECOOLG	0.0003	0.0514
PersonalServ	ECOOLG, ECOTTG	0.0011	0.044
Software	EINF	0.0078	0.0233
CarRental	ECOOLG	0.0004	0.0513
Movies	EM2G, EUNEMP, ECOPPG, ELEADG, EZINCG, ECOTTG	0.0028	0.0718
Casinos	EUNEMP	0.0498	0.0196
Healthcare	EM2G, ECOOLG, EBRENTG	0.0007	0.0581
Education	EFFR, EPRIME, ECOOLG, EALUMG, ELEADG	0.004	0.0582
SocialServices	ECOTTG	0.0471	0.0118
Engineering	EPRIME	0.0193	0.0174
Misc	EM2G, EIPG, EJETG, EGOLDG	0.0029	0.0786

6.1.5 Analyse comparative

Cette étape consiste à effectuer une analyse comparative des prévisions obtenues avec les modèles les plus performants pour chaque méthodologie. Chacune des ces séries de prévisions est comparée aux prévisions du modèle naïf à l'aide du test MSE-F de McCracken (2007). La performance de chaque modèle est évaluée en fonction du R^2 obtenu suite à ce test. Le tableau 21 ci-dessous résume les résultats de cette analyse comparative.

TABLEAU 21 – Analyse comparative US - In-Sample

Ce tableau présente une analyse comparative de chaque meilleur modèle avec le modèle naïf en tant que référence. Veuillez vous référer aux sections 6.1.1, 6.1.2, 6.1.3 et 6.1.4 du présent document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE, MSE-F et R^2) font référence au test de MSE-F de McCracken (2007) présenté à la section 4.7.

	Naïf	Univarié	Multivarié Backward	Multivarié Forward	Agrégé
MSE	0.1996%	0.1942%	0.1782%	0.1839%	0.1941%
MSE-F		6.85720	29.87882	21.20630	7.00867
R^2		0.02680	0.10714	0.07848	0.02738

Selon les résultats présentés au tableau 21, le modèle d'agrégation de modèles multivariés de prévision des primes sectorielles proposé dans cette étude génère des prévisions légèrement supérieures au modèle univarié de prévision de la prime de risque agrégée dans l'optique de prédire la prime de risque du marché américain. Par contre, le modèle proposé demeure nettement inférieur aux modèles multivariés de prévision de la prime de risque agrégée générés par les deux types de régressions *Stepwise*. Pour le marché américain et dans le cadre de cet échantillon, le modèle multivarié obtenu par un processus de régression *Stepwise Backward* fournit les meilleures prévisions de la prime de risque de marché en comparaison avec le modèle de référence avec un R^2 de 0.10714 selon le test de McCracken (2007). Ce modèle génère donc des prévisions supérieures de 10.7% au modèle de référence. Un rappel des variables indépendantes formant cette régression suit :

- **EINF** : Taux d'inflation annuel de la période vs la période précédente.
- **EM2G** : Croissance mensuelle de la masse monétaire M2 des États-Unis.
- **EGN2** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **2** ans.
- **EGN5** : Dernier taux journalier de la période pour les obligations gouvernementales émises

par la U.S. Federal Reserve ayant une échéance de **5** ans.

- **EGN10** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **10** ans.

TABLEAU 22 – Analyse comparative meilleur modèle US - In-Sample

Ce tableau présente une analyse comparative de chaque modèle avec le modèle multivarié *backwards* de la prime agrégée, identifié à l'étape précédente comme meilleur modèle, en tant que référence. Veuillez vous référer aux sections 6.1.1, 6.1.2, 6.1.3 et 6.1.4 du présent du document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (*MSE*, *MSE-F* et R^2) font référence au test de *MSE-F* de McCracken (2007) présenté à la section 4.7.

	Multivarié Backward	Univarié	Multivarié Forward	Agrégé
MSE	0.1782%	0.1942%	0.1839%	0.1941%
MSE-F		-20.55510	-7.74335	-20.41985
R^2		-0.08998	-0.03210	-0.08933

Les résultats présentés au tableau 22 confirment que le modèle multivarié *backwards* de la prime agrégée génère des prévisions supérieurs aux autres modèles étudiés.

6.1.6 Analyse Comparative - Out-Of-Sample

Cette étape consiste à effectuer une analyse comparative des prévisions *Out-Of-Sample* obtenues avec les modèles les plus performants pour chaque méthodologie. Les prévisions obtenues débutent à la 126^e période (juin 2001) avec une fenêtre de longueur variable de données historiques initiale de 125 périodes. La sélection des variables indépendantes pour l'ensemble de ces modèles *Out-Of-Sample* provient des résultats de l'analyse *In-Sample*. Cette étape consiste en un test complémentaire de la capacité de prévision de ces variables indépendantes sur la prime de risque de marché. Chacune des ces séries de prévisions est comparée aux prévisions du modèle naïf à l'aide du test *MSE-F* de McCracken (2007). La performance de chaque modèle est évaluée en fonction du R^2 obtenu suite à ce test. Le modèle de prévision par moyenne historique dans un contexte *Out-Of-Sample* consiste à effectuer une moyenne géométrique de la prime de risque de marché débutant à la 126^e période avec une fenêtre de longueur variable avec 125 périodes initiales. Le tableau 23 ci-dessous résume les résultats de cette analyse comparative.

TABLEAU 23 – Analyse comparative US - Out-Of-Sample

Ce tableau présente une analyse comparative de chaque modèle avec le modèle naïf en tant que référence dans un contexte *Out-Of-Sample*. Veuillez vous référer aux sections 6.1.1, 6.1.2, 6.1.3 et 6.1.4 du présent document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE , $MSE-F$ et R^2) font référence au test de $MSE-F$ de McCracken (2007) présenté à la section 4.7.

	Naïf	Univarié	Multivarié Backward	Multivarié Forward	Agrégé
MSE	0.2213%	0.2064%	0.1720%	0.1821%	0.2041%
MSE-F		8.93907	35.55609	26.64487	10.45340
R^2		0.06724	0.22284	0.17687	0.07775

Selon les résultats présentés au tableau 23, l'ensemble des modèles de prévisions présente un R^2 supérieur à celui obtenu lors de l'analyse *In-Sample*. Ceci est possible étant donné que le modèle de référence utilisé pour le test de McCracken (2007) dans un contexte *In-Sample* et dans un contexte *Out-Of-Sample* n'est pas identique. Il est possible que le modèle de référence (modèle naïf) soit plus performant *In-Sample* où il est statique dans le temps que dans le contexte *Out-Of-Sample* où il évolue dans le temps.

La conclusion sur le meilleur modèle demeure identique dans un cadre *Out-Of-Sample*. Pour le marché américain et dans le cadre de cet échantillon, le modèle multivarié de prévision de la prime de risque agrégée obtenu par un processus de régression *Stepwise Backward* demeure le modèle offrant les meilleures prévisions de la prime de risque de marché en comparaison avec le modèle de référence avec un R^2 de 0.22284 selon le test de McCracken (2007). Ce modèle obtient donc des prévisions supérieures de 22.28% lorsque comparé au modèle de référence. Les résultats présentés au tableau 24 confirment que le modèle multivarié backwards de la prime agrégée génère des prévisions supérieurs aux autres modèles étudiés dans un contexte *Out-Of-Sample*.

TABLEAU 24 – Analyse comparative meilleur modèle US - Out-Of-Sample

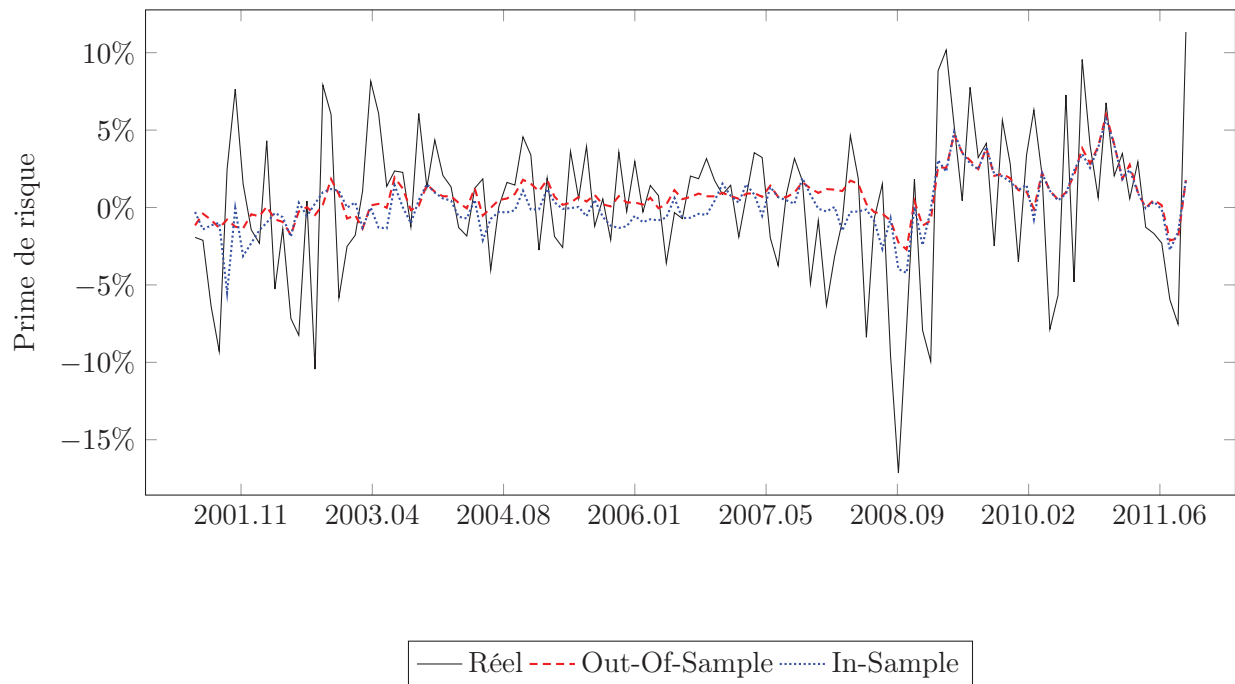
Ce tableau présente une analyse comparative de chaque modèle avec le modèle multivarié backwards de la prime agrégée, identifié à l'étape précédente comme meilleur modèle, en tant que référence dans un contexte *Out-Of-Sample*. Veuillez vous référer aux sections 6.1.1, 6.1.2, 6.1.3 et 6.1.4 du présent document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE, MSE-F et R^2) font référence au test de MSE-F de McCracken (2007) présenté à la section 4.7.

	Multivarié Backward	Univarié	Multivarié Forward	Agrégé
MSE	0.1720%	0.2064%	0.1821%	0.2041%
MSE-F		-20.68559	-6.92542	-19.50871
R^2		-0.20022	-0.05915	-0.18670

Dans un contexte *Out-Of-Sample* le modèle multivarié *Stepwise Backward* semble obtenir une meilleure performance notamment dans la prévision de décroissance significative de la prime de risque américaine. De plus, une meilleure performance du modèle dans les premières années d'observation (2001 à 2005) est notable grâce à l'exclusion des grandes variations affectant les dernières années de la période d'observation (2008 à 2011). Cette comparaison est représentée graphiquement à la figure 6 de la présente section.

FIGURE 6 – Comparaison des résultats du modèle *Stepwise Backward*

Cette figure compare les résultats obtenus par le modèle multivarié *Stepwise Backward* dans un contexte *In-Sample* de la section 6.1.5 et *Out-Of-Sample* de la section 6.1.6 avec la prime de risque réelle observée pour le marché américain.



6.2 Marché canadien

6.2.1 Modèle de prévision de la prime de risque basés sur la moyenne historique

Le modèle de prévision par moyenne historique dans un contexte In-Sample consiste à effectuer une moyenne géométrique de la prime de risque de marché sur l'ensemble des périodes observées. Le résultat de cette opération est une moyenne de **0.4%**.

6.2.2 Modèles univariés de prévision de la prime de risque agrégée

Suite à l'exécution des modèles univariés de l'ensemble des variables économiques sur la prime de risque du Canada du mois subséquent et tel que présenté au tableau 25, un seul modèle obtient une performance ayant un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%, soit celui employant la croissance mensuelle des commandes de biens durables mensuelles enregistrées par les entreprises manufacturières des États-Unis. Ce modèle obtient un R^2 ajusté de .0314 qui est assez faible. Tout comme pour le marché américain, ceci n'est pas surprenant considérant le nombre de facteurs considérables ayant un impact sur l'économie canadienne. Cependant, le lien entre les commandes de biens durables aux États-Unis et la performance du marché canadien est plausible étant donné l'étroite relation entre les deux pays, spécialement au niveau du secteur manufacturier.

Précédemment, Rapach, Wohar et Rangvid (2005) ont démontré que les taux d'intérêt sur les obligations canadiennes prédisent significativement la prime de risque canadienne sur un horizon d'un mois. Or, ce résultat n'est pas concordant avec ceux obtenus ici. Cette divergence pourrait faire l'objet d'une avenue de recherche pour une étude future.

TABLEAU 25 – Résultats des modèles univariés de prévision de la prime de risque agrégée (CA)

Ce tableau présente les résultats de modèles de prévision univariés de la prime de risque canadienne agrégée avec une seule variable économique de la période précédente comme variable indépendante. Les statistiques présentées sont le R^2 , le R^2 ajusté, la p -value de la régression, le coefficient B_1 ainsi que le coefficient B_0 obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.2. Les lignes en gras désignent les régressions ayant un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%. Veuillez vous référer à la section 5.3 du présent document pour les descriptions des variables économiques étudiées.

Variables	R2	R2 Adj	Prob > chi2	B1	B0
EDGNOG	0.0314	0.0273	0.0132	0.2087	0.0045
EIPG	0.0282	0.0241	0.0553	1.1077	0.0031
EINF	0.0229	0.0187	0.0644	(0.6126)	0.0206
EPLATG	0.0182	0.0141	0.0795	0.1018	0.0043
ECOOLG	0.0203	0.0162	0.0850	(0.0002)	0.0095
ECOPPG	0.0105	0.0063	0.0867	0.0612	0.0046
EUSAGC	0.0129	0.0087	0.1097	1.5456	0.0029
ECOFFG	0.0092	0.0050	0.1276	0.0371	0.0047
ERAWGC	0.0099	0.0057	0.2111	0.1322	0.0045
EALUMG	0.0052	0.0009	0.2505	0.0580	0.0049
EGDPG	0.0071	0.0028	0.2724	0.7648	0.0035
ECRUDEG	0.0062	0.0020	0.2776	0.0393	0.0047
EOPECG	0.0033	(0.0010)	0.2875	0.2373	0.0048
EUNEMPC	0.0040	(0.0003)	0.2904	0.1674	(0.0085)
EPALLG	0.0072	0.0030	0.2948	0.0382	0.0047
EDURSGC	0.0052	0.0009	0.3093	0.2391	0.0042
ELEADG	0.0058	0.0015	0.3258	0.0429	0.0048
EZINCG	0.0037	(0.0005)	0.3448	0.0274	0.0048
EECG	0.0036	(0.0006)	0.3596	(0.2304)	0.0054
ECHINAG	0.0035	(0.0008)	0.3940	0.0137	0.0048
EM2G	0.0045	0.0003	0.4026	(0.7473)	0.0084
ECOMPGC	0.0032	(0.0011)	0.4123	(0.4589)	0.0070
EPIIG	0.0036	(0.0007)	0.4136	0.4141	0.0035
EPRIMEC	0.0024	(0.0018)	0.4154	(0.1293)	0.0122
EUNEMP	0.0026	(0.0016)	0.4201	0.1402	(0.0032)
ECARGC	0.0018	(0.0025)	0.4540	0.0096	0.0049
ECARG	0.0019	(0.0024)	0.4620	0.0141	0.0050
EBRENTG	0.0031	(0.0011)	0.4712	0.0338	0.0047
EHEATG	0.0023	(0.0019)	0.4888	0.0000	0.0034
EUGASG	0.0018	(0.0025)	0.5114	0.0181	0.0049
ETBILL12C	0.0014	(0.0028)	0.5203	(0.0811)	0.0080
EHOILG	0.0019	(0.0024)	0.5433	0.0215	0.0049
ESOYG	0.0013	(0.0029)	0.5663	0.0219	0.0050
EM1G	0.0021	(0.0022)	0.5828	(0.2141)	0.0059

6.2.3 Modèles multivariés de prévision de la prime de risque agrégée

TABLEAU 26 – Résultats du modèle multivarié *backward* de prévision de la prime de risque agrégée (CA)

Ce tableau présente les résultats du modèle de prévision multivarié de la prime de risque canadienne agrégée par régression **Stepwise** avec **backward selection**. Cette méthode exclut chaque variable économique du modèle sauf celles étant significative avec un niveau de confiance supérieur ou égal à 95%. Les éléments présentés sont le R^2 ajusté du modèle, la p -value du modèle, ainsi que le coefficient, l'écart-type, la statistique t et la p -value de la statistique t de chaque variable indépendante obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.3. Veuillez vous référer à la section 5.3 du présent document pour les descriptions des variables indépendantes retenues.

R^2 ajusté du modèle		0.0433		
P-value du modèle		0.0021		
Variabes	Coef.	Écart Type	t	P> t
ECOOLG	-0.000231	0.0001039	-2.22	0.027
EDGNOG	0.2077488	0.750106	2.77	0.006
Constante	0.0089223	0.003471	2.57	0.011

En employant cette méthode de régression multivariée, le modèle obtient un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%. Les variables retenues par le modèle sont les suivantes :

- **ECOOLG** : Écart mensuel moyen de température des jours ayant une moyenne supérieure à 65°F aux États-Unis versus la température de référence de 65°F.
- **EDGNOG** : Croissance mensuelle des commandes de biens durables mensuelles enregistrées par les entreprises manufacturières des États-Unis.

On y retrouve la variable EDGNOG qui avait également été identifiée comme variable significative dans la prévision de la prime de risque sur un plan univarié. À cette variable s'ajoute la variable ECOOLG représentant l'écart mensuel moyen de température des jours ayant une moyenne supérieure à 65°F aux États-Unis versus la température de référence de 65°F. Ce modèle offre un pouvoir explicatif supérieur au modèle univarié avec un R^2 ajusté de .0433. Ces deux variables sont en lien avec le marché américain, mais encore une fois, l'étroite relation économique entre les États-Unis et le Canada peut expliquer le lien entre celles-ci et la prime de risque canadienne. Il faut cependant noter que ce modèle souffre possiblement d'une régression fallacieuse en raison de

la nature temporelle des données. Cet effet est souligné par la diminution marquée du niveau de signification de ce modèle dans un contexte *Out-Of-Sample*, tel que présenté à la section 6.2.6. Cet effet est atténué pour les modèles multivarié *Forward* et par agrégation de modèles multivariés de prévisions des primes sectorielles en raison du nombre plus élevés de variables composant ces modèles.

TABLEAU 27 – Résultats du modèle multivarié *forward* de prévision de la prime de risque agrégée (CA)

Ce tableau présente les résultats du modèle de prévision multivarié de la prime de risque canadienne agrégée par régression **Stepwise** avec **forward selection**. Cette méthode inclut chaque variable économique étant significative avec un niveau de confiance supérieur ou égal à 95% sur base univariée dans le modèle multivarié. Les éléments présentés sont le R^2 ajusté du modèle, la p-value du modèle, ainsi que le coefficient, l'écart-type, la statistique t et la p-value de la statistique t de chaque variable indépendante obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.3. Veuillez vous référer à la section 5.3 du présent document pour les descriptions des variables indépendantes retenues.

R^2 ajusté du modèle		0.0652		
P-value du modèle		0.0003		
Variables	Coef.	Écart Type	t	P> t
EDGNOG	0.2055442	0.741546	2.77	0.006
EINF	-0.6498284	0.2555485	-2.54	0.012
ECOOLG	-0.0002501	0.000103	-2.43	0.016
Constante	0.257577	0.0074569	3.45	0.001

En employant cette méthode de régression multivariée, le modèle obtient un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95%. Les variables retenues par le modèle sont les suivantes :

- **EDGNOG** : Croissance mensuelle des commandes de biens durables mensuelles enregistrées par les entreprises manufacturières des États-Unis.
- **EINF** : Taux d'inflation annuel des États-Unis de la période vs la période précédente.
- **ECOOLG** : Écart mensuel moyen de température des jours ayant une moyenne supérieure à 65°F aux États-Unis versus la température de référence de 65°F.

On y retrouve les variables EDGNOG et ECOOLG qui avaient été sélectionnées dans le modèle multivarié *backward*. À celles-ci s'ajoute la variable EINF représentant la variation mensuelle du taux d'inflation annuel aux États-Unis. Ce modèle obtient un R^2 ajusté de 0.0652, soit une

amélioration en comparaison avec le modèle multivarié *backward* et le modèle univarié.

Encore une fois, toutes les variables indépendantes du modèle font référence à des facteurs du marché américain. Dans l'ensemble, les variables économiques provenant du marché américain obtiennent une meilleure performance de prévision de la prime de risque canadienne en comparaison avec les variables économiques canadiennes. Ce phénomène serait possiblement dû au fait que les événements économiques américains ont tendance à affecter le marché canadien à retardement. Étant donné que cette étude tente de prédire la prime de risque du mois suivant l'observation des variables économiques, il est possible que cet effet de retardement ait un impact positif sur le pouvoir explicatif des variables économiques américaines.

6.2.4 Agrégation de modèles multivariés de prévision des primes sectorielles

Cette section résume les résultats obtenus dans le cadre de la prévision des rendements mensuels de chaque regroupement industriel par le biais de variables économiques. Le tableau 28 reporte les modèles univariés et multivariés ayant un pouvoir explicatif avec un niveau de confiance supérieur ou égal à 95% pour chaque regroupement industriel.

TABLEAU 28 – Résultats des modèles de prévision de la prime de risque de chaque regroupement industriel (CA)

Ce tableau présente les résultats de modèles de prévision univariés et multivariés ayant un pouvoir explicatif avec un niveau de confiance supérieur à 95% pour chaque regroupement industriel. Les statistiques présentées sont le R^2 , le R^2 ajusté et la p-value de la régression obtenus par le logiciel STATA en suivant la méthodologie présentée à la section 4.4. Veuillez vous référer à la section 5.3 du présent document pour les descriptions des variables économiques étudiées, ainsi qu'au tableau 5 pour les descriptions des regroupements industriels.

Regroupement	Variables	Chi 2	R2	Adj R2
MetalMining	ECPIGC	0.0120	0.0238	0.0196
	EMONGC	0.0467	0.0133	0.0091
	EPCIG	0.0000	0.0667	0.0627
	EPPIG	0.0347	0.0255	0.0214
	EGOLDG	0.0489	0.0254	0.0212
	Combined	0.0006	0.0948	0.0752
Oil	ERAWGC	0.0142	0.0265	0.0223
	ECARGC	0.0206	0.0176	0.0134
	EIPG	0.0098	0.0425	0.0384
	Combined	0.0019	0.0716	0.0596
Printing	EECG	0.0137	0.0271	0.0229
	EGOV30C	0.0224	0.0181	0.0135
	EM1G	0.0203	0.0266	0.0225
	EIPG	0.0206	0.0249	0.0208
	EDGNOG	0.0127	0.0297	0.0255
	ECOOLG	0.0171	0.0357	0.0300
	Combined	0.0011	0.1157	0.0926
Chemicals	EHEATG	0.0477	0.0217	0.0159
	EPLATG	0.0364	0.0381	0.0341

Suite à la page suivante

TABLEAU 28 – SUITE

Regroupement	Variables	Chi 2	R2	Adj R2
	Combined	0.0320	0.0442	0.0360
Computers	EPLATG	0.0322	0.0177	0.0135
TransEquip	ERETAILGC	0.0418	0.0111	0.0069
	ECARGC	0.0086	0.0285	0.0244
	EUNEMPC	0.0447	0.0124	0.0082
	ECOMPGC	0.0261	0.0289	0.0247
	ETSEIGC	0.0001	0.0782	0.0743
	EUSAGC	0.0403	0.0189	0.0148
	EM2G	0.0111	0.0325	0.0284
	ECARG	0.0311	0.0239	0.0198
	ECOOLG	0.0049	0.0472	0.0416
	ECOTTG	0.0207	0.0291	0.0249
	ECOWSG	0.0345	0.0156	0.0114
	Combined	0.0009	0.1579	0.1168
Railroad	EHSG	0.0148	0.0186	0.0144
	EM1G	0.0079	0.0229	0.0187
	ECOOLG	0.0369	0.0228	0.0170
	ESOYG	0.0088	0.0282	0.0240
	Combined	0.0000	0.0867	0.0710
Comm	EGDPGC	0.0211	0.0339	0.0297
	ECOMPGC	0.0136	0.0199	0.0157
	EEG	0.0095	0.0225	0.0183
	Combined	0.0141	0.0443	0.0320
PowerGasWaste	EIPG	0.0412	0.0207	0.0165
	EWHEATG	0.0236	0.0152	0.0110
	EUSDOLG	0.0173	0.0170	0.0128
	Combined	0.0179	0.0458	0.0335
Grocery	EGOV30C	0.0274	0.0149	0.0103
	EPPIG	0.0102	0.0196	0.0154
	ECRUDEG	0.0211	0.0285	0.0243
	EUGASG	0.0448	0.0180	0.0138
	EHOILG	0.0075	0.0359	0.0318
	EGOLDG	0.0145	0.0269	0.0228
	ESILVG	0.0034	0.0277	0.0236
	EPALLG	0.0177	0.0220	0.0171
	EPLATG	0.0019	0.0335	0.0294
	ECOPPG	0.0018	0.0313	0.0271
	Combined	0.0208	0.0834	0.0429
Banks	EECG	0.0452	0.0168	0.0126
	ERAWGC	0.0442	0.0225	0.0183
	ETSEIGC	0.0291	0.0310	0.0269
	EM1G	0.0074	0.0313	0.0272
	EM2G	0.0458	0.0165	0.0123
	EBRENTG	0.0365	0.0382	0.0318
	ECRUDEG	0.0042	0.0377	0.0336
	EUGASG	0.0283	0.0240	0.0199
	EHOILG	0.0165	0.0310	0.0269
	ECOPPG	0.0459	0.0212	0.0170
	EZINCG	0.0030	0.0382	0.0338
	Combined	0.0016	0.1151	0.0719
InsuranceCarry	EGOV10C	0.0163	0.0199	0.0157
	EGOV30C	0.0346	0.0178	0.0132
	EGN10C	0.0357	0.0180	0.0138
	ERAWGC	0.0459	0.0268	0.0226
	EUNEMPC	0.0187	0.0173	0.0131
	EGN10	0.0457	0.0212	0.0170
	EGN2	0.0354	0.0243	0.0202
	EGN3	0.0351	0.0247	0.0205
	EGN5	0.0411	0.0228	0.0186
	EGN7	0.0388	0.0235	0.0194
	EM2G	0.0178	0.0203	0.0161
	ETBILL3	0.0428	0.0237	0.0196

Suite à la page suivante

TABLEAU 28 – SUITE

Regroupement	Variables	Chi 2	R2	Adj R2
	ETBILL6	0.0404	0.0230	0.0189
	EGOLDG	0.0439	0.0176	0.0134
	Combined	0.0188	0.1203	0.0648
RealEstate	ERAWGC	0.0339	0.0248	0.0206
	ECOMPGC	0.0282	0.0261	0.0219
	EMONGC	0.0034	0.0273	0.0231
	ETSEIGC	0.0046	0.0514	0.0474
	EUSAGC	0.0026	0.0395	0.0354
	ECRUDEG	0.0416	0.0196	0.0154
	Combined	0.0005	0.1101	0.0869
Holdings	EGOV10C	0.0058	0.0325	0.0284
	EGN10C	0.0164	0.0241	0.0199
	EUNEMPC	0.0059	0.0396	0.0355
	EMONGC	0.0086	0.0205	0.0164
	EPCIG	0.0450	0.0095	0.0053
	EM2G	0.0195	0.0221	0.0180
	Combined	0.0000	0.0960	0.0724

Le tableau 29 ci-dessous présente la variable ou combinaison de variables fournissant la meilleure prévision des rendements pour chaque regroupement industriel. La sélection du meilleur modèle est effectuée en fonction de celui obtenant le coefficient de détermination ajusté le plus élevé.

TABLEAU 29 – Résultats des prévisions de rendements sectoriels canadiens

Ce tableau présente la variable ou combinaison de variables fournissant la meilleure prévision des rendements pour chaque regroupement industriel ainsi que leur χ^2 et R^2 ajusté. La sélection du meilleur modèle est effectuée en fonction de celui obtenant le coefficient de détermination ajusté le plus élevé. Veuillez vous référer à la section 5.3 du présent document pour les descriptions des variables économiques étudiées, ainsi qu'au tableau 5 pour les descriptions des regroupements industriels.

Regroupement	Variables	Chi 2	Adj R2
MetalMining	ECPIGC, EMONGC, EPCIG, EPPIG, EGOLDG	0.0006	0.0752
Oil	ERAWGC, ECARGC, EIPG	0.0019	0.0596
Printing	EECG, EGOV30C, EM1G, EIPG, EDGNOG, ECOOLG	0.0011	0.0926
Chemicals	EHEATG, EPLATG	0.0320	0.0360
Computers	EPLATG	0.0322	0.0135
TransEquip	ERETAILGC, ECARGC, EUNEMPC, ECOMPGC, ETSEIGC, EUSAGC, EM2G, ECARG, ECOOLG, ECOTTG, ECOWSG	0.0009	0.1168
Railroad	EHSG, EM1G, ECOOLG, ESOYG	0.0000	0.0710
Comm	EGDPGC, ECOMPGC, EEG	0.0141	0.0320
PowerGasWaste	EIPG, EWHEATG, EUSDOLG	0.0179	0.0335
Grocery	EGOV30C, EPPIG, ECRUDEG, EUGASG, EHOILG, EGOLDG, ESILVG, EPALLG, EPLATG, ECOPPG	0.0208	0.0429
Banks	EECG, ERAWGC, ETSEIGC, EM1G, EM2G, EBRENTG, ECRUDEG, EUGASG, EHOILG, ECOPPG, EZINCG	0.0016	0.0719
InsuranceCarry	EGOV10C, EGOV30C, EGN10C, ERAWGC, EUNEMPC, EGN10, EGN2, EGN3, EGN5, EGN7, EM2G, ETBILL3, ETBILL6, EGOLDG	0.0188	0.0648
RealEstate	ERAWGC, ECOMPGC, EMONGC, ETSEIGC, EUSAGC, ECRUDEG	0.0005	0.0869
Holdings	EGOV10C, EGN10C, EUNEMPC, EMONGC, EPCIG, EM2G	0.0000	0.0724

6.2.5 Analyse comparative

Cette étape consiste à effectuer une analyse comparative des prévisions obtenues avec les modèles les plus performants pour chaque méthodologie. Chacune des ces séries de prévisions sont comparées aux prévisions du modèle naïf à l'aide du test MSE-F de McCracken (2007). La performance de chaque modèle est évaluée en fonction du R^2 obtenu suite à ce test. Le tableau 30 ci-dessous résume les résultats de cette analyse comparative.

TABLEAU 30 – Analyse comparative CA - In-Sample

Ce tableau présente une analyse comparative de chaque modèle avec le modèle naïf en tant que référence. Veuillez vous référer aux sections 6.2.1, 6.2.2, 6.2.3 et 6.2.4 du présent du document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE, MSE-F et R^2) font référence au test de MSE-F de McCracken (2007) présenté à la section 4.7.

	Naïf	Univarié	Multivarié Backward	Multivarié Forward	Agrégé
MSE	0.2002%	0.1938%	0.1898%	0.1846%	0.1881%
MSE-F		7.76782	12.90860	19.85409	15.09176
R2		0.03187	0.05186	0.07760	0.06010

Selon les résultats présentés aux tableau 30 et 31, le modèle par agrégation sectorielle proposé dans cette étude génère des prévisions significativement supérieures au modèle univarié de prévision de la prime de risque canadienne agrégée. Cependant, il demeure que pour le marché canadien et dans le cadre de cet échantillon, le modèle multivarié de prévision de la prime de risque agrégée obtenu par un processus de régression *Stepwise Forward* fournit les meilleures prévisions de la prime de risque de marché en comparaison avec le modèle de référence avec un R^2 de 0.07760 selon le test de McCracken (2007). Ce modèle obtient donc des prévisions supérieures de 7.8% lorsque comparé au modèle de référence. Les facteurs composant ce modèle sont décrits dans la section précédente.

TABLEAU 31 – Analyse comparative meilleur modèle CA - In-Sample

Ce tableau présente une analyse comparative de chaque modèle avec le modèle multivarié forwards de la prime agrégée, identifié à l'étape précédente comme meilleur modèle, en tant que référence. Veuillez vous référer aux sections 6.2.1, 6.2.2, 6.2.3 et 6.2.4 du présent du document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE, MSE-F et R^2) font référence au test de MSE-F de McCracken (2007) présenté à la section 4.7.

	Multivarié Forward	Univarié	Multivarié Backward	Agrégé
MSE	0.1846%	0.1938%	0.1898%	0.1881%
MSE-F		-11.14838	-6.40653	-4.39278
R2		-0.04958	-0.02790	-0.01897

Les résultats présentés au tableau 31 confirment que le modèle multivarié *forwards* de la prime agrégée génère des prévisions supérieures aux autres modèles étudiés.

6.2.6 Analyse Comparative - Out-Of-Sample

Cette étape consiste à effectuer une analyse comparative des prévisions *Out-Of-Sample* obtenues avec les modèles les plus performants pour chaque méthodologie. Les prévisions obtenues débutent à la 119^e (décembre 2001) période avec une fenêtre récursive de données historiques initiale de 118 périodes. La sélection des variables indépendantes pour l'ensemble de ces modèles *Out-Of-Sample* provient des résultats de l'analyse *In-Sample*. Cette étape consiste en un test complémentaire de la capacité de prévision de ces variables indépendantes sur la prime de risque de marché. Chacune de ces séries de prévisions sont comparées aux prévisions du modèle naïf à l'aide du test MSE-F de McCracken (2007). La performance de chaque modèle est évaluée en fonction du R^2 obtenu suite à ce test. Le modèle de prévision par moyenne historique dans un contexte *Out-Of-Sample* consiste à effectuer une moyenne géométrique de la prime de risque de marché débutant à la 119^e période avec une fenêtre récursive avec 125 périodes initiales. Le tableau 32 ci-dessous résume les résultats de cette analyse comparative.

TABLEAU 32 – Analyse comparative CA - Out-Of-Sample

Ce tableau présente une analyse comparative de chaque modèle avec le modèle naïf en tant que référence dans un contexte *Out-Of-Sample*. Veuillez vous référer aux sections 6.2.1, 6.2.2, 6.2.3 et 6.2.4 du présent du document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE, MSE-F et R^2) font référence au test de MSE-F de McCracken (2007) présenté à la section 4.7.

	Naïf	Univarié	Multivarié Backward	Multivarié Forward	Agrégé
MSE	0.1757%	0.1755%	0.1736%	0.1609%	0.1445%
MSE-F		0.15778	1.43627	10.91204	25.72119
R2		0.00132	0.01193	0.08400	0.17773

Selon les résultats présentés au tableau 32, le pouvoir explicatif du modèle univarié et du modèle multivarié *Stepwise Backward* est nettement diminué dans un contexte *Out-Of-Sample*. Cependant, dans le cas du modèle multivarié *Stepwise Forward* de la prime agrégée, le pouvoir explicatif connaît une légère amélioration. Or, dans le cas du modèle de prévision par agrégation sectorielle, le pouvoir explicatif passe près de tripler dans un contexte *Out-Of-Sample*. L'explication pour cette amélioration est la même que dans le contexte du marché américain. Il s'agit donc du modèle offrant les meilleures prévisions de la prime de risque de marché canadienne en comparaison avec le modèle de référence avec un R^2 de 0.17773 selon le test de McCracken (2007). Ce modèle

obtient donc des prévisions supérieures de 17.77% lorsque comparé au modèle de référence. Les résultats présentés au tableau 33 confirment que le modèle d'agrégation de modèles multivariés de prévision des primes sectorielles pour prévision de la prime de risque agrégée génère des prévisions supérieurs aux autres modèles étudiés dans un contexte *Out-Of-Sample*.

TABLEAU 33 – Analyse comparative meilleur modèle CA - Out-Of-Sample

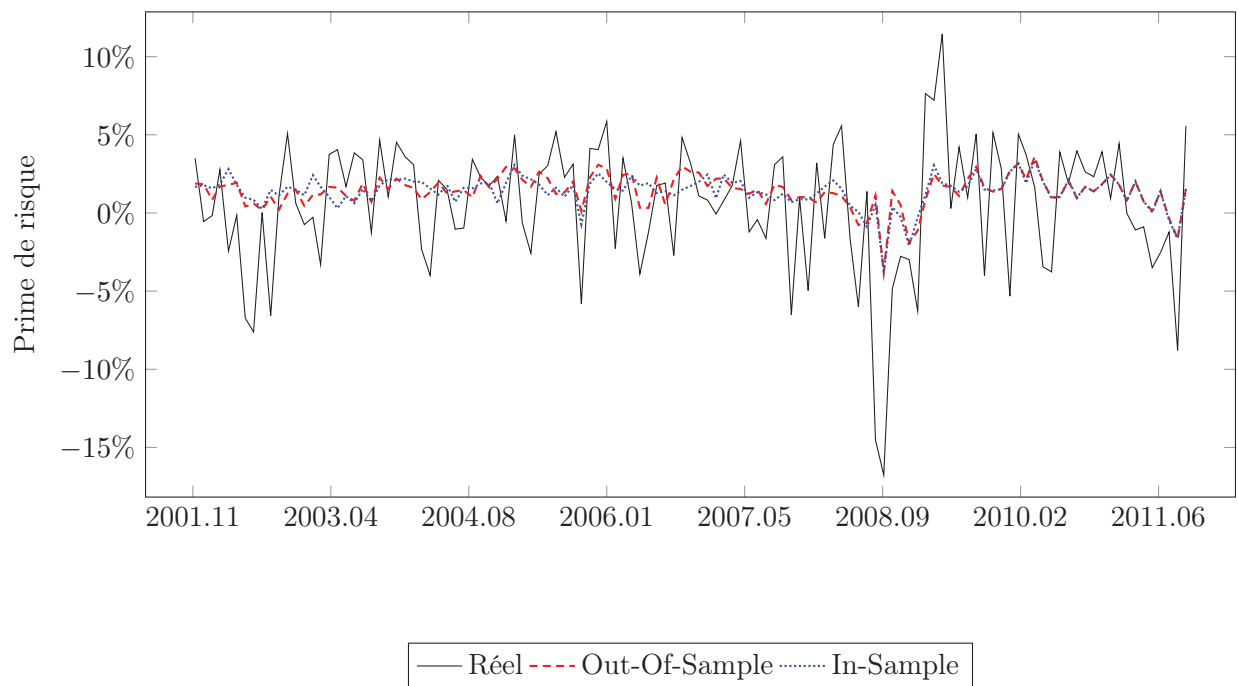
Ce tableau présente une analyse comparative de chaque modèle avec le modèle d'agrégation de modèles multivariés de prévision des primes sectorielles pour prévision de la prime de risque agrégée, identifié à l'étape précédente comme meilleur modèle, en tant que référence dans un contexte *Out-Of-Sample*. Veuillez vous référer aux sections 6.2.1, 6.2.2, 6.2.3 et 6.2.4 du présent du document pour davantage de détails sur les résultats obtenus par chacun des modèles de prévision et sur la sélection des meilleurs modèles. Les statistiques présentées (MSE, MSE-F et R^2) font référence au test de MSE-F de McCracken (2007) présenté à la section 4.7.

	Agrégé	Univarié	Multivarié Backward	Multivarié Forward
MSE	0.1445%	0.1755%	0.1736%	0.1609%
MSE-F		-21.02005	-19.96878	-12.17714
R2		-0.21453	-0.20164	-0.11399

Dans un contexte *Out-Of-Sample* le modèle de prévision par agrégation sectorielle semble obtenir une meilleure performance notamment dans les premières années d'observation (2001 à 2007) grâce à l'exclusion des grandes variations affectant les dernières années de la période d'observation (2008 à 2011). Cette comparaison est représentée graphiquement à la figure 7 de la présente section.

FIGURE 7 – Comparaison des résultats du modèle de prévision par agrégation sectorielle

Cette figure compare les résultats obtenus par le modèle de prévision par agrégation sectorielle dans un contexte *In-Sample* de la section 6.2.5 et *Out-Of-Sample* de la section 6.2.6 avec la prime de risque réelle observée pour le marché canadien.



7 Conclusion

Il est possible de tirer plusieurs conclusions des résultats de cette étude. Tout d'abord, on observe que l'ensemble des modèles retenus, tant univariés, multivariés ou par agrégation sectorielle obtiennent un pouvoir explicatif supérieur au modèle naïf, et ce tant pour le marché américain que le marché canadien.

Pour le marché américain, le modèle multivarié *Stepwise Backward* obtient la meilleure performance du groupe étudié avec un R^2 de 0.10714 dans un contexte *In-Sample* et un R^2 de 0.22284 dans un contexte *Out-Of-Sample* selon le test de McCracken (2007). Ce modèle permet donc d'expliquer 22.28% de la variation mensuelle de la prime de risque américaine dans le contexte de l'échantillon observé dans cette étude. Il faut cependant noter qu'un problème de multicollinéarité affecte ce modèle pouvant biaiser sa validité dans le cadre de son application sur une période d'observation où la corrélation entre ces variables diminuerait substantiellement. Or, ce risque demeure faible étant donné que suite au retrait du modèle de l'une des variables problématiques, les coefficients de régression des variables restantes ne sont pas significativement affectés. De plus, chacune des variables problématiques est significative dans le modèle multivarié et le coefficient de détermination de la régression multivariée est affectée négativement de façon significative suite au retrait de l'une de ces variables. Le modèle en question est composé des variables économiques suivantes :

- **EINF** : Taux d'inflation annuel de la période vs la période précédente.
- **EM2G** : Croissance mensuelle de la masse monétaire M2 des États-Unis.
- **EGN2** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **2** ans.
- **EGN5** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **5** ans.
- **EGN10** : Dernier taux journalier de la période pour les obligations gouvernementales émises par la U.S. Federal Reserve ayant une échéance de **10** ans.

Ces cinq variables économiques ont toutes un impact important sur l'économie américaine et leur inclusion au niveau du modèle offrant la meilleure prévision de la prime de risque n'est pas surprenante. Il n'est pas difficile de croire qu'une variation au niveau de ces facteurs économiques peut influencer la prime de risque de marché du mois suivant.

Pour le marché canadien, le modèle multivarié *Stepwise Forward* obtient la meilleure performance du groupe étudié dans un contexte *In-Sample* avec un R^2 de 0.07760 selon le test de

McCracken (2007). Cependant, dans un contexte *Out-Of-Sample*, le modèle de prévision par agrégation sectorielle élaboré à la section 4.4 obtient de loin la meilleure performance avec un R^2 de 0.17773 selon le test de McCracken (2007), permettant donc d'expliquer 17.77% de la variation mensuelle de la prime de risque canadienne.

La performance accrue des modèles de prévision dans un contexte *Out-Of-Sample* semble provenir d'une meilleure performance notamment dans les premières années d'observation (2001 à 2007) grâce à l'exclusion des grandes variations affectant les dernières années de la période d'observation (2008 à 2011). Une étude approfondie de cet effet dans le futur pourrait possiblement offrir des solutions pour améliorer davantage le pouvoir explicatif des modèles étudiés.

Le modèle de prévision par agrégation sectorielle élaboré à la section 4.4 semble obtenir une meilleure performance dans le contexte du marché canadien en raison d'un nombre beaucoup plus restreint de regroupements industriels composant la construction de la prévision (48 vs 14). Dans le cadre d'une prochaine étude, une réduction du nombre de regroupements industriels pourrait potentiellement améliorer le pouvoir explicatif de ce modèle dans le cadre la prévision de la prime de risque américaine. Il serait également intéressant d'utiliser alternativement une agrégation des rendements des regroupements industriels sur une base équipondérée au lieu de la moyenne pondérée basée sur la capitalisation boursière et de comparer les résultats obtenus.

De plus, il existe définitivement un potentiel d'amélioration du pouvoir explicatif de l'ensemble des approches étudiées en testant l'impact de prévision de la prime de risque de marché de périodes subséquentes au mois suivant l'observation de la variable économique. Cette approche beaucoup plus complexe décèlerait possiblement davantage de corrélations en testant les relations entre variables économiques et primes de risque avec plus d'une période d'écart. Ceci permettrait d'effectuer des prévisions à moyen terme (e.g. un trimestre) et à long terme (e.g. un an).

Enfin, étant donné que la majorité des données utilisées dans cette étude ne sont pas distribuées normalement avec un niveau de confiance supérieur ou égal à 95%, il mériterait de répéter cette étude en faisant l'utilisation de tests non paramétriques lors de la génération des prévisions de la prime de risque de marché et de comparer les résultat obtenus. Aussi, il serait intéressant d'utiliser des moyennes mobiles exponentielles dans le calcul du modèle naïf et de comparer avec les résultats obtenus en utilisant ce nouveau modèle de référence dans un contexte *Out-of-Sample*.

Références

- [1] Baker, M. & Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *The Journal of Finance*, 55(5), 2219-2257.
- [2] Bodie, Z. (1976). Common stocks as a hedge against inflation. *The Journal of Finance*, 31(2), 459-470.
- [3] Campbell, J. Y. & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3), 195-228.
- [4] Campbell, J. Y. & Thompson, S. B. (2008). Predicting excess stock returns out of sample : Can anything beat the historical average?. *The Review of Financial Studies*, 21(4), 1509-1531.
- [5] Chan, L. K. C., Jegadeesh, N. & Lakonishok, J. (1995). Evaluating the performance of value versus glamour stocks : The impact of selection bias. *Journal of Financial Economics*, 38, 269-296.
- [6] Chen, N., Roll, R. & Ross, S. A. (1986). Economic forces and the stock market. *The Journal of Business*, 59(3), 383-403.
- [7] Chrétien, S. & Coggins, F. (2011). Information variables and equity premium predictability in canada. *Forthcoming*.
- [8] Cochrane, J. K. (2008). The Dog That Did Not Bark : A Defense of Return Predictability. *The Review of Financial Studies*, 21(4), 1533-1575.
- [9] Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2), 385-407.
- [10] D'Agostino, R. B., A. J. Belanger, & R. B. D'Agostino Jr. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician*, 44, 316-321.
- [11] Davis, J. L. (1994). The cross-section of realized stock returns : The pre-COMPUSTAT evidence. *The Journal of Finance*, 49(5), 1579-1593.
- [12] Dow, C. H. (1920). Scientific stock speculation. *The Magazine of Wall Street*.
- [13] Draper, N. R. & H. Smith. (1998). *Applied regression analysis* (pp. 307-312). Hoboken, NJ : Wiley-Interscience.
- [14] Efron, B. (1979). Bootstrap methods : Another look at the jackknife. *Annals of Statistics*, 7, 1-26.
- [15] Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *The American Economic Review*, 71(4), 545-565.

- [16] Fama, E. F. & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1), 3-25.
- [17] Fama, E. F. & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427-465.
- [18] Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.
- [19] Fama, E. F. & Schwert, G. W. (1977). Asset returns and inflation. *Journal of Financial Economics*, 5(2), 115-146.
- [20] Flood, R. P., Hodrick, R. J. & Kaplan, P. (1987). An evaluation of recent evidence on stock market bubbles. *National Bureau of Economic Research*.
- [21] Geske, R. & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *The Journal of Finance*, 38(1), 1-33.
- [22] Goetzmann, W. N. & Jorion, P. (1993). Testing the predictive power of dividend yields. *The Journal of Finance*, 48(2), 663-679.
- [23] Gordon, M. J. (1962). *The investment, financing and valuation of the corporation*. Homewood, IL : Irwin.
- [24] Keim, D. B. & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, 17(2), 357-390.
- [25] Kendall, M. G. (1953). The analysis of economic time-series-part I : Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1), 11-34.
- [26] Kothari, S. P. & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns : A time-series analysis. *Journal of Financial Economics*, 44(2), 169-203.
- [27] Lamont, O. (1998). Earnings and expected returns. *The Journal of Finance*, 53, 1563-1587.
- [28] Lintner, J. (1956). Distribution of incomes of corporations among dividends, retained earnings, and taxes. *The American Economic Review*, 46(2), 97-113.
- [29] McCracken, M. W. (2007). Asymptotics for out of sample tests of granger causality. *Journal of Econometrics*, 140(2), 719-752.
- [30] Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5), 867-887.
- [31] Miller, M. H. & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *The Journal of Business*, 34(4), 411-433.
- [32] Newey, W. K. & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703-708.

- [33] Rapach, D. E., Strauss, J. K. & Zhou, G. (2010b). Out-of-sample equity premium prediction : Combination forecasts and links to the real economy. *The Review of Financial Studies*, 23(2), 821-862.
- [34] Rapach, D. E., Strauss, J. K. & Zhou, G. (2010a). International Stock Return Predictability : What is the Role of the United States? *SSRN Working Paper*.
- [35] Rapach, D. E., Wohar, M. E. & Rangvid, J. (2005). Macro variables and international stock return predictability. *International Journal of Forecasting*, 21(1), 137-166.
- [36] Ross, S. A. (1972). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13, 341-360.
- [37] Shanken, J. (1990). Intertemporal asset pricing : An empirical investigation. *Journal of Econometrics*, 45(1), 99-120.
- [38] Shapiro, S. S. & Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67, 215.
- [39] Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika*, 52, 591-611.
- [40] Welch, I. & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4), 1455-1508.
- [41] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817-838.

A Description des industries par code SIC

TABLEAU 34 – Description des industries par code SIC

Description des industries selon le code SIC à deux chiffres utilisés dans le modèle de prévision par agrégation sectorielle décrit à la section 4.4.

Code SIC à deux chiffres	Description
10	Metal Mining
12	Coal Mining
13	Oil and Gas Extraction
14	Mining and Quarrying of Nonmetallic Minerals, Except Fuels
15	Building Cnstrctn - General Contractors & Operative Builders
16	Heavy Cnstrctn, Except Building Construction - Contractors
17	Construction - Special Trade Contractors
20	Food and Kindred Products
21	Tobacco Products
22	Textile Mill Products
23	Apparel, Finished Prdcts from Fabrics & Similar Materials
24	Lumber and Wood Products, Except Furniture
25	Furniture and Fixtures
26	Paper and Allied Products
27	Printing, Publishing and Allied Industries
28	Chemicals and Allied Products
29	Petroleum Refining and Related Industries
30	Rubber and Miscellaneous Plastic Products
31	Leather and Leather Products
32	Stone, Clay, Glass, and Concrete Products
33	Primary Metal Industries
34	Fabricated Metal Prdcts, Except Machinery & Transport Eqpmnt
35	Industrial and Commercial Machinery and Computer Equipment
36	Electronic, Elctrcal Eqpmnt & Cmpnts, Exept Computer Eqpmnt
37	Transportation Equipment
38	Mesr/Anlyz/Cntrl Instrmnts; Photo/Med/Opt Gds; Watches/Clocks
39	Miscellaneous Manufacturing Industries
40	Railroad Transportation
41	Local, Suburban Transit & Interurbn Hgwy Passenger Transport
42	Motor Freight Transportation
43	United States Postal Service
44	Water Transportation
45	Transportation by Air
46	Pipelines, Except Natural Gas
47	Transportation Services
48	Communications
49	Electric, Gas and Sanitary Services
50	Wholesale Trade - Durable Goods
51	Wholesale Trade - Nondurable Goods
52	Building Matrials, Hrdwr, Garden Supply & Mobile Home Dealsr
53	General Merchandise Stores
54	Food Stores
55	Automotive Dealers and Gasoline Service Stations
56	Apparel and Accessory Stores
57	Home Furniture, Furnishings and Equipment Stores
58	Eating and Drinking Places
59	Miscellaneous Retail
60	Depository Institutions
61	Nondepository Credit Institutions
62	Security & Commodity Brokers, Dealers, Exchanges & Services
63	Insurance Carriers
64	Insurance Agents, Brokers and Service
65	Real Estate
67	Holding and Other Investment Offices
70	Hotels, Rooming Houses, Camps, and Other Lodging Places
72	Personal Services
73	Business Services
75	Automotive Repair, Services and Parking
76	Miscellaneous Repair Services
78	Motion Pictures
79	Amusement and Recreation Services
80	Health Services
81	Legal Services
82	Educational Services
83	Social Services
87	Engineering, Accounting, Research, Management & Related Svcs
99	Nonclassifiable Establishments

B Statistiques descriptives des industries de départ

TABLEAU 35 – Statistiques descriptives des industries de départ (É-U)

Ce tableau présente les statistiques descriptives pour chacune des industries de départ (regroupés par code SIC) du marché américain générées par cette étude pour la période d'échantillon. On y retrouve, en ordre, la moyenne, l'écart-type et les valeurs minimales et maximales.

Variable	Mean	Std. Dev.	Min.	Max.
IND10	0.02	0.081	-0.28	0.43
IND12	0.02	0.104	-0.364	0.378
IND13	0.018	0.082	-0.26	0.313
IND14	0.011	0.065	-0.279	0.181
IND15	0.02	0.094	-0.235	0.339
IND16	0.017	0.082	-0.271	0.282
IND17	0.016	0.102	-0.304	0.391
IND20	0.012	0.041	-0.147	0.152
IND21	0.014	0.072	-0.246	0.297
IND22	0.016	0.094	-0.269	0.61
IND23	0.017	0.065	-0.21	0.307
IND24	0.011	0.072	-0.309	0.255
IND25	0.017	0.068	-0.313	0.491
IND26	0.011	0.052	-0.144	0.254
IND27	0.01	0.05	-0.175	0.219
IND28	0.014	0.04	-0.098	0.154
IND29	0.013	0.052	-0.123	0.169
IND30	0.018	0.065	-0.193	0.272
IND31	0.026	0.089	-0.26	0.43
IND32	0.015	0.085	-0.22	0.318
IND33	0.019	0.09	-0.352	0.388
IND34	0.013	0.058	-0.239	0.255
IND35	0.021	0.075	-0.253	0.226
IND36	0.021	0.083	-0.276	0.312
IND37	0.015	0.056	-0.168	0.273
IND38	0.017	0.05	-0.191	0.222
IND39	0.017	0.067	-0.187	0.275
IND40	0.015	0.06	-0.207	0.211
IND42	0.012	0.055	-0.202	0.238
IND44	0.017	0.08	-0.286	0.312
IND45	0.015	0.07	-0.208	0.198
IND46	0.014	0.045	-0.127	0.167
IND47	0.02	0.066	-0.177	0.259
IND48	0.018	0.076	-0.141	0.655
IND49	0.013	0.074	-0.389	0.629
IND50	0.018	0.055	-0.197	0.212
IND51	0.015	0.05	-0.208	0.16
IND52	0.016	0.077	-0.183	0.288
IND53	0.013	0.06	-0.164	0.222
IND54	0.013	0.054	-0.175	0.147
IND55	0.022	0.074	-0.184	0.236
IND56	0.021	0.077	-0.233	0.233
IND57	0.024	0.096	-0.26	0.439
IND58	0.016	0.052	-0.136	0.196
IND59	0.02	0.059	-0.16	0.222
IND60	0.014	0.06	-0.233	0.257
IND61	0.014	0.087	-0.324	0.509
IND62	0.021	0.088	-0.286	0.346
IND63	0.014	0.058	-0.261	0.212
IND64	0.014	0.061	-0.218	0.416
IND65	0.018	0.067	-0.334	0.447
IND67	0.013	0.04	-0.167	0.194
IND70	0.014	0.075	-0.252	0.418
IND72	0.014	0.067	-0.217	0.3
IND73	0.022	0.075	-0.187	0.37
IND75	0.016	0.112	-0.318	0.563
IND78	0.017	0.073	-0.277	0.35
IND79	0.027	0.102	-0.3	0.899
IND80	0.017	0.066	-0.206	0.239
IND81	0.003	0.094	-0.311	0.369
IND82	0.029	0.092	-0.205	0.478
IND83	0.024	0.159	-0.466	0.857
IND87	0.024	0.07	-0.182	0.567
IND99	0.013	0.059	-0.205	0.212
N	250			

TABLEAU 36 – Statistiques descriptives des industries de départ (Canada)

Ce tableau présente les statistiques descriptives pour chacun des industries de départ (regroupés par code SIC) du marché canadien générés par cette étude pour la période d'échantillon. On y retrouve, en ordre, la moyenne, l'écart-type et les valeurs minimales et maximales.

Variable	Mean	Std. Dev.	Min.	Max.
IND10	0.02	0.09	-0.37	0.452
IND12	0.024	0.16	-0.416	0.639
IND13	0.02	0.072	-0.244	0.292
IND14	0.045	0.16	-0.394	0.889
IND15	0.023	0.096	-0.306	0.507
IND16	0.016	0.113	-0.319	0.6
IND17	0.003	0.062	-0.289	0.459
IND20	0.015	0.046	-0.147	0.188
IND21	0.014	0.057	-0.356	0.206
IND22	0.008	0.116	-0.327	0.662
IND23	0.013	0.133	-0.333	0.607
IND24	0.014	0.076	-0.211	0.252
IND25	0.011	0.076	-0.293	0.415
IND26	0.012	0.103	-0.342	0.687
IND27	0.007	0.053	-0.173	0.326
IND28	0.022	0.072	-0.254	0.329
IND29	0.014	0.061	-0.2	0.258
IND30	0.013	0.072	-0.207	0.237
IND31	-0.006	0.08	-0.547	0.625
IND32	0.01	0.072	-0.181	0.251
IND33	0.023	0.111	-0.402	0.48
IND34	0.014	0.062	-0.205	0.184
IND35	0.019	0.066	-0.171	0.275
IND36	0.017	0.14	-0.452	0.642
IND37	0.014	0.078	-0.314	0.294
IND38	0.028	0.128	-0.38	0.842
IND39	0.011	0.123	-0.516	0.680
IND40	0.014	0.063	-0.165	0.208
IND41	0	0.042	-0.51	0.152
IND42	0.016	0.07	-0.232	0.295
IND44	0.015	0.059	-0.218	0.176
IND45	0.015	0.095	-0.224	0.451
IND46	0.011	0.032	-0.137	0.131
IND48	0.014	0.056	-0.269	0.255
IND49	0.013	0.045	-0.125	0.46
IND50	0.016	0.057	-0.193	0.172
IND51	0.012	0.054	-0.155	0.214
IND52	-0.012	0.091	-0.405	0.478
IND53	0.01	0.079	-0.282	0.223
IND54	0.013	0.05	-0.192	0.162
IND55	0.009	0.074	-0.285	0.252
IND56	0.021	0.086	-0.287	0.343
IND57	0.015	0.057	-0.167	0.277
IND58	0.012	0.068	-0.202	0.395
IND59	0.02	0.091	-0.189	0.676
IND60	0.014	0.052	-0.283	0.156
IND61	0.011	0.044	-0.193	0.117
IND62	0.018	0.066	-0.289	0.272
IND63	0.015	0.057	-0.217	0.237
IND64	0.014	0.137	-0.457	0.75
IND65	0.015	0.056	-0.244	0.143
IND67	0.02	0.05	-0.205	0.211
IND70	0.009	0.067	-0.209	0.471
IND72	-0.001	0.065	-0.827	0.357
IND73	0.025	0.103	-0.312	0.643
IND75	0.015	0.117	-0.396	0.462
IND78	0.014	0.07	-0.195	0.451
IND79	0.012	0.093	-0.256	0.409
IND80	0.017	0.142	-0.318	0.695
IND82	-0.002	0.12	-0.618	0.613
IND87	0.02	0.07	-0.219	0.284
N	238			