

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

TEST EFFICACE DU MÉDAF SOUS L'HYPOTHÈSE DE DISTRIBUTION À  
SYMÉTRIE ELLIPTIQUE DES RENDEMENTS

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN ÉCONOMIQUE

PAR  
RICARDO DUBRÉSIL

JUILLET 2015

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

Je voudrais remercier mon directeur de recherche, M. Douglas Hodgson, pour son aide et sa disponibilité dans la réalisation de ce travail.

Je voudrais aussi témoigner ma reconnaissance envers les professeurs Alain Paquet et Alain Guay, qui ont accepté d'être mes lecteurs critiques et m'ont fourni des recommandations d'une précieuse utilité pour la finalisation du mémoire.

J'exprime toute ma gratitude à l'endroit de l'Université du Québec à Montréal pour la formation reçue ainsi qu'à tous les membres de l'administration du département des sciences économiques qui m'ont offert leur soutien au cours de mon cheminement à la maîtrise en économie financière. Un mot de remerciement spécial à madame Martine Boisselle pour son assistance.

Je remercie aussi mes camarades de classe et étudiants qui ont eu la générosité d'échanger avec moi des idées, documents et outils pertinents.

Merci à ma famille et à mes amis pour leurs encouragements.

## TABLE DES MATIÈRES

LISTE DES FIGURES.....	vi
LISTE DES TABLEAUX.....	vii
LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES .....	viii
RÉSUMÉ .....	ix
INTRODUCTION .....	1
CHAPITRE I	
PRÉSENTATION DU MÉDAF .....	4
1.1 Le MÉDAF.....	4
1.1.1 Fondements théoriques.....	4
1.1.2 Limites et extensions.....	7
1.2 Test du MÉDAF.....	9
1.2.1 Les premiers tests des années 70.....	9
1.2.2 La critique de Roll (1977).....	11
1.2.3 Tests modernes du MÉDAF.....	12
1.3 Test de Gibbons, Ross et Shanken (1989).....	12
1.3.1 Test en grand échantillon et en échantillon fini.....	12
1.3.2 Interprétation économique du test.....	15
CHAPITRE II	
GÉNÉRALISATION DU MÉDAF AUX LOIS À SYMÉTRIE ELLIPTIQUE .....	16
2.1 Non normalité des rendements et implications pour le MÉDAF.....	16
2.1.1 Évidence empirique de la non normalité et alternatives.....	16
2.1.2 Conséquences sur le MÉDAF.....	17

2.2	Généralisation du MÉDAF avec la symétrie elliptique.....	18
2.2.1	Définition des lois à symétrie elliptique.....	18
2.2.2	Compatibilité avec le MÉDAF.....	20
2.3	Test GRS sous l'hypothèse de symétrie elliptique. ....	21
2.3.1	L'approche paramétrique. ....	21
2.3.2	L'approche semi-paramétrique. ....	22
2.3.3	Correction de Zhou (1993) sur la statistique de Wald. ....	23
2.3.4	L'approche de Beaulieu, Dufour et Khalaf (2002). ....	23
CHAPITRE III		
CADRE MÉTHODOLOGIQUE : LA PROCÉDURE H.L.V. (2002) .....		25
3.1	Fondements statistiques de la procédure H.L.V. (2002). ....	25
3.1.1	Principes de l'estimation adaptative.....	25
3.2.1	Implications de la symétrie elliptique pour l'estimation adaptative. ....	29
3.2	Procédure du test semi-paramétrique de H.L.V. (2002).....	30
3.2.1	Mise en œuvre de l'estimation. ....	30
3.2.2	Conditions nécessaires à l'adaptabilité. ....	32
3.2.3	Test de Wald adaptatif.....	35
3.4	Test de symétrie elliptique de Li, Fang et Zhu (1997). ....	35
3.4.1	Principe du test. ....	35
3.4.2	Mise en oeuvre du test. ....	36
CHAPITRE IV		
PRÉSENTATION DES RÉSULTATS ET ANALYSES .....		38
4.1	Description de l'échantillon.....	38
4.2	Traitement des données.....	39
4.1	Statistiques descriptives des données. ....	41
4.4	Résultats et analyses. ....	43

4.4.1 Résultats de la procédure semi-paramétrique.....	45
4.4.2 Sensibilité de l'efficacité semi-paramétrique.....	46
CONCLUSION.....	48
APPENDICE A	
STATISTIQUES DESCRIPTIVES.....	50
APPENDICE B	
TESTS D'ADÉQUATION STATISTIQUE ET ANALYSES.....	53
BIBLIOGRAPHIE	

## LISTE DES FIGURES

Figure	Page
1.1 Frontières efficientes, <i>CML</i> et <i>SML</i> . .....	7
2.1 Distributions elliptiques symétriques et asymétriques. ....	19
4.1 Test de Li, Fang et Zhu (1997) sur les 80 actions. ....	43
B.1 Test de Li, Fang et Zhu (1997), simulations. ....	56

## LISTE DES TABLEAUX

Tableau	Page
4.1 Statistiques descriptives des portefeuilles.....	41
4.2 Test de Mardia (1970) sur les 3 portefeuilles.....	42
4.3 Estimation paramétrique du MÉDAF. ....	44
4.4 Estimation semi-paramétrique du MÉDAF. ....	45
4.5 Tests paramétriques vs semi-paramétriques. ....	46
A.1 Statistiques descriptives et $\beta$ des actifs. ....	50
B.1 Test de normalité sur les résidus des modèles paramétriques. ....	53
B.2 Tests d'hétéroscédasticité sur les modèles paramétriques.....	54
B.3 Tests d'autocorrélation sur les modèles paramétriques.....	54
B.4 Tests de RESET sur les modèles paramétriques. ....	55
B.5 Test de stabilité de Quandt-Andrews. ....	55
B.6 Analyse de sensibilité de l'estimation semi-paramétrique. ....	56

## LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AMEX	American Stock Exchange
ARCH	Autoregressive Conditional Heteroskedasticity
BJS	Black, Jensen et Scholes
CANSIM	Canadian Socio-economic Information Management
CML	Capital Market Line
D.D.L	Degrés de liberté
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GMM	Generalized Method of Moments
GRS	Gibbons Ross Shanken
H.L.V.	Hodgson, Linton et Vorkink
I.I.D.	Indépendamment et identiquement distribués.
MCO	Moindres Carrés Ordinaires
MÉDAF	Modèle d'Évaluation des Actifs Financiers
NASDAQ	National Association of Securities Dealers Automated Quotations
NYSE	New York Stock Exchange
SML	Securities Market Line
S&P TSX	Standard and Poor Toronto Stock Exchange
RAI	Régressions Apparemment Indépendantes

## RÉSUMÉ

Dans le présent mémoire, le MÉDAF non conditionnel est testé avec un modèle semi-paramétrique qui suppose les rendements distribués selon la famille de lois elliptiques symétriques. La normalité multivariée imposée par le modèle de Sharpe-Lintner et Mossin est incompatible avec les données empiriques. Une erreur de spécification sur leur distribution expose le test conventionnel du MÉDAF à un risque de perte d'efficacité et de puissance.

Plusieurs auteurs, dont Owen et Rabinovitch (1983) et Chamberlain (1983) ont montré la cohérence du MÉDAF avec cette classe de lois, qui impose très peu de restrictions sur la forme fonctionnelle de la distribution. Aussi, une application du test de Gibbons, Ross et Shanken (1989) sous cette conjecture, si elle est confirmée, peut permettre une inférence plus robuste. Conceptuellement, un modèle semi-paramétrique est nécessaire pour établir une solution optimale dans ce contexte. Il ne requiert en effet, qu'une spécification partielle de la distribution conditionnelle de la variable endogène pour en estimer les paramètres d'intérêt.

Selon Stein (1956) et Bickel (1982), lorsqu'un tel estimateur est asymptotiquement optimal, il est adaptatif. Cette propriété est vérifiable en termes de convergence en probabilité des dérivées première et seconde de la fonction log-vraisemblance des paramètres du modèle semi-paramétrique sur un ensemble de solutions qu'il admet. En principe, avec une estimation non paramétrique efficace de la distribution inconnue, il est possible de satisfaire cette condition. Cette efficacité est toutefois, inversement proportionnelle à la dimension de la distribution à estimer et exige un très grand nombre d'observations pour être garantie. Hodgson, Linton et Vorkink (2002) s'appuient sur les propriétés de la famille à symétrie elliptique pour réduire la dimension de la loi à estimer et construire une estimation adaptative qui rejette le MÉDAF avec un test Wald plus précis.

Ce travail propose un exposé des fondements statistiques et économétriques de la procédure de Hodgson, Linton et Vorkink (2002). Ses limitations ainsi que des stratégies d'amélioration y sont discutées. L'hypothèse de symétrie elliptique est testée selon l'approche de Li, Fang et Zhu (1997), qui est d'une relative facilité de mise en œuvre comparativement aux autres méthodes existantes. La méthodologie est appliquée à un échantillon canadien qui est plus récent et plus large que celui d'origine. La symétrie elliptique n'est pas rejetée. Le MÉDAF, est estimé et testé avec des coefficients plus précis mais n'est pas validé par l'échantillon. Une analyse de sensibilité des résultats par rapport au nombre de portefeuilles, au nombre de titres par portefeuille et la taille de l'échantillon est aussi présentée et révèle leur effet bénéfique sur la procédure.

**MOTS-CLÉS:** semi-paramétrique, estimateur adaptatif, Modèle d'Évaluation des Actifs Financiers, distributions à symétrie elliptique, test d'efficacité moyenne-variance.

## INTRODUCTION

Le présent travail a pour but de montrer, qu'avec une estimation semi-paramétrique efficace en présence de distributions à symétrie elliptique des rendements, la précision du test de Gibbons, Ross et Shanken (1989) du MÉDAF peut être améliorée significativement. Le résultat attendu n'est donc pas la confirmation des prédictions du MÉDAF, mais plutôt une plus grande précision du test, lorsque l'hypothèse de symétrie elliptique est empiriquement confirmée. La procédure de Hodgson, Linton et Vorkink (2002) constitue le cadre méthodologique du mémoire.

Le cadre conceptuel le plus moderne utilisé pour tester le MÉDAF de Sharpe-Lintner-Mossin ou de Black (1972) est l'approche de Gibbons, Ross et Shanken (1989), qui sera exposée plus loin en détail. La performance de cette approche est étroitement liée aux hypothèses sur la distribution jointe des rendements. En dépit des faits empiriques signalés par Mandelbrot (1963) et Fama (1965), prouvant l'inadéquation de la loi normale multivariée, celle-ci a largement servi de paradigme statistique pour la construction des tests du MÉDAF dans les années 70 et 80. MacKinley et Richardson (1991) et Zhou (1993) ont montré dans quelle mesure cette erreur sur la vraie distribution biaise les statistiques du test de Gibbons, Ross et Shanken (1989) et en altère la puissance.

Plusieurs lois et familles de lois ont été proposées dans la littérature pour généraliser le MÉDAF. Entre autres, Owen et Rabinovitch (1983), Chamberlain (1983), et Beck (1997) ont soutenu la pertinence de la famille à symétrie elliptique. Celle-ci regroupe des lois et familles de lois telles la normale, la  $t$  de Student, certaines combinaisons d'échelle de distributions normales et les lois Pareto-stables couramment utilisées en économie financière. Dans les années 90, des tests compatibles avec cette hypothèse ont été développés. Ils reposaient cependant sur une estimation paramétrique procédant par un choix parcimonieux d'une distribution *à priori* dans la famille, pour spécifier l'estimateur du Maximum de Vraisemblance. Les distributions  $t$  de Student ont retenu une attention particulière en raison

adéquation empirique, notamment avec l'excès de kurtosis. Toutefois, le risque d'erreur de spécification n'a jamais été entièrement écarté. Comme l'a montré Sentana (2008), les répercussions sur l'inférence statistique peuvent être aussi néfastes que l'imposition à tort de la normalité multivariée.

Il est techniquement possible de construire un estimateur du Maximum de Vraisemblance à partir d'une estimation non paramétrique de la vraie distribution. Un estimateur non paramétrique est conçu pour laisser les données dicter elles-mêmes, la forme de la distribution. Dans le cadre multivarié, son efficacité décline proportionnellement et rapidement avec la dimension de la loi à estimer, si un très grand nombre d'observations n'est pas utilisé. Ce problème, appelé le Fléau de la Dimension, a en quelque sorte limité les développements de tests du MÉDAF selon une approche non paramétrique. Cette problématique a été adressée par Hodgson, Linton et Vorkink (2002) qui ont proposé une stratégie de réduction de la dimension de la distribution à estimer qui impose la symétrie elliptique. Ils parviennent ainsi à estimer le MÉDAF de façon adaptative, avec des écart-type en moyenne 15% plus faibles que ceux fournis par les MCO. Grâce à des simulations de Monte-Carlo, ils ont mis en évidence une puissance accrue du test construit sur cette base.

Le travail de Hodgson, Linton et Vorkink (2002) a porté sur un échantillon de rendements journaliers annualisés de firmes cotées au NYSE, NASDAQ et AMEX de janvier 1996 à décembre 1997. Jusque-là, cette méthodologie n'a pas été implémentée avec un échantillon canadien. De plus, une période relativement plus longue et récente pourrait conduire à une plus grande efficacité du test. Un échantillon de 80 actions ordinaires membres de l'indice composite du S&P TSX sur la période allant de janvier 2009 à octobre 2014 est donc utilisé. Une analyse de sensibilité des résultats par rapport à la variation de la taille de l'échantillon ainsi qu'au nombre de portefeuilles utilisés et leur composition est proposée. Finalement, un test de symétrie elliptique est effectué selon une approche plus intuitive et relativement plus facile de mise en œuvre que celle du travail original.

Le mémoire s'organise en quatre chapitres. Le premier est consacré à la présentation du modèle faisant l'objet du test ainsi qu'à une discussion sur ses limites. Les avancées clef dans

l'histoire des tests du MÉDAF unifactoriel sont aussi dégagées. Le second chapitre propose une revue de littérature sur la problématique de la validité théorique et empirique du MÉDAF en présence de rendements à symétrie elliptique. Les propriétés de cette famille de lois, pertinentes à l'analyse, sont discutées. Les implications de la non normalité et plus précisément de la symétrie elliptique sur le test de Gibbons, Ross et Shanken (1989) sont aussi présentées. Le chapitre trois expose en détail, la procédure semi-paramétrique de Hodgson, Linton et Vorkink (2002). Les concepts statistiques et économétriques fondamentaux de cette approche, ses limitations et aussi les opportunités d'amélioration sont discutées. Le dernier chapitre détaille la méthodologie d'implémentation de la procédure à l'échantillon canadien et présente les principaux résultats et analyses.

## CHAPITRE I

### PRÉSENTATION DU MÉDAF

#### 1.1 Le MÉDAF.

##### 1.1.1 Fondements théoriques.

Le MÉDAF est un modèle de détermination de la rentabilité des titres financiers en fonction de leur niveau de risque associé au marché ou risque systémique. C'est l'un des modèles de référence en économie financière et en finance appliquée. Les fondements du MÉDAF prennent corps dans la théorie Moyenne-Variance élaborée par Markovitz (1956). La version canonique fût introduite par Treynor (1962), Lintner (1962), Sharpe (1964) et Mossin (1966). Eu égard à ses hypothèses restrictives et irréalistes, cette version dite de Sharpe-Lintner-Mossin a fait l'objet de nombreuses critiques et, ultérieurement, a subi plusieurs déclinaisons.

La dérivation du MÉDAF traditionnel fait intervenir plusieurs hypothèses. L'efficience des marchés est imposée en vue de créer des conditions favorables à la réalisation de l'équilibre général. Les résultats spécifiques à la version de Sharpe-Lintner-Mossin s'obtiennent en supposant l'absence de restrictions sur les ventes à découvert et l'emprunt qui sont illimités et accessibles à un taux unique : le taux sans risque. L'efficience des marchés implique aussi, pour ces auteurs, l'atomicité des investisseurs individuels qui forment des anticipations homogènes avec une connaissance commune de la distribution de probabilité des rendements financiers. Avec une fonction d'utilité unipériodique et quadratique, compatible avec l'aversion au risque, ces derniers résolvent leur problème d'optimisation (ou de sélection de

portefeuille) exclusivement en fonction de la moyenne et de la variance des titres risqués. La combinaison de tous ces portefeuilles optimaux détenus à l'équilibre, détermine le portefeuille de marché, qui admet dans sa composition les actifs échangeables, négociables et parfaitement divisibles de toute sorte, par exemple, les œuvres d'art ou le capital humain.

Treynor (1962), Sharpe (1964), Lintner (1965) et Mossin (1966) ont donné forme au MÉDAF en articulant leur démonstration sur ces conditions. La Théorie Moyenne-Variance prédit qu'à l'optimum, tout investisseur accomplit la diversification<sup>1</sup> de portefeuille. Dans l'ensemble des portefeuilles optimaux réalisables dans l'économie, appelée frontière moyenne-variance, seuls ceux économiquement efficaces (ou appartenant à la frontière efficiente<sup>2</sup>), sont effectivement désirés. La contribution de Sharpe, Lintner et Mossin a consisté premièrement à introduire l'hypothèse d'homogénéité anticipations et des préférences pour montrer l'unicité et l'efficacité globalement maximale du portefeuille de marché. Celle-ci sert d'ailleurs d'étalon de mesure à l'efficacité de tout autre portefeuille désiré et est une fonction linéaire de sa covariance avec ce dernier. En vertu du Théorème de Séparation en Deux Fonds de Tobin (1952), les auteurs ont établi qu'à l'équilibre, même le plus averse au risque des investisseurs détiendra une portion non nulle de ce portefeuille.

À l'équilibre, l'offre excédentaire du titre sans risque est nulle et pour chaque individu, toute opportunité d'investissement financier se décompose en une combinaison linéaire optimale de ce dernier avec le portefeuille de marché.

La version canonique du MÉDAF détermine l'espérance de rendement de tout titre efficace  $J$ , par la relation

$$E(\bar{R}_J) = Rf + \beta_J [E(\bar{R}_M) - Rf],$$

<sup>1</sup> La diversification consiste, pour un investisseur, à déterminer dans un portefeuille, l'allocation optimale des titres qui réduit au minimum son niveau de risque total par rapport au rendement attendu. Le risque ainsi diversifié est appelé risque de marché ou risque systématique.

<sup>2</sup> La notion de frontière efficiente a été introduite par Markowitz (1952) et désigne un ensemble de portefeuilles, de la frontière moyenne-variance qui maximisent le ratio rendement anticipé-volatilité (ou écart-type). Dans le MÉDAF, l'efficacité d'un actif  $A$  se mesure par son ratio de Sharpe donné par  $\frac{E(\bar{R}_A) - Rf}{[\sigma(\bar{R}_A)]}$  où  $E(\bar{R}_A) - Rf$  est le rendement excédentaire attendu du titre par rapport au taux sans risque  $Rf$  et  $\sigma(\bar{R}_A)$  sa volatilité sur la période.

appelée Ligne Caractéristique des Titres ou *SML*, où le taux sans risque et le rendement espéré du portefeuille de marché sont respectivement représentés par  $Rf$  et  $E(\tilde{R}_M)$  et le coefficient  $\beta_J$  mesure le niveau de risque systémique associé au titre  $J$  selon  $\beta_J = \frac{\sigma(\tilde{R}_J, \tilde{R}_M)}{\sigma^2(\tilde{R}_M)}$ . Par cette relation, tout titre non corrélé avec le portefeuille de marché est nécessairement assimilable au taux sans risque.

L'équation précédente découle d'un autre résultat du modèle appelé la Droite de Marché des Capitaux ou *CML* donnée par

$$E(\tilde{R}_J) = Rf + [E(\tilde{R}_M) - Rf] \frac{\sigma(\tilde{R}_J)}{\sigma(\tilde{R}_M)},$$

où  $\sigma(\tilde{R}_J)$  et  $\sigma(\tilde{R}_M)$  désignent respectivement les volatilités du titre  $J$  et du portefeuille de marché. Elle pose une relation linéaire entre le rendement espéré et la volatilité de tout titre  $J$  quelconque. Par cette équation, le MÉDAF établit aussi une mesure intuitive du degré d'aversion au risque des agents dans l'économie. La disponibilité du taux sans risque dans l'économie élargit l'ensemble des possibilités d'investissements financiers, en altérant la forme de la frontière efficiente de Markowitz (1952) en une demi-droite  $[Rf, M]$ , d'ordonnée à l'origine  $Rf$ .

La Ligne Caractéristique des Titres et la Droite de Marché des Capitaux<sup>3</sup>, les frontières efficientes et le portefeuille de marché sont représentés à la figure 1.1 ci-dessous. La partie de gauche de cette figure illustre la Droite de Marché des Capitaux, où le point  $M$  désigne le portefeuille de marché (ou portefeuille de tangence) de paramètres  $(\mu_M, \sigma_M)$  sur l'espace moyenne-variance défini par le repère d'axes  $(\mu, \sigma)$ . D'une part,  $M$  est le point de tangence entre la *CML* et la frontière efficiente de Markowitz (1956), la section supérieure de la frontière Moyenne-Variance (désignée par *EVM*) partant du point  $O^4$ , et d'autre part, entre la *CML* et la fonction d'utilité concave (illustrée par la courbe en pointillés, notée  $V_M$ ) de l'investisseur représentatif averse au risque. L'égalité des pentes de la *CML* et de  $V_M$ , au

<sup>3</sup> *SML* et *CML* sont des abréviations des expressions anglaises *Securities Market Line* et *Capital Market Line* désignant respectivement la Ligne Caractéristique des Titres et la Droite de Marché des Capitaux.

<sup>4</sup>  $O$  représentait, selon Markovitz (1952) le portefeuille efficient à variance globalement minimale.

point d'équilibre  $M$ , permet mesurer le degré d'aversion au risque dans l'économie par la pente de la  $CML$  donnée par  $\frac{\sigma(\bar{R}_I)}{\sigma(\bar{R}_M)}$ . Sur la section de droite de la figure, est représentée la Ligne Caractéristique des Titres, notée  $SML$ , qui pose une relation linéaire et positive entre le rendement espéré  $\mu$  et le niveau de risque systématique de tout actif. On en déduit que  $\beta_M$  est toujours égal à un pour le portefeuille de marché.

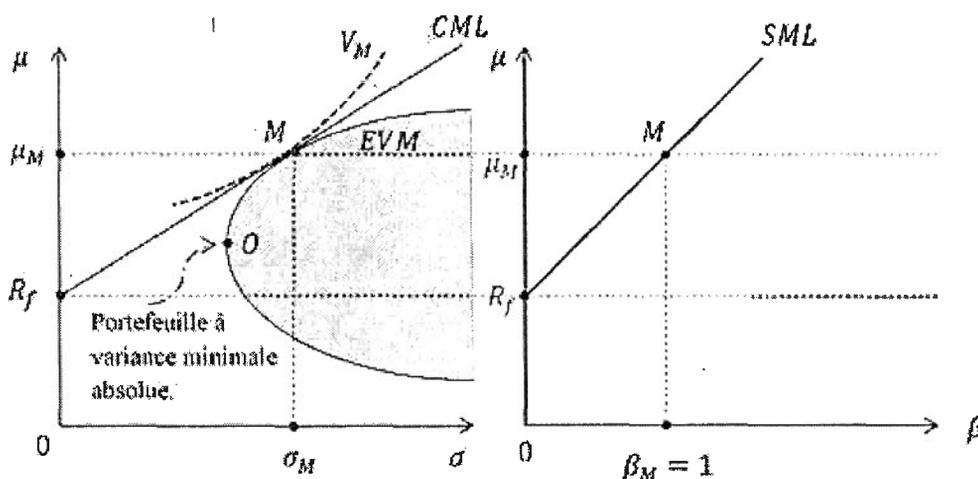


Figure 1.1 Frontières efficaces,  $CML$  et  $SML$ .

### 1.1.2 Limites et extensions.

Sur le plan théorique, le MÉDAF est un modèle logiquement cohérent. Les critiques portent sur la validité empirique de ses hypothèses et résultats. L'une des principales<sup>5</sup> faiblesses réside dans l'hypothèse d'efficacité des marchés, qui n'est jamais vérifiée dans toute l'acception du terme. En ce sens, les innovations apportées au MÉDAF original se sont articulées dans deux directions. On distingue les versions dites non conditionnelles, qui relâchent certaines des conditions d'efficacité, tout en conservant un horizon temporel

<sup>5</sup> D'autres conditions telles la normalité multivariée des titres, la fonction d'utilité quadratique, l'inobservabilité du portefeuille de marché ont aussi été remises en question.

unipériodique pour l'analyse et celles du MÉDAF conditionnel qui ajoutent à la prime de risque de marché d'autres variables pertinentes, lesquelles dictent périodiquement la formation des anticipations des agents et induisent à long terme une variation dans les  $\beta$ .

Parmi les modèles unipériodiques, on peut citer celui avec les anticipations hétérogènes de Lintner (1969), celui sans les ventes à découvert illimitées de Vasicek (1971) et le MÉDAF zéro-beta de Black (1972). Dans ce dernier,  $Rf$  est remplacé par un titre dont le rendement espéré  $E(\tilde{R}_{zc\{m\}}) > Rf$ , est non corrélé avec le marché selon

$$E(\tilde{R}_J) = E(\tilde{R}_{zc\{m\}}) + \beta_J[E(\tilde{R}_M) - E(\tilde{R}_{zc\{m\}})].$$

Brennan (1970) a été le premier à incorporer la taxation et Lindenberg (1979) à prendre en compte la présence d'investisseurs influents. Ils aboutissent tous à la linéarité de la *SML* et prédisent une prime de risque de marché strictement positive. Cependant, ils n'arrivent pas à expliquer certaines anomalies<sup>6</sup> observées dans les rendements qui contredisent la notion d'efficacité du portefeuille de marché. De plus, Ross (1977) a montré qu'il y a une limitation au nombre de conditions d'efficacité pouvant être relâchées simultanément si on veut aboutir aux mêmes résultats du MÉDAF. Aussi, aucune de ces versions n'a caractérisé la non efficacité des marchés sous tous ses angles envisageables et conséquemment n'est apte à expliquer tous les traits comportementaux des rendements boursiers. Plus tard, Ross (1978) va s'inspirer du MÉDAF pour développer le Modèle d'Évaluation par Arbitrage qui n'impose pas l'efficacité du marché et établit une relation rendement-risque multifactorielle de la même nature.

D'autres modèles multifactoriels et multipériodiques ont aussi été introduits. On peut citer entre autres, celle de Friend, Landskroner et Losq (1976) avec l'ajout du taux d'inflation anticipé, celle de Mayers (1972) avec les titres non négociables comme le stock de capital humain, celle de Basu (1977) incluant le ratio cours-bénéfice, celle de Banz (1981) démontrant l'importance de la taille de la firme et de Bhandari (1988) incorporant le ratio dette-valeur marchande de la firme. Le MÉDAF à trois facteurs de Fama et French (1993)

---

<sup>6</sup> Il s'agit de rendements de titres qui systématiquement accusent un ratio de Sharpe plus élevé que celui du portefeuille de marché.

constitue une généralisation de cette approche. Projetant l'analyse moyenne-variance sur un horizon multipériodique, Lucas (1978), Breeden (1979) et Merton (1973) ont été les pionniers du MÉDAF inter temporel, basé par exemple sur la consommation avec Lucas (1978), ou la production avec Cochrane (1991).

## 1.2 Test du MÉDAF.

### 1.2.1 Les premiers tests des années 70.

Dans les années 70, les premiers tests cherchaient à valider les principaux résultats des modèles de Sharpe-Lintner-Mossin et de Black (1972). Il s'agit de la linéarité de l'équation de l'équation de la *SML*, la positivité de la prime de risque de marché, son unicité comme variable explicative et finalement, le fait que tout titre non corrélé avec le portefeuille de marché soit identique au taux sans risque. Initialement, trois principales méthodologies adaptées à différentes structures d'échantillon ont été proposées. Blume et Friend (1973) ont développé une procédure pour des données en coupe transversale<sup>7</sup>. Black, Jensen et Scholes (1972) ont construit un test incorporant des séries chronologiques<sup>8</sup> et Fama et MacBeth (1973) ont proposé une adaptation du test de Blume et Friend (1973) aux données de panel. La linéarité de la *SML* est unanimement validée, de même que la prime de risque de marché positive. Les conclusions sont toutefois partagées sur les autres résultats du MÉDAF. Dans une discussion, Huang et Litzenberg (1993), montrent qu'empiriquement c'est la version de Black (1972) qui a connu le plus de succès. Un bref survol des deux premières approches est donné dans les lignes qui suivent.

Le test de Blume et Friend (1970) est construit à partir du modèle suivant estimé par MCO

---

<sup>7</sup> Cela signifie que l'échantillon est composé de titres individuels dont les rendements sont observés sur une seule période.

<sup>8</sup> L'échantillon est composé ici d'un ensemble d'observations de rendements réalisés pour chaque un actif sur un intervalle de temps de plusieurs périodes. Dans le cas de données de panel, les rendements moyens de plusieurs titres sont observés sur plusieurs périodes simultanément.

$$(\bar{R}_j - R_f) = \gamma_0 + \gamma_1 b_j + \eta_j \text{ avec } j = [1, 2, 3, \dots, N].$$

$\bar{R}_j$  et  $R_f$  sont respectivement le rendement moyen de chaque titre  $j$  et le taux sans risque moyen sur la période échantillonnale de  $T$  observations.  $b_j$  est une variable construite lors d'une première étape d'estimation<sup>9</sup> qui représente le vecteur de coefficients  $\hat{\beta}_j$  pour chaque titre.  $\eta_j$  est le terme d'erreur de la régression en coupe transversale qui incorpore une erreur de mesure<sup>10</sup> sur les variables endogène et exogène du modèle.  $\gamma_1$  représente la prime de risque de marché et  $\gamma_0$ , la constante de régression. Les résultats de Blume et Friend (1970) n'appuient pas le modèle de Sharpe-Lintner-Mossin à la faveur d'un  $\gamma_0$  statistiquement significatif, impliquant  $E(\tilde{R}_{zc\{m\}}) \neq R_f$ . Toutefois, la linéarité<sup>11</sup> et la prime de risque marché  $\gamma_1$  strictement positive ne sont pas rejetées.

Le test de Blume et Friend (1970) n'a pas été construit sur des estimations robustes. Les MCO utilisés ne sont pas efficaces en présence d'erreurs de mesure sur les variables. La condition d'orthogonalité entre régresseurs  $b_j$  et terme d'erreur et  $\eta_j$  est donc violée. Par conséquent, les coefficients  $\gamma_0$  et  $\gamma_1$  ont été respectivement surestimé et sous-estimé. De plus, l'erreur de mesure est accentuée sur la variable  $b_j$ , car dans son estimation lors de la première étape, la corrélation transversale entre les termes d'erreurs des  $N$  équations<sup>12</sup>, n'est pas prise en compte par les MCO. Les écart-types estimés des  $b_j$  sont biaisés. Black, Jensen et Scholes (1972) ont développé une stratégie de regroupement des données qui minimise ce problème.

---

<sup>9</sup>  $\hat{\beta}_j$  est estimé dans une première étape, à partir de  $N$  régressions en séries temporelles appliquées aux Lignes Caractéristique des titres sur la même période de  $T$  observations.

<sup>10</sup> Le terme d'erreur  $\eta_j$  de cette régression intègre aussi des erreurs de mesure  $v_j$  et  $\epsilon_j$  sur les espérances des variables qui sont approximées par leur moyenne empirique  $\bar{R}_j$  et  $b_j$  tel que  $\bar{R}_j = [E(\tilde{R}_j) - v_j]$  et  $b_j = [\beta_j - \epsilon_j]$ .

<sup>11</sup> La linéarité est testée à partir du modèle non restreint  $(\bar{R}_j - R_f) = \gamma_0 + \gamma_1 b_j + \gamma_2 b_j^2 + \eta_j$  où un test de Student est appliqué sur le coefficient  $\gamma_2$ .

<sup>12</sup> La plupart des actions échangées sur un marché ou membres d'un indice de référence partagent en général des facteurs de risques communs à leur secteur d'activité ou industrie. Cela peut être source de covariance dans les termes d'erreur de régressions multiples en coupe longitudinale. L'omission de cette structure de dépendance temporelle, est source de biais dans les écart-types estimés des coefficients  $\hat{\beta}_j$ .

Black, Jensen et Scholes (1972) ont adapté leur test aux séries temporelles stationnaires avec le modèle  $(\bar{R}_{j,t} - R_{f,t}) = \alpha_j + \beta_j(\bar{R}_{M,t} - R_{f,t}) + \varepsilon_{j,t}$ , avec  $j = [1,2,3, \dots, N]$  et  $t = [1,2,3, \dots, T]$  où  $(\bar{R}_{j,t} - R_{f,t})$  sont les rendements excédentaires moyens de chacune des  $N$  actions ordinaires régressés sur  $(\bar{R}_{M,t} - R_{f,t})$  qui est l'excès de rendement du portefeuille de marché. La période échantillonnale couvre de janvier 1926 à décembre 1965, à fréquence mensuelle. Avec les  $\beta_j$  estimés sur une période allant de janvier 1926 à décembre 1930, les titres sont triés par ordre décroissant du  $\hat{\beta}_j$  et répartis sur 10 portefeuilles affectés chacun d'un décile de l'échantillon. Les rendements excédentaires mensuels des portefeuilles sont ensuite calculés pour toute l'année 1931. Ces étapes d'estimation des  $\hat{\beta}_j$ , de regroupement en portefeuilles et de calculs des rendements moyens sur une année ultérieure, sont reprises année après année jusqu'à 1965. Finalement, avec les séries de rendements calculés de janvier 1931 à décembre 1965, la Ligne Caractéristiques des Titres de chaque portefeuille est estimée. L'hypothèse nulle est  $\alpha_j = 0, \forall j = [1,2,3, \dots, 10]$ . Celle-ci est rejetée, impliquant  $E(\tilde{R}_{zc\{m\}t}) \neq R_{f,t}$ . En effet, lorsque  $\hat{\beta}_j > 1, \hat{\alpha}_j < 0$  et vice-versa. Ainsi la version Sharpe-Lintner-Mossin est invalidée. Ces conclusions sont aussi appuyées par les résultats d'une régression en coupe transversale sur le même échantillon, qui utilise les  $\beta_j$  estimés, ce qui donne  $E(\tilde{R}_{zc\{m\}}) > R_f$ .

### 1.2.2 La critique de Roll (1977).

Roll (1977) a prouvé le caractère non testable des implications théoriques du MÉDAF. D'une part, il rappelle que le portefeuille de tangence n'est ni observable, ni approximable<sup>13</sup>. Son efficacité *ex post* est encore moins vérifiable. D'autre part, il montre que la validation apparente de l'hypothèse de linéarité par les tests des années 70, n'est qu'une conséquence mathématique plus que toute autre chose. En effet, si le portefeuille figurant dans la spécification du MÉDAF, est choisi sur la vraie frontière moyenne-variance efficiente (qui

<sup>13</sup> Aucun portefeuille ou indice de référence ne peut satisfaire son critère de composition ni même l'approximer. Il doit inclure tous les titres négociables tels les œuvres d'art, par exemple, les timbres postaux collectionnés et le capital humain.

peut être estimée), alors il existe déjà par construction, nonobstant toute considération d'ordre statistique, une relation linéaire entre le rendement espéré des titres et leur  $\beta_j$  estimé. Plus encore, il souligne que l'efficience *ex ante* d'un portefeuille de référence ne peut être testée si celui-ci est choisi selon des critères d'efficience *ex post* (sur la frontière efficiente). En somme, Roll (1977) montre une tautologie dans les résultats du MÉDAF.

### 1.2.3 Tests modernes du MÉDAF.

La critique de Roll (1977), loin de mettre au rancart le MÉDAF, a donné lieu à un débat autour des stratégies de sa validation empirique. Entre autres, Stambaugh (1982) et Shanken (1987) ont joué un rôle déterminant dans l'émergence d'une nouvelle génération de tests unifactoriels<sup>14</sup>. Le premier a montré que la précision des tests est peu sensible au choix du portefeuille représentatif, car la volatilité du prix des catégories d'actifs admissibles dans ce dernier est largement dominée par celle des actions. Shanken (1987) consolide cet argument en indiquant que l'erreur d'approximation du portefeuille de marché est une fonction décroissante du degré de corrélation entre celui-ci et le portefeuille figurant. Suivant une approche bayésienne qui suppose une forte corrélation, il teste le MÉDAF en estimant la probabilité du portefeuille de tangence d'être non efficient. Dans l'esprit de ces travaux, Gibbons, Ross et Shanken (1989) un cadre méthodologique de référence.

## 1.3 Test de Gibbons, Ross et Shanken (1989).

### 1.3.1 Test en grand échantillon et en échantillon fini.

Le test GRS comporte plusieurs innovations. Contrairement aux méthodes antérieures, il utilise l'estimateur des Régressions Apparemment Indépendantes<sup>15</sup> ou RAI, qui intègre la

---

<sup>14</sup> D'autres tests modernes basés sur une approche multifactorielle comme celle de Fama et French (1992) mais ne sont pas passés en revue ici. Ils ne sont pas pertinents au sujet développé.

<sup>15</sup> Cet estimateur sera exposé en détail dans les lignes qui suivent.

structure de corrélation transversale entre les équations individuelles et donne une estimation plus précise. Ensuite, il procède à un test joint de l'hypothèse nulle de l'approche de Black, Jensen et Scholes (1972). Finalement, en il fournit une interprétation économique intuitive et non tautologique. Avec des séries temporelles<sup>16</sup>, un système de  $N$  équations multiples est spécifié sous la forme vectorielle

$$R_t = \alpha + \beta R_{mt} + \varepsilon_t^{17},$$

avec les hypothèses  $E[\varepsilon_t] = 0$ ,  $E[\varepsilon_t \varepsilon_t'] = \Sigma$ ,  $E[R_{mt}] = \mu_m$ ,  $E[(R_{mt} - \mu_m)^2] = \sigma_m^2$  et  $Cov(R_{mt}, \varepsilon_t) = 0$ . La possibilité de corrélations contemporaines entre les  $N$  séries de termes d'erreurs n'est donc pas exclue.

L'estimateur RAI, introduit par Zellner (1962), fait office d'un estimateur des Moindres Carrés Généralisés pour un système d'équations multiples. Soit le modèle  $Y = X\theta + \varepsilon$ , où  $Y \in \mathbb{R}^{T \times N}$ ,  $X \in \mathbb{R}^{T \times K}$ ,  $\theta = [\alpha, \beta]'$  avec  $\theta \in \mathbb{R}^{2 \times N}$  et  $\varepsilon \in \mathbb{R}^{T \times N}$ . Le système peut être réarrangé comme suit

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_N \end{pmatrix}_{(T \times N) \times 1}, \quad X = \begin{pmatrix} X_1 & 0 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 & 0 \\ 0 & 0 & X_3 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \dots & X_K \end{pmatrix}_{(T \times N) \times (K \times N)}, \quad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \vdots \\ \theta_N \end{pmatrix}_{(K \times N)} \quad \text{et} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_N \end{pmatrix}_{(T \times N) \times 1}.$$

Pour cette spécification l'estimateur RAI est donné par

$$\hat{\theta}_{RAI} = [X'(\Omega^{-1})X]^{-1}X'(\Omega^{-1})Y,$$

où  $\Omega^{-1} = \Sigma^{-1} \otimes I_T$ <sup>18</sup> avec  $\Sigma = E[\varepsilon_t \varepsilon_t']$ . Lorsque  $\Sigma$  n'est pas connu, elle est estimée de façon convergente par  $\hat{\Sigma} = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{T}$  et  $\hat{\Omega}^{-1} = \hat{\Sigma}^{-1} \otimes I_T$ .

<sup>16</sup> La technique de regroupement des  $M$  actifs d'origine en portefeuilles  $N$  avec  $N < M$  est aussi appliquée selon l'approche de Black, Jensen et Scholes (1972) pour minimiser l'erreur de mesure sur les variables.

<sup>17</sup>  $t = [1, 2, 3, \dots, N]$  représente l'indice des actifs ou portefeuilles,  $R_t \in \mathbb{R}^{T \times 1}$ , le vecteur de séries de  $T$  observations des rendements excédentaires des  $N$  actifs,  $R_{mt}$  est la série temporelle de rendements excédentaires du portefeuille de référence,  $\varepsilon_t \in \mathbb{R}^{T \times 1}$ , le vecteur de séries de termes d'erreur des  $N$  équations et finalement  $\alpha$  et  $\beta$  sont respectivement des vecteurs de paramètres de dimension  $(N \times 1)$ .  $K$  désigne le nombre de régresseurs (incluant la constante).

<sup>18</sup>  $\otimes$  désigne le produit de Kronecker et  $I_T$  représente une matrice identité de rang  $(T \times N)$ .

Les paramètres estimés par RAI robuste sont asymptotiquement normaux selon

$$\hat{\alpha} \sim N\left(\alpha, \frac{1}{T} \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right] \Sigma\right) \text{ et } \hat{\beta} \sim N\left(\beta, \frac{1}{T} \left[1 + \frac{1}{\hat{\sigma}_m^2}\right] \Sigma\right).$$

S'il y a normalité dans la distribution jointe des termes d'erreur, les trois familles de test<sup>19</sup> paramétriques connues permettent de faire de l'inférence asymptotique, avec une loi de  $\chi^2$  à  $N$  degrés de liberté. Avec le test de Wald, on forme la statistique

$$J_0 = \hat{\alpha}' (\widehat{\text{Var}}[\hat{\alpha}])^{-1} \hat{\alpha} \quad ^{20}$$

qui pour des valeurs significativement non nulles, rejette  $H_0$  au seuil de significativité retenu. L'approche du ratio de vraisemblance procède par comparaison des fonctions de log-vraisemblance, évaluées pour le MÉDAF non contraint  $R_t = \alpha + \beta R_{mt} + \varepsilon_t$  et celui contraint  $R_t = \beta R_{mt} + \varepsilon_t$ . Si on désigne par  $\mathcal{L}^c(\alpha^*, \beta^*)$ , la fonction de log-vraisemblance du MÉDAF contraint et par  $\mathcal{L}^{nc}(\alpha^{nc}, \beta^{nc})$  celle du MÉDAF non contraint, sous  $H_0$  vrai, le ratio doit satisfaire la condition  $\lambda = \left[ \frac{\mathcal{L}^c(\alpha^*, \beta^*)}{\mathcal{L}^{nc}(\alpha^{nc}, \beta^{nc})} \right] \in (0, 1]$ . La statistique du test se forme suivant  $J_2 = -2 \ln \lambda$ . Finalement, l'approche du multiplicateur de Lagrange se base sur la vérification des conditions du premier ordre sur la fonction de log-vraisemblance associée aux paramètres du MÉDAF contraint (c'est-à-dire estimé sans le vecteur de constantes  $\alpha$ ). Si la contrainte est effectivement fondée, par les conditions du premier ordre, on doit nécessairement avoir  $S(\alpha^*, \beta^*) = \frac{\partial \mathcal{L}^c(\alpha^*, \beta^*)}{\partial \alpha^*} = 0$ , avec  $S(\alpha^*, \beta^*)$  représentant le vecteur du *Score*<sup>21</sup>. La statistique du test du Multiplicateur de Lagrange se construit selon  $LM = TR^2$  avec  $T$  la taille de l'échantillon et  $R^2$  le coefficient de détermination multiple du modèle.

<sup>19</sup> Les trois familles des tests statistiques sont celle Wald, celle du multiplicateur de Lagrange et celle du ratio de vraisemblance.

<sup>20</sup>  $\widehat{\text{Var}}[\hat{\alpha}] = \frac{1}{T} \left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right] \hat{\Sigma}$ , avec  $\hat{\mu}_m^2$  et  $\hat{\sigma}_m^2$  étant le carré de la moyenne et de la volatilité du rendement excédentaire du portefeuille de marché et  $\hat{\Sigma}$  la matrice variance-covariance des termes d'erreur de la régression RAI.

<sup>21</sup> Rappelons que le vecteur du *Score* est le vecteur de dérivées première de la fonction de log-vraisemblance par rapport à ses arguments. Pour un modèle  $Y = X\theta + \varepsilon$ , si on note la fonction de log-vraisemblance  $\ln \mathcal{L}(\theta|X, Y)$ , le vecteur du *Score* s'écrit  $\frac{\partial \ln \mathcal{L}(\theta|X, Y)}{\partial \theta}$  en notation vectorielle.

MacKinley (1985) et Campbell, Lo et MacKinley (1997) ont prouvé qu'en échantillon fini, les statistiques  $J_0$  et  $J_2$  sont biaisées. Des différences significatives entre les valeurs- $p$ <sup>22</sup> sont notées pour les trois familles de tests en petit échantillon. Des corrections ont été proposées par Gibbons, Ross et Shanken (1989) pour  $J_0$  et par Jobson et Korkie (1982) pour  $J_2$  selon

$$J_1 = \left( \frac{T-N-1}{TN} \right) J_0 \sim F_{N, T-N-1} \text{ et } J_3 = \left( \frac{T-\frac{N}{2}-2}{T} \right) J_2 \sim \chi_N^2,$$

en tenant compte de la taille du système d'équations  $N$  dont l'ampleur des biais est une fonction croissante, tel que discuté par Campbell, Lo et MacKinley (1997).

### 1.3.2 Interprétation économique du test.

La procédure de Gibbons, Ross et Shanken (1989) permet de tester intuitivement l'efficacité Moyenne-Variance du portefeuille de marché. En formalisant la statistique de Wald selon  $J_0 \equiv \mathbb{T} \left[ 1 + \frac{\hat{\beta}_m^2}{\hat{\sigma}_m^2} \right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$ , ces derniers démontrent pour tout portefeuille  $P$  approximant le vrai portefeuille de marché  $M$ , la forme quadratique  $\hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha}$  est représentative de la distance entre leur ratio de Sharpe respectif selon  $\hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \equiv \left[ \frac{\hat{\alpha}_P^2}{\hat{\sigma}_P^2} - \frac{\hat{\beta}_m^2}{\hat{\sigma}_m^2} \right]$ . Or, en vertu du MÉDAF on doit avoir  $J_0 \leq 0$  puisque  $\frac{\hat{\beta}_m^2}{\hat{\sigma}_m^2}$  est un maximum global. S'il y a donc évidence empirique de  $J_0 > 0$ , cela s'interprète comme la non efficacité du portefeuille de marché, et le rejet du MÉDAF. Toutefois, si la normalité n'est pas respectée, ce test perd sa puissance et ces interprétations ne sont plus justes.

<sup>22</sup> Valeur- $p$  est l'équivalent de l'expression anglaise *p-value* qui s'interprète comme la probabilité d'erreur de type-1 calculée pour une statistique (d'un test) à partir des données de l'échantillon.

<sup>23</sup> Ce résultat s'appuie sur le théorème de Muirhead (1983). Si  $T$  est le nombre d'observations et  $N$  le nombre de variables aléatoires, pour un vecteur de variables  $Z \in \mathbb{R}^{N \times 1}$  avec  $Z \sim N(0, \Omega)$  et une matrice  $\hat{\Omega} \in \mathbb{R}^{N \times N}$  distribuée selon une loi de Wishart (une généralisation de la loi  $\chi^2$  multivariée) selon  $T\hat{\Omega} \sim W_1(T, \Omega)$  avec  $T \geq N$ , si  $Z$  et  $\Sigma$  sont indépendants alors on a  $\left( \frac{T-N-1}{TN} \right) Z' \hat{\Omega}^{-1} Z \sim F_{N, T-N-1}$ . Avec  $Z = \left[ 1 + \frac{\hat{\beta}_m^2}{\hat{\sigma}_m^2} \right]^{\frac{1}{2}} \hat{\alpha}$  et  $\Omega = T\hat{\Sigma}$  on obtient la statistique  $J_1$ .

## CHAPITRE II

### GÉNÉRALISATION DU MÉDAF AUX LOIS À SYMÉTRIE ELLIPTIQUE

#### 2.1 Non normalité des rendements et implications pour le MÉDAF.

##### 2.1.1 Évidence empirique de la non normalité et alternatives.

L'hypothèse de normalité est incompatible avec de nombreux faits stylisés de la finance de marchés. Des amas de volatilité<sup>24</sup>, une corrélation transversale non constante entre les rendements et un excès de kurtosis sont fréquemment observables dans les titres boursiers à différentes périodicités. Mis en lumière par Mandelbrot (1963) et Fama (1965), l'excès de kurtosis est le fait qui a retenu le plus d'attention. Depuis ces découvertes, plusieurs alternatives ont été proposées, comme la distribution log-normale avec Clark (1973), la loi exponentielle selon Gray et French (1990) et lois  $t$  de Student promues par Blattberg et Gonedes (1974), en raison de leur nature leptokurtique.

L'idée d'utiliser des familles de lois pour caractériser les rendements a aussi fait son chemin dans la littérature. Mandelbrot (1963) et Fama (1965) avaient défendu l'adéquation de la famille de lois Pareto-stables<sup>25</sup> non gaussiennes à variance non bornée avec les queues

---

<sup>24</sup> Les amas de volatilité se traduisent par le fait que les périodes de marché haussier ou baissiers ont tendance à se succéder. En d'autres termes, il s'agit d'une autocorrélation dans la variance conditionnelle des rendements.

<sup>25</sup> Une variable aléatoire  $X$  suit une distribution stable s'il existe une séquence de  $n$  variables  $\{Y_i\}$  *i.i.d.* et de constantes  $\{a_n\} \in \mathbb{R}$  et  $\{b_n\} \in \mathbb{R}^+$  tel que  $\lim_{n \rightarrow \infty} \left\{ \frac{1}{b_n} \sum_{i=1}^n Y_i - a_n \right\} \rightarrow^d X$ . On écrit alors  $X \sim S_\alpha(\beta, \sigma, \mu)$  où

épaisses de la distribution des rendements. Cet argument sera repris plus tard, notamment par Byczkowski, Nolan et Rajput (1993), Abdoul-Hamid et Nolan (1998) et Rachev et Mittnik (2000) qui ont établi la preuve analytique de leur cohérence avec Théorie Moyenne-Variance. La possibilité de variance non bornée nuit toutefois à la mesure cohérente du risque, selon Fama (1965). Chamberlain (1983), Owen et Rabinovich (1983), Beck (1987) et Ingersoll (1987) ont proposé les lois à symétrie elliptique, qui incluent les combinaisons d'échelle de distributions normales symétriques discutées par Engle (1982) et Bollerslev (1986) pour caractériser le phénomène d'amas de volatilité et l'excès de kurtosis. Toutefois, la symétrie elliptique ne prévaut pas toujours. Aussi, Clark (1973), Fielitz et Rozelle (1983) ont suggéré une généralisation avec la famille des combinaisons d'échelles de distributions normales<sup>26</sup>, qui sont plus largement, membres de la classe elliptique asymétrique. Chu (1973), Eaton (1981) et Gupta et Varga (1993) ont démontré leur relation avec la famille à symétrie elliptique. Finalement, la famille plus englobante de lois dites hyperboliques<sup>27</sup> a été introduite dans la littérature financière par Eberlein et Keller (1995). Plusieurs travaux empiriques, dont ceux de Prause (1997, 2000) et Eberlein (2001), Bingham et Kiesel (2001) et Eberlein et Hammerstein (2004) ont prouvé son adéquation empirique, notamment avec l'excès de kurtosis et l'asymétrie.

### 2.1.2 Conséquences sur le MÉDAF.

Du point de vue théorique, l'hypothèse de normalité ne constitue pas une condition *sine qua non* à la dérivation du MÉDAF<sup>28</sup>. Entre autres, Ross (1978), Ingersoll (1987) et Berk (1997)

---

$\beta, \sigma, \mu$  sont respectivement les paramètres d'asymétrie, d'échelle ou encore dispersion et localisation. Le paramètre  $\alpha \in (0, 2]$  mesure le degré d'aplatissement de la distribution.

<sup>26</sup> Une combinaison d'échelle de distributions normales est une distribution sous laquelle la réalisation d'une variable aléatoire provient d'une combinaison linéaire de deux ou plusieurs lois normales sous-jacentes de paramètres  $\mu_i$  et  $\sigma_i$  respectifs. Sans perte de généralité, si  $X_1 \sim N(\mu_1, \sigma_1)$  et  $X_2 \sim N(\mu_2, \sigma_2)$  et  $Z = (\pi_1 X_1 + \pi_2 X_2)$  avec  $\mu_1 \neq \mu_2$  et/ou  $\sigma_1 \neq \sigma_2$ ,  $\pi_1 + \pi_2 = 1$  et  $0 \leq \pi_i \leq 1 \forall i$  alors  $Z$  suit une combinaison d'échelle de distributions normales. Lorsque  $\mu_1 \neq \mu_2$  et  $\sigma_1 = \sigma_2$  ou  $\sigma_1 \neq \sigma_2$  la distribution de  $Z$  est asymétrique, lorsque  $\mu_1 = \mu_2$  et/ou  $\sigma_1 \neq \sigma_2$   $Z$  sera symétrique et elliptique.

<sup>27</sup> Elle regroupe la distribution inverse gaussienne et celle hyperbolique.

<sup>28</sup> Lorsqu'on omet les préférences de l'investisseur représentées par une fonction d'utilité quadratique.

ont établi des conditions moins restrictives sur les propriétés stochastiques des données. Selon Ross (1978), il suffit que tout actif soit décomposable en une combinaison linéaire de deux ou plusieurs fonds auxquels s'ajoute un bruit blanc. Owen et Rabinovitch (1983) ont montré que les distributions à symétrie elliptique satisfont cette propriété de séparabilité de Ross (1978). Ingersoll (1987) et Berk (1997) ont ainsi établi que la symétrie elliptique, constituée en soi, est une condition suffisante au MÉDAF. Ingersoll (1987) a illustré ceci en considérant le cas trivial d'une distribution  $t$  de Student multivariée pour les rendements, sans émettre d'hypothèses supplémentaires sur les préférences des agents.

Pour le test de Gibbons, Ross et Shaken (1989) traditionnel, le non-respect de la normalité peut entraîner une plus ou moins grande perte d'efficacité et de puissance. Avec un estimateur robuste à la non-normalité, celui de la Méthode des Moments Généralisés, MacKinley et Richardson (1991), ont illustré l'ampleur des biais dans les statistiques construites avec des données simulées sous diverses distributions alternatives. Les valeurs- $p$  de celles-ci sont significativement différentes de celles issues de données normalement distribuées. De surcroît, dans certains cas, les biais persistent durablement malgré l'ajout d'un nombre considérable d'observations et peuvent être positifs ou négatifs selon l'erreur de spécification, impliquant soit un risque d'erreur de type-1 ou d'erreur de type-2 accru. Entre autres, Zhou (1993) et Sentana (2008) ont tiré les mêmes conclusions en reprenant l'analyse dans le contexte de lois elliptiques symétriques.

## 2.2 Généralisation du MÉDAF avec la symétrie elliptique.

### 2.2.1 Définition des lois à symétrie elliptique.

Les distributions à symétrie elliptique<sup>29</sup> sont une famille appartenant à la classe de lois dites elliptiques asymétriques. Un vecteur de variables aléatoires  $X = [X_1, X_2, X_3, \dots, X_N]$  est dit distribué à symétrie elliptique<sup>30</sup> si sa fonction de densité, lorsqu'elle existe, s'écrit de la forme

---

<sup>29</sup> Elles ont été mises en évidence par Kelker (1970).

$$f(X) = c_N |\Sigma|^{-1/2} g_N \{(X - \mu)' \Sigma^{-1} (X - \mu)\}.$$

L'expression  $g_N$  désigne une fonction<sup>31</sup> génératrice de densité, dont la forme fonctionnelle est indépendante de  $X$ . Le terme  $\mu = [\mu_1, \mu_2, \mu_3, \dots, \mu_N]'$  représente le vecteur de moyennes lorsqu'il existe et  $\Sigma \in \mathbb{R}^{n \times n}$  est une matrice symétrique définie positive qui représente la matrice variance-covariance, si elle est bornée. La constante  $c_N$  est déterminée par la forme fonctionnelle  $g_N$  et finalement  $|\Sigma|$  désigne le déterminant de la matrice  $\Sigma$ .

L'intuition géométrique derrière la symétrie elliptique peut être illustrée en considérant l'exemple d'une distribution normale bivariée d'un vecteur de variables  $X = [X_1, X_2]$  de paramètres  $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  et  $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$ , représentée sur la figure 2.1, à la section supérieure. La fonction de densité jointe de  $X_1$  et  $X_2$  décrit une surface à l'allure d'une cloche.

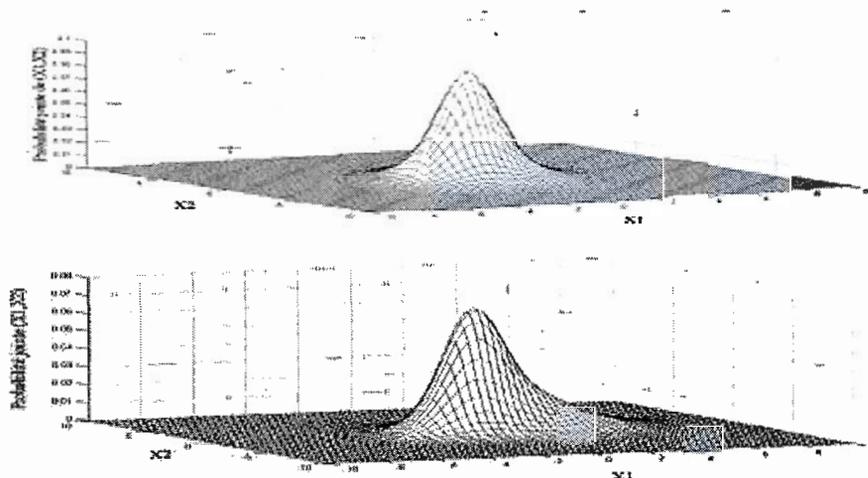


Figure 2.1 Distributions elliptiques symétriques et asymétriques.

L'intersection de cette surface avec les plans orthogonaux à sa hauteur décrit un ensemble de courbes de niveau, de forme elliptique qui sont toutes centrées sur les coordonnées du vecteur

<sup>30</sup> On écrit  $X \sim E_N(\mu, \Sigma, g_N)$  où  $\mu$  et  $\Sigma$  sont respectivement le vecteur de moyenne et la matrice variance-covariance lorsqu'ils sont bornés et  $g_n$  la forme fonctionnelle génératrice de la densité de probabilité.

<sup>31</sup> On note  $g_N$  avec  $N$  en indice parce qu'il s'agit d'une fonction de plusieurs variables. Plus loin on verra que la fonction  $g_N$  peut correspondre à une fonction scalaire (ou à une dimension).

de moyennes  $\mu$ . Les axes de symétrie de ces ellipses sont exclusivement déterminés par la forme quadratique  $\{(X - \mu)' \Sigma^{-1} (X - \mu)\}$ . La forme fonctionnelle  $g_N$  renseigne sur le degré d'épaisseur des queues de la distribution<sup>32</sup>. La section inférieure de la figure 2.1 illustre une combinaison d'échelle<sup>33</sup> (bivariée) de lois normales de moyennes et variance inégales, qui est un exemple de loi elliptique asymétrique.

### 2.2.2 Compatibilité avec le MÉDAF.

Les propriétés distinctives<sup>34</sup> de la famille à symétrie elliptique ont des implications théoriques et empiriques pour le MÉDAF. D'une part, Bawa (1975) avait élargi la portée de l'analyse moyenne-variance en termes de paramètres de localisation et de dispersion  $\mu$  et  $\Sigma$ , lesquels caractérisent les lois à symétrie elliptique. D'autre part, Ross (1978) avait généralisé les conditions suffisantes au MÉDAF avec des propriétés stochastiques des titres, qui ne requièrent pas d'analyse en termes moyenne-variance, ni une spécification particulière de la fonction d'utilité de l'agent. Il a montré que si chaque titre du marché  $Z_t$  peut être stochastiquement décomposé en une somme selon  $Z_t = \theta_0 + \theta_1 Z_t^* + v_t$ , avec  $Z_t^*$  faiblement préféré à  $Z_t$  et  $v_t$  un bruit blanc d'espérance nulle, les choix des investisseurs restent compatibles avec le MÉDAF. Owen et Rabinovitch (1983) ainsi que Chamberlain (1983) ont démontré que les lois à symétrie elliptique sont cohérentes avec le critère de Ross (1978) en vertu de leur divisibilité infinie et leur stabilité par sommation.

---

<sup>32</sup> Pour une distribution normale  $g_N(z) = \exp\{-\frac{1}{2}z^2\}$  avec  $z = \Sigma^{-1/2}(X - \mu)$ . Pour la distribution  $t$  de Student,  $g_N(z) = \left\{1 + \frac{z^2}{\nu}\right\}^{-\frac{(N+\nu)}{2}}$  où  $\nu \in \mathbb{N}^+$  est le nombre de degrés de liberté. Pour une distribution exponentielle,  $g_N(z) = \exp\{-r \cdot z^s\}$  avec  $r$  et  $s$  des nombres positifs.

<sup>33</sup> Sur la partie inférieure est illustrée une combinaison d'échelle de lois normales selon  $X = \sum_{i=1}^2 0.5 X_i$  avec  $X_1 \sim N(\mu_1, \Sigma_1)$  où  $\mu_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  et  $\Sigma_1 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$  et  $X_2 \sim N(\mu_2, \Sigma_2)$  où  $\mu_2 = \begin{pmatrix} 2 \\ 0.5 \end{pmatrix}$  et  $\Sigma_2 = \begin{pmatrix} 4.5 & 0.5 \\ 0.5 & 3.5 \end{pmatrix}$ . On note un étalement de la distribution jointe vers la droite. Ces graphiques ont été réalisés par l'auteur à partir de Matlab R2014b.

<sup>34</sup> Pour un exposé détaillé sur les propriétés générales de la famille, le lecteur peut se référer à Fan, Kotz et Ng (1990).

S'appuyant sur ces arguments, les auteurs généralisent le modèle où les investisseurs ne s'accordent plus sur une même et unique loi pour former leurs anticipations, mais ont connaissance commune de leurs paramètres de localisation et de dispersion. Pour la plupart des distributions usitées en théorie financière, cela ramène à l'analyse moyenne-variance, mais les résultats tiennent aussi pour les lois à variance non bornées de Mittnik (1998) et Rachev (2000). Du point de vue de l'analyse économétrique, la stabilité par sommation<sup>35</sup>, par convolution et par transformation affine des rendements distribués selon cette famille en garanti le caractère infiniment divisible. Ainsi, pour tout titre, si les rendements journaliers sont à symétrie elliptiques, ceux hebdomadaires, mensuels ou annuels et intra-journaliers le sont systématiquement. Leur regroupement en portefeuille n'affecte pas leur distribution. Owen et Rabinovitch (1983) ont montré en ce sens que lorsque les variables d'une régression sont ainsi caractérisées, les termes d'erreur issus des MCO ou Moindres Carrés Pondérés et les résidus standardisés le sont aussi, car leur fonction de densité jointe est invariante par rapport à toute transformation orthogonale de celles-ci. Le test de symétrie elliptique de Li, Fang et Zhu (1997), exposé et appliqué plus loin, repose sur ce principe. Finalement, dans ce même ordre d'idées, la fonction génératrice de densité  $g_N$  peut toujours correspondre à une fonction univariée ou scalaire notée  $g$ , en appliquant une transformation du vecteur de variables en une variable univariée.

## 2.3 Test GRS sous l'hypothèse de symétrie elliptique.

### 2.3.1 L'approche paramétrique.

Le test GRS, lorsqu'il est construit sur une estimation par Maximum de Vraisemblance, implique la spécification de la distribution jointe des termes d'erreurs. En principe, l'application des MCO sous-entend une inférence statistique sous l'hypothèse de normalité jointe, ce qui est trop restrictif. L'approche paramétrique du test GRS généralisée aux lois

---

<sup>35</sup> Toute variable  $Y$  elliptiquement distribuée peut en effet être décomposée en une somme de variables sous-jacentes  $X_i$  selon  $Y \equiv^d \sum_{i=1}^n X_i / b_i$  où les termes  $b_i$  représentent des constantes. L'expression  $\equiv^d$  désigne l'identité en distribution de probabilité.

elliptiques symétriques consiste donc à sélectionner dans cette famille, une distribution alternative plus compatible avec la réalité. Dans la littérature, la loi  $t$  de Student multivariée s'est imposée comme un choix parcimonieux pour spécifier le Maximum de Vraisemblance. En principe, dans un modèle à  $N$  portefeuilles distribués selon  $E_N(\mu, \Sigma, g_N)$ , la forme fonctionnelle  $g_N$  est indicatrice de l'épaisseur des queues la distribution, ce qui est aisément modélisable avec le nombre de degrés de liberté d'une loi  $t$  de Student ou bien la mesure d'excès de kurtosis d'une loi de Kotz.

Zhou (1993) et Sentana (2008) ont cependant souligné les faiblesses de l'approche paramétrique. L'erreur de spécification sur la distribution, même si la vraie famille est correctement identifiée, induit toujours des biais dans les statistiques du test GRS. Si la symétrie elliptique est dans les termes d'erreurs, une augmentation suffisante de la taille de l'échantillon peut garantir l'optimalité asymptotique des coefficients estimés par un Pseudo-Maximum de Vraisemblance<sup>36</sup> basé sur une distribution  $t$  de Student. En revanche, en cas de symétrie elliptique dans les régresseurs, l'efficacité du Pseudo-Maximum de Vraisemblance est remise en question. Zhou (1993) illustre cela en rapportant des différences significatives dans les statistiques du test *GRS* à partir de simulation de Monte-Carlo<sup>37</sup> sous des distributions comme la  $t$  de *Student*, celle de Kotz et des combinaisons d'échelle de distributions normales symétriques. De plus, selon l'erreur de spécification, le risque d'erreur de type-1 ou 2 peut être considérablement accru et durable.

### 2.3.2 L'approche semi-paramétrique.

Face à la fragilité de l'approche paramétrique, la littérature s'est tournée vers des techniques plus flexibles. MacKinley et Richardson (1991) ont proposé un test robuste à la non normalité basé sur l'estimateur de la Méthode de Moments Généralisés ou GMM. Ce dernier n'impose

---

<sup>36</sup> Le Pseudo-Maximum de Vraisemblance est un Maximum de Vraisemblance comportant une erreur de spécification sur la fonction de densité mais dont la matrice variance-covariance est ajustée pour minimiser asymptotiquement la perte d'efficacité résultant de cette erreur. Bollerslev et Wooldridge (1992) ont montré que sous certaines conditions cet estimateur est asymptotiquement optimal.

<sup>37</sup> Ces probabilités critiques sont calculées en utilisant des simulations de Monte Carlo.

aucune restriction sur la forme de la distribution des termes d'erreur de la régression. Cependant, Sentana et Amengual (2008) ont montré que le test basé sur le GMM n'est pas plus puissant que celui reposant le Pseudo-Maximum de Vraisemblance basée sur une loi  $t$  de Student, si la symétrie elliptique prévaut empiriquement. Hodgson, Linton et Vorkink (2002) ont en ce sens, développé une méthodologie de test semi-paramétrique efficace en présence de lois à symétrie elliptique. Ils ont estimé des coefficients du MÉDAF avec des écart-types en moyenne 15% plus faibles que ceux des MCO. Cependant, lorsque la symétrie elliptique n'est pas respectée, la procédure de Hodgson et *alii* (2002) risque de perdre son efficacité asymptotique, comme l'indique Sentana (2008). À ce moment le GMM s'impose comme une solution plus parcimonieuse.

### 2.3.3 Correction de Zhou (1993) sur la statistique de Wald.

Zhou (1993) a proposé un ajustement de la statistique de Wald  $J_0$  du test GRS pour le généraliser. Celle-ci devient  $J_0^* = \left(\frac{J_0}{\xi}\right) \sim \chi_N^2$  avec  $\xi = \left(1 + \frac{\kappa_Y}{N(N+2)}\right)$ , où  $\kappa_Y$  désigne la mesure multivariée du kurtosis de la distribution jointe du vecteur de  $N$  rendements excédentaires (variables endogènes du système d'équations multiples à estimer). Un estimateur convergent du coefficient d'aplatissement multivarié est proposé par Mardia (1970). Lorsque la vraie distribution jointe des données est une normale,  $\kappa_Y = 0$  et  $J_0^* = J_0$ . Sinon  $J_0^* < J_0$ . L'approche de Zhou (1993), est conçue pour rendre la statistique de Wald robuste à la non normalité au sens large, mais non spécifiquement à la présence à symétrie elliptique. Aussi, Hodgson et *alii* (2002) soulignent que leur approche dérivée explicitement sur cette hypothèse doit conduire à un test plus puissant.

### 2.3.4 L'approche de Beaulieu, Dufour et Khalaf (2002).

Les limitations des approches discutées ci-haut ont conduit certains chercheurs à prôner l'abandon de l'objectif d'estimation robuste pour la construction des tests. Beaulieu, Dufour et Khalaf (2002) ont conçu une stratégie visant à estimer les valeurs-p exactes de la

statistique GRS construite par MCO robuste avec pour seule condition imposée, l'hypothèse *i.i.d.* sur les termes d'erreur. La distribution en échantillon finie, sous l'hypothèse nulle, des statistiques de Wald, du Multiplicateur de Lagrange et du Ratio de Vraisemblance est reconstituée par échantillonnage répété avec des simulations de Monte Carlo sur des résidus standardisés *i.i.d.* issus de lois appartenant à la famille à symétrie elliptique.

Sentana (2008) souligne toutefois que l'hypothèse *i.i.d.* sur la distribution conditionnelle des résidus est trop restrictive pour la famille à symétrie elliptique, ce qui fait que l'approche de Beaulieu, Dufour et Khalaf (2002) n'est pas non plus à l'abri du risque de perte de puissance en cas d'hypothèse erronée sur le processus générateur des données. Sentana (2008) propose en ce sens de recourir à des techniques d'échantillonnages répétés n'imposant aucune forme d'hypothèse sur les données comme celle du *bootstrap*.

## CHAPITRE III

### CADRE MÉTHODOLOGIQUE : LA PROCÉDURE H.L.V. (2002)

#### 3.1 Fondements statistiques de la procédure H.L.V. (2002).

##### 3.1.1 Principes de l'estimation adaptative.

Avant de présenter la procédure de Hodgson, Linton et Vorkink (2002), les principes fondamentaux de l'estimation semi-paramétrique efficace et leur application au MÉDAF à symétrie elliptique seront abordés. En principe, la connaissance exacte de la vraie distribution jointe des données n'est pas toujours requise pour obtenir une estimation paramétrique sans biais, convergente et efficace. En revanche, elle joue un rôle déterminant dans l'efficacité de l'inférence statistique. En grand échantillon et en échantillon fini, la variance asymptotique de l'estimateur du Maximum de Vraisemblance, en vertu du Théorème de Cramer-Rao, est la borne inférieure de la variance de tout estimateur régulier, convergent et asymptotiquement normal. Dans le modèle économétrique

$$Y = X\theta + \varepsilon^{38},$$

un estimateur paramétrique va imposer une hypothèse implicite ou explicite<sup>39</sup> sur la vraie distribution jointe  $f^*(\varepsilon)$ . Sans perte de généralité, si on suppose  $\varepsilon \sim E_N(\mu, \Sigma, g_N)$ , la fonction

---

<sup>38</sup>  $Y \in \mathbb{R}^{T \times N}$  et  $X \in \mathbb{R}^{T \times K}$  sont respectivement, le vecteur de  $N$  variables endogènes et un vecteur de  $K$  variables exogènes (observées sur  $T$  périodes).  $\varepsilon \in \mathbb{R}^{T \times N}$  est un vecteur de  $N$  termes d'erreur et  $\theta \in \mathbb{R}^{K \times N}$  est un vecteur de  $N$  paramètres associés aux  $K$  régresseurs du modèle.

<sup>39</sup> Par exemple, dans un cas univarié, un estimateur du Maximum de Vraisemblance basé sur une distribution  $t$  de Student implique l'identification de la moyenne, de la variance ainsi que du nombre de degrés de libertés.

de densité jointe  $f^*(\varepsilon) = f^*(\mu, \Sigma, g_N^*)$  est partiellement caractérisée. Les paramètres<sup>40</sup>  $\mu \in \mathbb{R}^{N \times 1}$  et  $\Sigma \in \mathbb{R}^{N \times N}$  sont identifiables s'ils sont bornés. Cependant un autre ensemble de paramètres  $\eta^* \in \mathbb{R}^{N \times 1}$  associés à la forme fonctionnelle<sup>41</sup>  $g_N^*$  ne le sont pas. Aussi, sont-ils désignés par paramètres de nuisance. Dans ce contexte, on est en présence d'un modèle semi-paramétrique. Pour chaque valeur de  $\eta$  envisageable sous les hypothèses du modèle, correspond une spécification distincte de la fonction de log-vraisemblance des paramètres  $\mathcal{L}(\theta|\varepsilon)$  selon  $\{\mathcal{L}(\theta|\varepsilon)\}_j = \ln \prod_{t=1}^T f(\varepsilon_t|\mu, \Sigma, \eta_j)$  avec  $j = [1, 2, 3, \dots, J]$ .

Un estimateur paramétrique donne pour solution optimale, un vecteur de paramètres  $\hat{\theta}^*$  correspondant au vecteur  $\eta_j^*$  qui satisfait les conditions<sup>42</sup> d'optimalité sur la fonction de log-vraisemblance  $\{\mathcal{L}(\theta|\varepsilon)\}$ . Un estimateur semi-paramétrique décrit comme solution une séquence de vecteurs de paramètres  $\tilde{\theta}_j$  avec  $j = [1, 2, 3, \dots, J]$  admissibles ou estimables pour le modèle, pour chaque  $\eta_j$  sur la famille. Dans cette séquence, la solution globalement optimale est le  $\tilde{\theta}_j^*$  qui correspond à la distribution exacte de l'échantillon.

Un choix arbitraire de  $\eta^*$  selon une approche paramétrique, expose à un risque d'erreur de spécification et conséquemment à la non convergence et la non optimalité asymptotique, ont montré White (1982), Bollerslev et Wooldridge (1992) et Zhou (1993). Ils ont, à cet effet, proposé une estimation robuste par la méthode du Pseudo-Maximum de Vraisemblance<sup>43</sup>, qui est convergente et asymptotiquement optimale sous certaines conditions. Sentana (2008) a toutefois soulevé les limitations de cette stratégie et indiqué les enjeux liés à son emploi à l'aveuglette. Dans certains cas, les biais induits dans les statistiques des tests peuvent persister durablement, en dépit de l'augmentation de la taille de l'échantillon. Le Pseudo-Maximum de Vraisemblance ne peut donc s'imposer comme une solution universelle.

<sup>40</sup> Ils représentent respectivement des vecteurs de moyenne et la matrice variance-covariance.

<sup>41</sup> Un exemple de paramètre de nuisance  $\eta^* \in \mathbb{R}^{N \times 1}$ , est le vecteur du nombre de degrés de liberté d'une distribution multivariée  $t$  de Student à  $N$  variables.

<sup>42</sup> Les conditions du premier ordre sur la fonction de log-vraisemblance sont  $E \left[ \frac{\partial \{\mathcal{L}(\theta|\varepsilon)\}}{\partial \theta} \right] = 0$  et les conditions du second ordre sont  $-E \left[ \frac{\partial^2 \{\mathcal{L}(\theta|\varepsilon)\}}{\partial \theta \partial \theta'} \right] = I(\theta)$  où  $I(\theta)$  désigne la matrice d'information de Fisher.

<sup>43</sup> Un estimateur du Pseudo-Maximum de Vraisemblance gaussien opère un ajustement de la matrice variance-covariance asymptotique pour tenir compte du biais de spécification de la fonction log-vraisemblance.

La vertu de l'approche semi-paramétrique est d'imposer le moins de restrictions possibles sur la distribution jointe  $f^*(\varepsilon)$  en vue d'approximer le  $\tilde{\theta}_j$  (dans l'ensemble des solutions décrites par le modèle) qui satisfait l'optimalité asymptotique. D'un point de vue conceptuel, lorsqu'il est convergent, asymptotiquement normal et à la limite aussi efficace<sup>44</sup> que le véritable estimateur du Maximum de Vraisemblance, un estimateur semi-paramétrique est dit adaptatif. Dans le cadre des distributions à symétrie elliptique, l'adaptabilité de l'estimateur se traduit par sa capacité à accomplir cette optimalité asymptotique, indépendamment de la connaissance exacte de  $g_N$  ou du vecteur de paramètres de nuisance  $\eta^*$ . Les conditions nécessaires et suffisantes à cette propriété sont extensivement discutées par Stein (1956), Hájek (1962), Bérán (1974), Bickel (1982) et Pagan et Ullah (1999).

La portée de l'adaptabilité se limite exclusivement une classe d'estimateurs dit réguliers<sup>45</sup>, qui convergent asymptotiquement à la vitesse  $\sqrt{T}$ , où  $T$  représente la taille de l'échantillon. En effet, certains modèles semi-paramétriques peuvent admettre dans leur ensemble de solutions  $\tilde{\theta}_j$ , quelques estimateurs super-convergers<sup>46</sup>. Ces derniers sont exclus de la définition de l'adaptabilité. Selon Stein (1956), un estimateur semi-paramétrique est régulier si et seulement si, localement sa distribution limite est normale selon

$$\sqrt{T} (\tilde{\theta}_j - \theta_j) \xrightarrow{d} N(0, I(\theta_j)^{-1})^{47}, \forall j = [1, 2, 3, \dots, J].$$

Cette définition a des implications directes sur les conditions du premier ordre et du second ordre de la fonction log-vraisemblance de tout modèle semi-paramétrique efficace.

---

<sup>44</sup> Rappelons qu'un estimateur du Maximum de Vraisemblance est convergent si la dérivée première de la fonction de vraisemblance appelée *Score*, s'annule à la vraie valeur des paramètres de la distribution à priori. Il est asymptotiquement optimal ou efficace si sa variance atteint la borne de Cramer-Rao.

<sup>45</sup> Les conditions de régularité sont plus formellement étudiées par Le Cam (1969) et discutées dans Bickel (1982). Entre autres, il est nécessaire que la fonction de log-vraisemblance  $\mathcal{L}(\cdot)$  soit différentiable par rapport à tout paramètre  $\theta_j$ , que la matrice d'information  $I(\theta)$  existe et qu'il existe parmi la séquence de paramètres décrits par le modèle semi-paramétrique, un sous ensemble qui convergent à la vitesse  $\sqrt{T}$ . Cette dernière condition réfère à la Normalité Asymptotique Locale.

<sup>46</sup> Un estimateur superconvergent est celui dont la variance asymptotique est localement inférieure à la borne de Cramer-Rao mais qui n'est pas efficace en échantillon fini. De plus, sa vitesse de convergence est plus élevée que celle de Maximum de Vraisemblance.

<sup>47</sup> Ici  $I^{-1} = \{I(\theta_j, \eta_j)\}^{-1}$  est la matrice variance-covariance asymptotique évaluée à la vraie valeur de l'ensemble des paramètres  $\theta_j$  et  $\eta_j$ .

Si on désigne par  $\Delta(\theta_j) = \frac{\partial}{\partial \theta_j} \{\mathcal{L}(\theta|\varepsilon)\}_j$  et  $\Delta(\eta_j) = \frac{\partial}{\partial \eta_j} \{\mathcal{L}(\theta|\varepsilon)\}_j$  les ensembles de vecteurs de *Score* associés respectivement à chaque vecteur de paramètres  $\theta_j \in \mathbb{R}^K$  identifiables et chaque vecteur de paramètres de nuisance  $\eta_j \in \mathbb{R}^K$ , la normalité asymptotique locale de l'estimateur implique l'orthogonalité entre  $\Delta(\theta_j)$  et  $\Delta(\eta_j)$ . Cela sous-entend que la convergence des vecteurs de *Score*  $\Delta(\theta_j)$  doit être réalisable, indépendamment de la convergence de  $\Delta(\eta_j)$ . Bickel (1982) montre que cette condition implique que la matrice variance-covariance asymptotique de l'estimateur adaptatif soit bloc diagonale selon

$$\{I(\theta_j, \eta_j)\}^{-1} = \begin{pmatrix} -E[\Delta(\theta_j)\Delta(\theta_j)']^{-1} & -E[\Delta(\theta_j)\Delta(\eta_j)'] \\ -E[\Delta(\eta_j)\Delta(\theta_j)'] & -E[\Delta(\eta_j)\Delta(\eta_j)']^{-1} \end{pmatrix}, \forall j = [1, 2, 3, \dots, J],$$

avec  $E[\Delta(\theta_j)\Delta(\eta_j)'] = 0$  et  $E[\Delta(\eta_j)\Delta(\theta_j)'] = 0$  et  $-E[\Delta(\theta_j)\Delta(\theta_j)']^{-1}$  correspondant à la borne de Cramer-Rao pour chaque  $\theta_j$  considéré, en vertu du théorème de Hájek (1972).

Il est toutefois impossible de valider empiriquement les conditions de Stein (1956) et de Bickel (1982) telles que stipulées plus haut, sans connaître l'ensemble des  $\eta_j$  sous la famille de lois de probabilité. Bickel (1982) et Hodgson et alii (2002) ont toutefois explicité les critères d'adaptabilité en ces termes plus concrets. Sur une famille de lois considérée, vérifier l'adaptabilité d'un estimateur semi-paramétrique revient à vérifier la convergence en probabilité de son vecteur du *Score* et de sa matrice variance-covariance asymptotique évalués pour chaque  $\theta_j$  donné selon

$$\hat{\Delta}(\theta_j) - \Delta(\theta_j) \rightarrow^P 0 \forall j \text{ et } \hat{I}(\theta_j)^{-1} - I(\theta_j)^{-1} \rightarrow^P 0, \forall j.$$

Dans cet ordre d'idées, tout estimateur  $\tilde{\theta}$  régulier quelconque, dans l'ensemble de solutions, qui satisfait cette convergence est admissible pour le modèle semi-paramétrique. Intuitivement, il est concevable d'estimer de façon non paramétrique la densité jointe tel que  $\hat{f}(\varepsilon) \rightarrow^P f^*(\varepsilon)$  et de calculer par la suite la fonction log-vraisemblance,  $\hat{\Delta}(\theta_j)$  et  $\hat{I}(\theta_j)^{-1}$ . Cela estimerait de façon convergente, le  $\theta_j$  correspondant à la vraie distribution. Suivant cette logique, Rothenberg et Leenders (1968) ont proposé un estimateur adaptatif en deux étapes, où, premièrement, on estime le vecteur de termes d'erreur  $\varepsilon$  par une régression paramétrique

robuste et ensuite, par la mise en œuvre de techniques non paramétriques, on estime  $f^*(\varepsilon)$ , le Score  $\Delta(\theta_j)$  et de la matrice variance-covariance asymptotique  $I(\theta_j)^{-1}$ . Hodgson, Linton et Vorkink (2002), ont adapté avec succès cette procédure au test GRS au contexte des lois à symétrie elliptique.

### 3.2.1 Implications de la symétrie elliptique pour l'estimation adaptative.

Dans le cas d'une loi univariée, les méthodes d'estimation non paramétriques, telles que discutées par Silverman (1986) et Pagan et Ullah (1999), peuvent être appliquées à la procédure de Rothenberg et Leenders (1968). Cependant, sur le plan multivarié, Härdle et Linton (1994) et Hodgson (1998) montrent que l'efficacité de l'estimateur non paramétrique est mise à mal par le problème du Fléau de la Dimension. En effet, sa précision est inversement proportionnelle à la dimension de la loi à estimer et se détériore rapidement avec toute augmentation de celle-ci. En théorie, l'élargissement de la taille de l'échantillon pourrait contrecarrer cet effet pervers. Mais c'est au coût élevé d'un nombre d'observations supplémentaires qui croît marginalement à un taux exponentiel.

Hodgson, Linton et Vorkink (2002) montrent comment, en exploitant les propriétés de stabilité par sommation et d'invariance en loi des variables elliptiques symétriques sous toute transformation orthogonale, la réduction de la dimension de la loi à estimer est possible. Sans perte de généralité, dans modèle de régression du MÉDAF représenté en notation vectorielle par  $R = \alpha + \beta R_m + u$ , en supposant  $u \sim E_N(0, \Sigma, g_N)$ , la distribution jointe  $p(u)$  s'écrit

$$p(u) = c_N |\Sigma|^{-1/2} g_N(u' \Sigma^{-1} u).$$

L'argument de fonction génératrice de densité  $g_N(\cdot)$  correspond en fait à une variable aléatoire univariée<sup>48</sup>  $v = (u' \Sigma^{-1} u)$ , distribuée selon  $v \sim E_1(\mu_v, \sigma_v^2, g_N)$ . Les deux variables  $u$  et  $v$  sont donc caractérisées par la même fonction génératrice de densité  $g$  scalaire (univariée)  $g_N(u' \Sigma^{-1} u) \equiv g(v)$ . Aussi,  $p(u) = c_N |\Sigma|^{-1/2} g_N\{u' \Sigma^{-1} u\}$  peut être estimé de

<sup>48</sup>  $v_t = (\varepsilon_t' \varepsilon_t)$  à chaque période  $t = [1, 2, 3, \dots, T]$ , où  $\varepsilon_t = \Sigma^{-1/2} u_t$  est le vecteur de termes d'erreur standardisés du modèle. En effet,  $u_t' \Sigma^{-1} u_t = u_t' \Sigma^{-1/2} \Sigma^{-1/2} u_t$ .

manière non paramétrique selon  $\hat{p}(\hat{u}) = \hat{c}_N |\hat{\Sigma}|^{-1/2} \hat{g}(\hat{v})$ . Les auteurs ont développé une stratégie d'estimation convergente et efficace pour le vecteur de *Score* à partir de  $\hat{g}_N$ , qui vérifie la condition d'adaptabilité de Bickel (1986). En outre, ils montrent que la performance de l'estimateur de  $\hat{p}(u)$  est sensible à une transformation de Box-Cox (1964)<sup>49</sup> appliquée sur la variable  $v$  selon  $z = \tau(v) = \left(\frac{v^\lambda - 1}{\lambda}\right)$ . En réécrivant  $\hat{p}(\hat{u})$  en fonction la distribution univariée  $\hat{\gamma}(\hat{z})$  on abouti à  $\hat{p}(\hat{u}) = |\hat{\Sigma}|^{-1/2} [\tau^{-1}(\hat{z})]^{1-N/2} J_\tau(\hat{z})^{-1} \hat{\gamma}(\hat{z})$ . En effet, par le théorème de Casella et Berger (1990), on a  $\hat{g}_N(v) = c_N^{-1} [\tau^{-1}(\hat{z})]^{1-N/2} J_\tau(\hat{z})^{-1} \hat{\gamma}(\hat{z})$  où  $J_\tau(z)$  représente le terme jacobien de la distribution de  $v$  et  $c_N$  une constante calculée proportionnellement à  $g_N$ .

Hodgson et *alii* (2002) ont estimé  $\gamma(z)$  par la méthode du noyau de densité avec une largeur de bande optimale fixe selon la méthode de Silverman (1986). Alternativement, l'approche des *K-plus proches voisins* ou *K-Nearest Neighbors* de Loftsgaarden et Quesenberry (1965), celle du Maximum de Vraisemblance Pénalisé de Good and Gaskins (1971), la largeur de bande variable d'Abramson (1982), ou le noyau variable de Breiman, Meisel et Purcell (1977) et finalement le noyau gamma, adaptatif de Chen (2000) pourraient apporter davantage de flexibilité et de précision dans l'estimation de  $\hat{\gamma}(\hat{z})$ .

### 3.2 Procédure du test semi-paramétrique de H.L.V. (2002).

#### 3.2.1 Mise en œuvre de l'estimation.

L'estimative adaptative du MÉDAF se construit en quatre étapes comme suit. Pour le modèle<sup>50</sup>  $R = \theta R_m + u$ , la première étape vise à construire la variable  $\hat{z}_t$ , pour laquelle la densité  $\gamma(\hat{z}_t)$  et sa dérivée première  $\gamma'(\hat{z}_t)$  sont à estimer. Les coefficients  $\hat{\theta} = [\hat{\alpha}, \hat{\beta}]'$  et les

<sup>49</sup> Davidson et MacKinnon (2004) ont présenté une discussion sur le rôle de la transformation de Box-Cox dans l'estimation économétrique.

<sup>50</sup>  $R$  est le vecteur de séries de rendements excédentaires des titres,  $R_m$  est la prime de risque de marché,  $\theta = [\alpha, \beta]'$  l'ensemble de paramètres du MÉDAF et  $u$  le vecteur de termes d'erreur de la régression.

résidus  $\hat{u}$  du modèle sont donc obtenus par Régressions Apparemment Indépendantes. Les résidus  $\hat{u} \in \mathbb{R}^{T \times N}$  sont ensuite standardisés<sup>51</sup> selon  $\hat{\varepsilon} = \hat{\Sigma}^{-1/2} \hat{u}$  pour servir à la construction de la variable univariée  $\hat{v}_t$  et en dernière instance  $\hat{z}_t$ .

Dans l'estimation du vecteur du *Score* associé à la variable  $\hat{z}_t$ <sup>52</sup>, interviennent les fonctions  $\hat{\gamma}_t(\hat{z}_t)$  et  $\hat{\gamma}'_t(\hat{z}_t)$  qui sont estimées lors de la seconde étape selon

$$\hat{\gamma}_t(\hat{z}_t) = \frac{1}{T-1} \sum_{\substack{s=1 \\ s \neq t}}^T K_{h_n}(\hat{z} - \hat{z}_s) \text{ et } \hat{\gamma}'_t(\hat{z}_t) = \frac{1}{T-1} \sum_{\substack{s=1 \\ s \neq t}}^T K'_{h_n}(\hat{z} - \hat{z}_s).$$

Hodgson, Linton et Vorkink (2002) utilisent un noyau gaussien<sup>53</sup>  $K_{h_n}$  et une largeur de bande fixe  $h_n$  optimale selon le règle de Silverman (1986). L'efficacité de cette estimation, peu sensible au choix de  $K_{h_n}$ , est étroitement liée à  $h_n$ .

Dans la troisième étape, l'ensemble de vecteur de *Score*  $\hat{\varphi}_t(\hat{u}_t)$  est estimé selon

$$\hat{\varphi}_t(\hat{u}_t) = \hat{\Sigma}^{-1/2} \hat{\varepsilon}_t \left[ s(\hat{v}_t) + \tau'(\hat{v}_t) \cdot \frac{\hat{\gamma}'_t(\hat{z}_t)}{\hat{\gamma}_t(\hat{z}_t)} \right] \text{ avec } t = [1, 2, 3, \dots, T],$$

où  $s(\hat{v}_t) = (1 - N/2)v^{-1} - \frac{J'_\tau(\hat{z})}{J_\tau(\hat{z})} \cdot \tau'(\hat{v}_t)$  avec  $J_\tau(\hat{z}) = J_\tau(\tau(\hat{v}_t)) = \left| \frac{\partial \tau^{-1}(z)}{\partial z} \right|$ ,  $J'_\tau(\hat{z})$  sa dérivée première et  $\tau'(\hat{v}_t)$  la dérivée première de la fonction de transformation de Box-Cox  $\tau(\hat{v}_t)$ . L'estimation convergente du *Score*  $\hat{\Delta}(\hat{\theta})$  ainsi que la matrice variance-covariance asymptotique  $[\hat{I}(\hat{\theta})]^{-1}$  s'achève par les calculs

$$\hat{\Delta}(\hat{\theta}) = -\frac{1}{\sqrt{T}} \sum_{t=1}^T \omega'_t \hat{\varphi}_t(\hat{u}_t) \text{ et } [\hat{I}(\hat{\theta})]^{-1} = \left[ \frac{1}{T} \sum_{t=1}^T \omega'_t \hat{\Omega} \omega_t T^{-1} \right]^{-1}.$$

<sup>51</sup>  $\hat{\Sigma} \in \mathbb{R}^{N \times N}$  est un estimateur convergent de leur matrice variance-covariance  $\hat{\Sigma} = \hat{\varepsilon}^{-1} \hat{\Sigma}_{\hat{u}}$  où  $\hat{\Sigma}_{\hat{u}} = \left[ \frac{1}{(T-K-N)} \sum_{t=1}^T \hat{u} \hat{u}'_t \right]$  et  $\hat{\varepsilon} = [\det \hat{\Sigma}_{\hat{u}}]^{1/N}$ . Selon Hodgson et alii (2002), la matrice  $\hat{\Sigma}_{\hat{u}}$  est ajustée par  $\hat{\varepsilon}^{-1}$  pour tenir compte de l'hypothèse de symétrie elliptique imposée sur le vecteur  $\hat{u}$ .

<sup>52</sup> Si à chaque période  $t$ , la fonction de log-vraisemblance estimée de  $\hat{z}_t$  est  $\ln \{ \hat{\gamma}_t(\hat{z}_t) \}$  alors le *Score* pour  $\hat{z}_t$ , s'obtient selon  $\frac{\hat{\gamma}'_t(\hat{z}_t)}{\hat{\gamma}_t(\hat{z}_t)}$ .

<sup>53</sup> Dans la procédure d'estimation, chaque observation de  $\hat{z}_t$  de l'échantillon est, à tour de rôle, utilisée comme centre d'intervalle notée  $\hat{z}_s$  pour évaluer localement le noyau de densité selon  $K_{h_n}(u) = \left[ (2\pi h_n^2)^{-\frac{1}{2}} \right] e^{-\frac{u^2}{2h_n^2}}$  et sa dérivée selon  $K'_{h_n}(u) = -\left( \frac{u}{h_n^2} \right) \left[ (2\pi h_n^2)^{-\frac{1}{2}} \right] e^{-\frac{u^2}{2h_n^2}}$  avec  $u_s = (\hat{z} - \hat{z}_s)$  pour  $s = 1, 2, \dots, T$  où  $t \in \mathbb{R}^{T \times 1}$  et  $h_n$  la largeur de bande. Cette technique est appelée *Leave-One-Out kernel density*.

où  $\omega_t$  représente le vecteur d'observations des  $K$  variables exogènes du modèle sur chaque période  $t$  et  $\hat{\Omega} = \frac{1}{T} \sum_{t=1}^T \{\hat{\varphi}_t(\hat{u}_t)\hat{\varphi}_t(\hat{u}_t)'\}$  représente un estimateur de la matrice hessienne  $E[\varphi_t(u_t)\varphi_t(u_t)']$  de  $\varphi_t(u_t)$ . L'estimateur adaptatif  $\tilde{\theta}$  s'écrit alors

$$\tilde{\theta} = \hat{\theta} + \left\{ \left( \frac{1}{\sqrt{T}} \right) [\hat{I}(\hat{\theta})]^{-1} \hat{\Delta}(\hat{\theta}) \right\}.$$

Le premier  $\tilde{\theta}$  obtenu par cette mise en œuvre est une solution admissible sur la famille à symétrie elliptique, au sens de Bickel (1982). Dans un souci d'optimalité semi-paramétrique, Hodgson, Linton et Vorkink (2002) proposent d'appliquer un algorithme de Newton-Raphson<sup>54</sup>, consistant à itérer les quatre étapes de la procédure, énumérées ci-haut, un nombre  $J$  de fois jusqu'à déterminer le  $j^{\text{ième}}$   $\tilde{\theta}$  qui minimise globalement la variance asymptotique pour la séquence  $\tilde{\theta}_j = \tilde{\theta}_{j-1} + \left\{ \left( \frac{1}{\sqrt{T}} \right) [\hat{I}(\tilde{\theta}_{j-1})]^{-1} \hat{\Delta}(\tilde{\theta}_{j-1}) \right\}$ . Ils ont ainsi montré que des gains d'efficacité peuvent résulter de cette étape supplémentaire.

### 3.2.2 Conditions nécessaires à l'adaptabilité.

Un ensemble de conditions sont nécessaires à la convergence du *Score* et de la matrice variance-covariance. *Primo*, il faut que l'hypothèse de distribution à symétrie elliptique ne soit pas rejetée par les données. Sinon la densité jointe  $p(u)$ , n'est plus représentable selon  $p(u) = c_N |\Sigma|^{-1/2} g_N(u'\Sigma^{-1}u)$  et il n'est plus certain de pouvoir l'estimer en fonction d'une densité univariée. On revient à ce moment à la problématique du Fléau de la Dimension. L'hypothèse de symétrie elliptique, sur la base d'un test de Béran (1979)<sup>55</sup> n'a pas été rejetée avec l'échantillon de Hodgson et *alii* (2002). Des tests plus modernes de symétrie elliptique existent. On peut mentionner celui de Romano (1988), de Li, Fang et Zhu (1997), celui de Manzotti et *al.* (2002), de Huffer et Park (2007). La plupart sont d'une grande complexité de mise en œuvre et font intervenir des notions statistiques qui exigeraient

<sup>54</sup> Les résidus de l'étape d'estimation adaptative précédente sont réintroduits dans un nouveau cycle d'estimation en quatre étapes jusqu'à ce que l'efficacité de l'estimateur atteigne sa limite locale.

<sup>55</sup> Pour un vecteur  $X_N \sim E_N(\mu, \Sigma, g_N)$ , le test de Béran (1979) mesure la distance entre la distribution de l'échantillon et la famille à symétrie elliptique avec une statistique calculée en fonction de  $\hat{\mu}$  et  $\hat{\Sigma}$ .

un développement à part, et qui dépassent la portée de la discussion dans ce travail. L'approche de Li, Fang et Zhu (1997), est la plus simple et la plus intuitive. Aussi, elle servira de cadre méthodologique du test.

La partie paramétrique du modèle, qui fournit les résidus  $\hat{u}_t$  doit être estimée de façon robuste à l'autocorrélation et l'hétéroscédasticité sur les  $N$  régressions individuelles incorporées dans l'estimation par RAI. Comme le montrent Zhou (1993) et Hodgson et alii (2002), l'hypothèse de rendements à symétrie elliptique implique une forme de dépendance temporelle du second ordre entre les rendements et le portefeuille de marché. Aussi, dans chaque équation, les variables (endogènes et exogènes) sont pondérées par l'inverse de la volatilité conditionnelle des termes d'erreur correspondants. Celle-ci est estimée d'abord, selon l'approche de Nadaraya (1964) et Watson (1964) avec

$$\hat{\sigma}_t = \left( \frac{\sum_{\substack{s=1 \\ s \neq t}}^T K_{h_n}(R_m - R_{m_s}) \hat{u}_t^2}{\sum_{\substack{s=1 \\ s \neq t}}^T K_{h_n}(R_m - R_{m_s})} \right)^{1/2},$$

où chaque  $\hat{\sigma}_t \in \mathbb{R}^{T \times 1} \forall t = [1, 2, 3 \dots, N]$  et ensuite selon une spécification GARCH (1,1)<sup>56</sup> avec

$$\hat{\sigma}_{j,t} = \sqrt{\hat{\gamma}_0 + \hat{\alpha} \hat{\varepsilon}_{j-1,t}^2 + \hat{\beta} \hat{\sigma}_{j-1,t}^2},$$

où  $\hat{\sigma}_t \in \mathbb{R}^{T \times 1} \quad t = [1, 2, 3 \dots, N]$  et  $j = [1, 2, 3 \dots, T]$ . Finalement, la matrice variance-covariance  $\Sigma$  des termes d'erreur est corrigée du nombre de degrés de liberté en petit échantillon selon  $\hat{\Sigma} = \hat{c}^{-1} \hat{\Sigma}_{\hat{u}}$  où  $\hat{\Sigma}_{\hat{u}} = \left[ \frac{1}{(T-K-N)} \sum_{t=1}^T \hat{u} \hat{u}' \right]$  et  $\hat{c} = [\det \hat{\Sigma}_{\hat{u}}]^{1/N}$ .

En supposant que la symétrie elliptique ne soit pas rejetée, l'adaptabilité de l'estimateur est liée à la précision des estimateurs non paramétriques  $\hat{\gamma}(\hat{z}_t)$  et  $\hat{\gamma}'(\hat{z}_t)$ , qui elle-même est tributaire du choix du paramètre optimal  $\lambda$  pour la transformation de Box-Cox  $\hat{z}_t = \left( \frac{\hat{\sigma}^\lambda - 1}{\lambda} \right)$ . Hodgson et alii (2002), en tenant compte des propriétés de la fonction  $g_N$ , ont établi que  $\lambda^* = \frac{N}{2}$ . D'autre part, le support de la distribution de  $\hat{z}_t$  admet la valeur 0 comme borne

<sup>56</sup> Hodgson, Linton et Vorkink (2002) ont trouvé que cette spécification est suffisamment parcimonieuse pour purger le modèle de l'hétéroscédasticité conditionnelle.

inférieure<sup>57</sup>. Or l'estimateur du noyau de densité est conçu pour une variable définie sur  $(-\infty, +\infty)$ . Cela entraîne un biais dans  $\hat{\gamma}(\hat{z}_t)$  et  $\hat{\gamma}'(\hat{z}_t)$  pour des valeurs de  $\hat{z}_t$  au voisinage de 0. Plusieurs techniques existent pour corriger ces biais aux frontières. Une correction de Schuster (1985)<sup>58</sup> a été appliquée par les auteurs. Finalement, la largeur de bande optimale a été déterminée en fonction de la convergence en probabilité du ratio  $\frac{\hat{\gamma}'(\hat{z}_t)}{\hat{\gamma}(\hat{z}_t)}$  plutôt que  $\hat{\gamma}(\hat{z}_t)$  et  $\hat{\gamma}'(\hat{z}_t)$  séparément.

Le calcul du *Score*  $\hat{\varphi}_t(\hat{u}_t)$  doit périodiquement respecter un ensemble de conditions de régularité, établies par Bickel (1982), dites conditions d'élimination<sup>59</sup>, qui permettent d'exclure de l'échantillon les observations qui peuvent, entre-autre, faire que l'estimateur soit localement non régulier ou superconvergent. Ils définissent ainsi des bornes pour des éléments de  $\hat{\varphi}_t(\hat{u}_t)$ , tels  $\hat{z}_t$  et  $\frac{\hat{\gamma}'(\hat{z}_t)}{\hat{\gamma}(\hat{z}_t)}$ . Par exemple, si sur une période on a  $\hat{\gamma}_t(\hat{z}_t) \rightarrow 0$ , alors

$$\frac{\hat{\gamma}'_t(\hat{z}_t)}{\hat{\gamma}_t(\hat{z}_t)} \rightarrow \infty \text{ et } \hat{\varphi}_t(\hat{u}_t) \rightarrow \infty.$$

Hodgson et alii (2002) ont adapté ces conditions selon

- i.  $\hat{\gamma}_t(\hat{z}_t) \geq d_n$ .
- ii.  $|\hat{z}_t| \leq e_n$ .
- iii.  $|\lambda(\hat{z}_t)| \leq b_n$ .
- iv.  $\left| \rho^{1/2}(\hat{z}_t) \hat{\gamma}'_t(\hat{z}_t) \right| \leq c_n \hat{\gamma}_t(\hat{z}_t)$ <sup>60</sup>.

<sup>57</sup>  $\hat{z}_t = \hat{\varepsilon}'_t \hat{\varepsilon}_t \equiv \sum_{i=1}^N \hat{\varepsilon}_i^2$  avec  $\hat{\varepsilon}_i \in \mathbb{R}_+^{T \times 1} \forall t = [1, 2, 3, \dots, N]$ .

<sup>58</sup> La correction de Schuster (1985) consiste à ajouter au noyau de densité  $K_{h_n}(u)$  de  $\hat{z}_t$  un noyau miroir évalué pour  $\hat{z}_t^* = -(\hat{z}_t + 2c) \forall t$ . Le calcul de la densité  $\hat{\gamma}_t(\hat{z}_t)$  se fait donc selon

$\hat{\gamma}_t(\hat{z}_t) = \frac{1}{nh_n} \sum_{i=1}^n \{K_{h_n}(\hat{z} - \hat{z}_s) + K_{h_n}(-\hat{z} - 2c + \hat{z}_s)\}$ .  $\hat{\gamma}'_t(\hat{z}_t)$  se calcule de façon similaire avec l'ajout de  $K'_{h_n}(-\hat{z} - 2c + \hat{z}_s)$ .

<sup>59</sup> Celles-ci sont désignées par *trimming conditions* en anglais et ont été largement discutées par Pagan et Ullah (1999).

<sup>60</sup> Les constantes  $b_n, c_n, d_n$  et  $e_n$  sont définies en fonction du nombre d'observations  $n = [1, 2, 3, \dots, T]$  de l'échantillon tel que  $c_n \rightarrow \infty$ ,  $b_n \rightarrow \infty$ ,  $e_n \rightarrow \infty$ ,  $h_n \rightarrow 0$ ,  $d_n \rightarrow 0$ ,  $c_n h_n \rightarrow 0$ ,  $e_n h_n^{-3} \rightarrow o(n)$  et  $b_n h_n^{-3} \rightarrow o(n)$  avec. Le terme  $\rho(\hat{z}_t) = v \tau'(v) J_\tau^{-1}(z)$  avec  $v = \tau^{-1}(z)$  et  $\lambda(\hat{z}_t) = (d/dz)^{-1} \rho^{1/2}(\hat{z}_t)$ .  $(d/dz)^{-1}$  désigne l'opérateur l'anti-dérivée.

### 3.2.3 Test de Wald adaptatif.

La statistique du test GRS est réadaptée, en substituant dans  $J_0$ , le vecteur de coefficients adaptatifs  $\tilde{\alpha}$  et leur matrice variance-covariance  $(\widehat{Var}[\tilde{\alpha}])^{-1}$  selon  $\tilde{J}_0 = \tilde{\alpha}'(\widehat{Var}[\tilde{\alpha}])^{-1}\tilde{\alpha}$ . On peut aussi écrire  $\tilde{J}_0 = (R\tilde{\theta} - q)' \{R[\hat{I}(\tilde{\theta})^{-1}]R'\}^{-1} (R\tilde{\theta} - q)$  avec  $\tilde{\theta} \in \mathbb{R}^{N \times 2} = [\tilde{\alpha}, \tilde{\beta}]'$ ,  $R \in \mathbb{R}^{L \times (2 \times N)}$ , une matrice de  $L$  restrictions linéaires sur les  $2 \times N$  paramètres  $\tilde{\theta}$ , de rang  $L < (2 \times N)$  et  $q \in \mathbb{R}^{L \times 1}$  un vecteur de zéros. Finalement,  $\hat{I}(\tilde{\theta})^{-1} \in \mathbb{R}^{(2 \times N) \times (2 \times N)}$  est la matrice variance-covariance de  $\tilde{\theta}$ . En vertu de la définition de l'adaptabilité de  $\tilde{\alpha}$ , on sait qu'il est asymptotiquement normal. Aussi, sous l'hypothèse nulle vraie, on a  $\tilde{J}_0 \sim \chi_N^2$ .

## 3.4 Test de symétrie elliptique de Li, Fang et Zhu (1997).

### 3.4.1 Principe du test.

Le test de symétrie elliptique utilisé dans ce travail est celui de Li, Fang et Zhu (1997) qui se base sur un examen visuel du diagramme Q-Q ou diagramme quantile-quantile de la distribution jointe de vecteur de variables distribuées à symétrie elliptique. Il s'inspire des méthodes de tests de normalité multivariée développées par des auteurs comme Cox et Small (1978) et Ahn (1992). Il repose sur la propriété d'invariance en distribution des statistiques dites robustes<sup>61</sup> par rapport à toute transformation orthogonale<sup>62</sup> des variables à partir desquelles elles sont formées. Li, Fang et Zhu (1997) montrent qu'un test d'invariance d'une statistique robuste équivaut à un test de symétrie elliptique, puisque les variables distribuées à symétrie elliptiques sont en elles-mêmes stables et invariantes en distribution sous toute transformation orthogonale. Ils ont proposé plusieurs statistiques satisfaisant cette propriété,

<sup>61</sup> Les propriétés des statistiques robustes sont discutées entre autres, par Fang, Kotz et Ng (1990).

<sup>62</sup> Une transformation orthogonale dans une variable consiste à pré multiplier celle-ci par une matrice de projection orthogonale, symétrique et idempotente pour obtenir une nouvelle variable qui s'interprète comme les écarts de la variable d'origine par rapport à sa moyenne.

comme la  $t$  de Student et celle de *Fisher*. Le présent travail privilégie uniquement une statistique  $t$  de Student étant donné sa facilité de mise en œuvre et d'interprétation.

Dans un contexte univarié, pour la moyenne empirique  $\bar{X}$  de la variable  $X \sim N(\mu, \sigma^2)$ , la statistique  $t = \sqrt{n} \left( \frac{\bar{X}_n - \mu}{\hat{\sigma}_n} \right) \sim t_{n-1}$ . Similairement, pour un vecteur  $X \sim E_N(\mu, \Sigma, g_N)$  avec  $X \in \mathbb{R}^{T \times N}$ , il est attendu que le vecteur de statistiques  $t\{Y_t\} = \sqrt{N} \left( \frac{Y_t - \bar{Y}_t}{\hat{\sigma}_t} \right) \sim t_{N-1} \quad \forall t = [1, 2, 3, \dots, T]$ , où  $Y = \Sigma^{-1/2} M X$  avec  $M = [I - (l'l)^{-1} l'l']$ , une matrice de projection orthogonale<sup>63</sup>. Finalement, la transformation orthogonale est pondérée par  $\Sigma^{-1/2}$  en vue de garantir que  $Y \sim N(0, I, g_N)$  et soit *i.i.d.*

### 3.4.2 Mise en oeuvre du test.

La mise en œuvre du test consiste à standardiser premièrement le vecteur de variables sur lequel il sera appliqué. Pour un vecteur  $Y$  on a  $\tilde{Y} = \Sigma^{-\frac{1}{2}} [Y - \mu] \sim E_N(0, I, g_N)$  où  $\mu$  et  $\Sigma$  désignent respectivement le vecteur de moyenne et la matrice variance-covariance de  $Y$ . Ces derniers peuvent être estimés par Maximum de Vraisemblance avec  $\hat{\mu}$  et  $\hat{\Sigma}$ . Ensuite, on forme le vecteur de dimension  $T \times 1$  de statistiques  $t_i\{\tilde{Y}_i\} = Z_i = \frac{\sqrt{N}\tilde{Y}_i}{s_i}$  avec  $i = [1, 2, 3, \dots, T]$ .  $\bar{Y}_i = \frac{1}{N} \sum_{j=1}^N \tilde{Y}_{i,t}$  et  $s_i = \frac{1}{N} \sum_{j=1}^N [\tilde{Y}_{i,t} - \bar{Y}_i]^2$  avec  $N$  est la dimension de chaque vecteur  $\tilde{Y}_i$  et  $T$ , la taille de l'échantillon. On peut tester la symétrie elliptique en évaluant graphiquement la distance entre la distribution cumulative de cette statistique  $t\{\tilde{Y}_i\}$  et celle d'une  $t$  de Student à  $N - 1$  degrés de liberté. Pour ce faire, Li, Fang et Zhu (1997) proposent de représenter sur un diagramme quantile-quantile, les statistiques  $t_{N-1}^{\tilde{\alpha}_i}$  et  $t\{\tilde{Y}_i\}$  où  $t_{N-1}^{\tilde{\alpha}_i}$  désigne la valeur lue dans une table d'une loi  $t$  de Student à  $N - 1$  degrés de liberté, associée à la probabilité  $\tilde{\alpha}_i = \left( \frac{2i-1}{2T} \right)$  avec  $i = [1, 2, 3, \dots, T]$ .  $t\{\tilde{Y}_i\}$  représente le vecteur de statistiques  $t$  calculées, triées par ordre croissant.

<sup>63</sup>  $l \in \mathbb{R}^{T \times 1}$  est un vecteur unitaire et  $I \in \mathbb{R}^{T \times T}$ , une matrice identité de rang  $T$ .

Lorsque le nuage de points de coordonnées  $(t_{n-1}^{\tilde{\alpha}_i}, t\{\tilde{Y}_i\})$  se confond avec la première bissectrice (axe faisant un angle de  $45^0$  avec l'axe des abscisses en partant de l'origine) il y a évidence empirique de symétrie elliptique dans distribution de la variable  $Y$ . Sur la figure B.1 à l'appendice B, à la colonne de gauche, sont rapportés des résultats de tests effectués sur simulations d'exemple de variables à distribution elliptique symétriques. Sur la colonne de droite de la figure, sont représentés les diagrammes Q-Q de variables dont la distribution n'appartient pas à cette classe. On peut noter que leur distribution cumulative de  $t\{\tilde{Y}_i\}$  s'écarte significativement de la première bissectrice pour celles-ci. Les auteurs proposent aussi un test basé sur la mesure du coefficient de corrélation simple  $\hat{\rho}$  entre  $t_{n-1}^{\tilde{\alpha}_i}$  et  $t\{\tilde{Y}_i\}$ , qui pour des valeurs significativement proche de l'unité, valide l'hypothèse symétrie elliptique. Une des faiblesses du test, mentionnée par Li, Fang et Zhu (1997) est de supposer le vecteur de variables *i.i.d.* Or, on sait aussi que la famille de distributions à symétrie elliptique admet aussi des lois violant l'hypothèse *i.i.d.*<sup>64</sup> Aussi, il n'a pas une puissance comparable à celle d'un test de Béran (1979) ou de Manzotti et alii (2002) par exemple.

---

<sup>64</sup> *i.i.d.* est l'abréviation de "indépendamment et identiquement distribués".

## CHAPITRE IV

### PRÉSENTATION DES RÉSULTATS ET ANALYSES

#### 4.1 Description de l'échantillon.

Les données brutes utilisées dans ce travail sont constituées de séries chronologiques du prix de fermeture d'un échantillon de 80 actions ordinaires membres de l'indice composite du *S&P TSX*. Les observations sont à fréquence journalière<sup>65</sup> et s'étalent sur la période allant du 5 janvier 2009 au 29 octobre 2014. À celles-ci s'ajoutent le taux sur les bons du trésors canadien à échéance de trois mois, représentant le titre sans risque et le cours de l'indice de référence ci-haut mentionné qui sert de portefeuille figurant. Le nombre d'observations s'élève à 1483 pour chaque titre. Le taux sans risque correspond au vecteur v39065 se trouvant au tableau 176-0048 de la base de données CANSIM de Statistiques Canada<sup>66</sup>. Toutes les autres séries proviennent de la base de données de Bloomberg, accessible à partir d'un terminal de la salle des marchés de l'École des Sciences de la Gestion de l'Université du Québec à Montréal.

Le choix de la période échantillonnale couverte pour l'application de la méthodologie du test du MÉDAF est motivé par plusieurs considérations. D'une part, l'utilisation de données plus récentes peut contribuer à actualiser la validité empirique de la procédure de Hodgson, Linton et Vorkink (2002), initialement appliquée sur un échantillon plus restreint et plus ancien. D'autre part, on peut s'attendre à ce que l'augmentation de la taille de l'échantillon améliore

---

<sup>65</sup> Il faut mentionner que seulement les journées ouvrables de la bourse de Toronto sont prises en considération dans le calcul des rendements. Cela totalise en moyenne 250 jours par année.

<sup>66</sup> L'adresse de référence est [www5.statcan.gc.ca/cansim](http://www5.statcan.gc.ca/cansim).

l'efficacité de la procédure, en favorisant entre autres, une meilleure convergence des vecteurs de Score de l'estimateur semi-paramétrique. Finalement, la période retenue coïncide à celle succédant à une crise financière mondiale majeure qui a plus ou moins perturbé le marché financier canadien et qui, potentiellement est riche en évènements pertinents aux hypothèses du modèle semi-paramétrique utilisé.

#### 4.2 Traitement des données.

La méthodologie de regroupement en portefeuilles de Black, Jensen et Scholes (1972) a été appliquée sur les 80 titres. Sur la période allant du 5 janvier 2009 au 31 décembre 2009, leur  $\beta$  est estimé dans un système d'équations multiples par la méthode des Régressions Apparemment Indépendantes<sup>67</sup>. Ils servent par la suite à ordonnancer les titres par ordre décroissant du niveau de risque et à les répartir en 10 portefeuilles composés de 8 actions chacune. Ces étapes sont réitérées année après année, jusqu'à la fin de l'échantillon. Les 255 premières observations utilisées pour initier le regroupement en portefeuilles sont écartées des calculs ultérieurs. Les principaux résultats du travail portent uniquement sur trois portefeuilles. Des analyses comparatives sont toutefois effectuées avec 2, 3 et 5 portefeuilles et avec des versions de trois portefeuilles renfermant respectivement 8, 10 et 16 titres.

Les excès de rendements des portefeuilles sont présentés en 3 versions. La première est celle des rendements d'origine, les deuxième et troisième versions sont ceux pondérées par l'inverse de volatilité conditionnelle des termes d'erreur de leur équation respective. Cette volatilité est estimée respectivement selon l'approche non paramétrique de Nadaraya (1964) et Watson (1964) et une spécification GARCH(1,1) estimé par Maximum de Vraisemblance gaussien. Le sommaire des statistiques descriptives des trois variantes de portefeuilles est disponible au tableau 4.1 ci-dessous. L'incorporation de données pondérées dans les  $N$  régressions individuelles de départ, sert à traiter les problèmes de non sphéricité des termes d'erreur, qui affectent la robustesse de l'estimateur par RAI.

---

<sup>67</sup> L'appellation Régressions Apparemment Indépendantes est la traduction française du terme *Seemingly Unrelated Regression (SUR)*.

Avec les trois versions de rendements, trois systèmes d'équations multiples ont été estimés. Ils sont désignés par *MODÈLE 1*, *2* et *3*<sup>68</sup>. Des tests d'hétéroscédasticité de White et *ARCH*<sup>69</sup> à un, trois et cinq retards sont effectués. Des tests d'autocorrélations de Breusch-Godfrey (1978) à un, trois et cinq retards et de Ljung-Box (avec un nombre optimal de retards déterminé selon un critère d'information Hannan-Quinn (1979) sont aussi appliqués. L'analyse de l'adéquation statistique des modèles paramétriques est complétée avec des tests de spécification de RESET et de stabilité structurelle de Quandt-Andrews<sup>70</sup>. Les résultats des tests d'adéquation statistique sont présentés aux tableaux B.2 à B.5 à l'appendice B.

Deux modèles semi-paramétriques, désignés par *MODÈLE AD1* et *AD2*, ont été estimés. Leur partie paramétrique incorpore des rendements, pondérés respectivement selon une approche non paramétrique et une spécification GARCH (1,1). Une correction de Zhou (1993) a été appliquée sur la statistique du test GRS construite sur le *MODÈLE 1*. La partie non paramétrique des modèles *AD1* et *AD2* a été calculée avec une largeur de bande fixe, optimale pour un noyau gaussien, selon  $h^* = \left[ 2.1621 T^{-1/7} \right] \hat{\sigma}(\hat{z}_t)$ . Celle-ci est utilisée à la fois dans l'estimation des distributions et de leur dérivée, garantissant ainsi la convergence en probabilité des vecteur de *Score*. Les biais aux frontières sont corrigés par la méthode de Schuster (1985). Finalement, les conditions d'élimination ont été imposées sur le calcul des

---

<sup>68</sup> Le *MODÈLE1* désigne celui incorporant des rendements non pondérés mais une correction de Newey-West (1987) est appliquée sur la matrice variance-covariance de ses paramètres estimés. Le *MODÈLE2* désigne celui avec les rendements pondérés par la volatilité conditionnelle estimée de façon non paramétrique et le *MODÈLE3* désigne celui ajusté avec une volatilité spécifiée selon un processus *GARCH(1,1)*.

<sup>69</sup> Un test *ARCH* permet de détecter la présence d'une forme autorégressive de l'hétéroscédasticité dans un modèle économétrique. Dans la notation *ARCH(p)*, le terme *p* désigne le nombre de retards du carré du terme d'erreur incorporé dans la régression auxiliaire spécifiée pour construire le test. Une représentation *ARCH(1)* peut en général, s'avérer trop restrictive. L'omission de retards additionnels pertinents fait courir le risque de perte de puissance du test. Par conséquent, des tests *ARCH(3)* et *ARCH(5)* sont aussi effectués.

<sup>70</sup> Le test de Quandt-Andrews ne fait aucune hypothèse *ex ante* sur la période charnière marquant un bris structurel. En éliminant au moins 15% des observations (au début et à la fin) de l'échantillon, il procède en une série de tests de Chow appliqués en considérant une-à-une, les observations de l'échantillon tronqué comme période de rupture. La statistique du test accusant la valeur maximale sur l'échantillon est considérée pour faire de l'inférence.

Scores avec une paramétrisation identique à celle utilisée<sup>71</sup> empiriquement par Hodgson et alii (2002). Tous les calculs ont été complétés à partir du logiciel Matlab, version R2014b.

#### 4.1 Statistiques descriptives des données.

Les statistiques descriptives des rendements (pondérés ou non) des portefeuilles rapportés au tableau 4.1 ci-dessous n'appuient pas l'hypothèse de normalité, sur la base d'un test de Jarque-Béra (1981) qui rejette fortement l'hypothèse nulle, avec l'évidence d'excès de kurtosis.

Tableau 4.1 Statistiques descriptives des portefeuilles.

Période : 4 Janvier 2010 au 29 octobre 2014.

Nombre d'observations : 1227

Titre	Moyenne	Écart-type	Kurtosis	Stat. de Jarque-Béra	valeur-p
Excès de rendements non pondérés					
Portefeuille 1	0,000	0,010	6,427	698,052	< 0,01
Portefeuille 2	0,000	0,008	5,584	373,909	< 0,01
Portefeuille 3	0,000	0,007	8,283	1426,834	< 0,01
Excès de rendements pondérés par la volatilité conditionnelle non paramétrique.					
Portefeuille 1	-0,043	1,120	5,590	353,981	< 0,01
Portefeuille 2	-0,072	1,138	5,649	360,641	< 0,01
Portefeuille 3	-0,122	1,181	5,749	413,806	< 0,01
Excès de rendements pondérés par la volatilité conditionnelle GARCH(1,1).					
Portefeuille 1	-0,058	0,982	3,558	18,613	< 0,01
Portefeuille 2	-0,087	0,988	3,383	8,268	0,02
Portefeuille 3	-0,110	0,993	3,755	29,618	< 0,01

<sup>71</sup> Les calculs de Hodgson, Linton et Vorkink (2002) ont été programmés sur le logiciel GAUSS et les archives sont disponibles en ligne à l'adresse suivante <http://marriottschool.net/emp/tsp/adsur.html>.

Des faits similaires sont observables pour les rendements des actions dont les statistiques descriptives sont présentées au tableau A.1 à l'appendice A. Le rejet de la normalité est aussi confirmé par les résultats du tableau 4.2 ci-dessous qui résume une analyse multivariée de Mardia (1970) sur les portefeuilles. Celle-ci confirme des coefficients de kurtosis multivarié significativement plus élevés que ceux observables sous une loi normale jointe. La présence d'asymétrie semble est aussi détectée par le test.

Tableau 4.2 Test de Mardia (1970) sur les 3 portefeuilles.

Version	Méthode	Coefficient	Statistique	d.d.l.	p-value	
Non pondérés	-	S	0,48	97,21	10	< 0,01
		S*	0,48	97,56	10	< 0,01
		K	29,44	46,18		< 0,01
Pondérés	Non paramétrique	S	0,78	159,47	10	< 0,01
		S*	0,78	160,06	10	< 0,01
		K	29,25	45,57		< 0,01
	GARCH(1,1)	S	0,28	56,82	10	< 0,01
		S*	0,28	57,03	10	< 0,01
		K	19,38	14,00		< 0,01

*S* désigne le coefficient d'asymétrie. *S\** désigne le coefficient d'asymétrie ajusté pour le biais en petit échantillon et *K* désigne le coefficient du kurtosis multivarié.

L'hypothèse de symétrie elliptique n'est pourtant pas rejetée sur la base du test de Li, Fang et Zhu (1997) appliqué sur les 80 actions. Cette contradiction s'explique par la sensibilité de la puissance du test de Mardia (1970) par rapport à la vraie distribution des données. Dans certains cas, la bimodalité, admise dans la famille de lois elliptiques symétriques, peut affecter sa puissance à rejeter l'asymétrie. Ceci a été discuté par Doornik et Hansen (1994) et étudié empiriquement par Eye (2005), à partir de simulations de Monte-Carlo. Il est donc plus prudent de s'appuyer sur le test de Li, Fang et Zhu (1997) qui est construit avec une statistique robuste et est plus apte à ne pas rejeter la symétrie elliptique lorsqu'elle est vraie.

L'examen visuel du diagramme Q-Q (ou diagramme quantile-quantile) sur la figure 4.1 ci-dessous, ne semble pas indiquer d'écart significatif de la distribution cumulative de la statistique du test par rapport à la première bissectrice (la ligne en pointillés). Un coefficient

de corrélation  $\hat{\rho}$ , entre la distribution cumulative de probabilité de la statistique du test de Li, Fang et Zhu (1997) et celle d'une distribution *t de Student* à 79 degrés de liberté indique une forte dépendance.

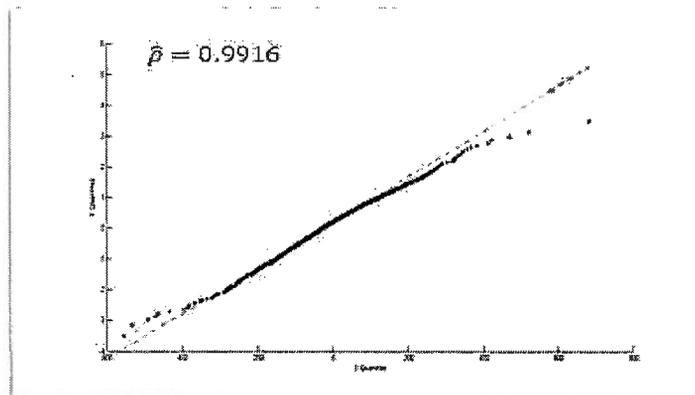


Figure 4.1 Test de Li, Fang et Zhu (1997) sur les 80 actions.

Le test n'a pas été appliqué sur les portefeuilles eux-mêmes étant donné leur nombre insuffisant pour en garantir la précision. La significativité des résultats peut cependant être étendue aux portefeuilles, en vertu de la propriété de stabilité des variables distribuées à symétrie elliptique (se référer à la sous-section 2.2.2 au chapitre 2).

#### 4.4 Résultats et analyses.

Le *MODÈLE3* s'avère le plus statistiquement adéquat pour garantir l'adaptabilité semi-paramétrique. Au tableau B.2, les tests d'hétéroscédasticité ne rejettent pas, au seuil de significativité de 5%, l'hypothèse nulle pour chaque équation de ce modèle. Idem pour le test d'autocorrélation de Ljung-Box (1978) avec 7 retards optimaux, rapporté au tableau B.3. Ce dernier est plus robuste que le test de Breusch-Godfrey (1978) spécifié arbitrairement avec 1, 3 et 5 retards qui indique un rejet de l'hypothèse nulle à 5%. Ces résultats renforcent la robustesse<sup>72</sup> et la validité du test de RESET de Ramsey (1969), rapporté au tableau B.4, qui ne rejette pas l'hypothèse nulle d'absence d'erreur de spécification au même seuil considéré.

<sup>72</sup> Il faut rappeler qu'un test de RESET construit avec un modèle économétrique affecté des problèmes discutés plus haut est biaisé et voit sa puissance systématiquement diminuée.

Cela est tout aussi valable pour le test de stabilité structurelle de Quandt-Andrews, développé par Quandt (1960) et Andrews (1993). Les résultats rapportés pour ce dernier au tableau B.5 n'indiquent pas de détection d'instabilité des paramètres au seuil de significativité de 5 %. Les *MODÈLE 1* et 2 n'accusent pas une telle robustesse. On note dans certaines équations, la présence d'autocorrélation, d'hétéroscédasticité et d'erreur de spécification.

Au tableau 4.3, on note que pour les trois modèles paramétriques, les coefficients  $\alpha$  sont estimés avec des valeurs significativement non nulles, au seuil de 5%. Ils sont aussi de signe négatif, ce qui traduit le fait que les portefeuilles ont enregistré des rendements inférieurs à ceux prédits par le MÉDAF de Sharpe-Lintner-Mossin. Les coefficients  $\beta$  ne sont pas significatifs au seuil de significativité de 5%. Ces résultats indiquent la non validité du modèle sur l'échantillon. Les résultats d'un test de Wald sur les coefficients  $\alpha$ , appliqué sur les différents modèles, rapportés au tableau 4.4, confirme le rejet du MÉDAF au seuil de significativité de 1%.

Tableau 4.3 Estimation paramétrique du MÉDAF.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

Version	Titre	$\hat{\alpha}$			$\hat{\beta}$		
		valeur	$\hat{\sigma}_{\hat{\alpha}}$	valeur-p	valeur	$\hat{\sigma}_{\hat{\beta}}$	valeur-p
MODÈLE 1	1	0,000	< 0,01	0,191	0,018	0,037	0,619
	2	0,000	< 0,01	0,083	0,036	0,040	0,363
	3	0,000	< 0,01	< 0,01	-0,017	0,027	0,530
MODÈLE 2	1	-0,042	0,032	0,185	-0,037	0,034	0,285
	2	-0,070	0,032	0,030	-0,056	0,030	0,065
	3	-0,120	0,034	< 0,01	-0,048	0,026	0,063
MODÈLE 3	1	-0,058	0,028	0,037	0,010	0,031	0,751
	2	-0,088	0,028	0,002	0,054	0,031	0,076
	3	-0,109	0,028	< 0,01	-0,019	0,024	0,429

## 4.4.1 Résultats de la procédure semi-paramétrique.

Les tests paramétrique et semi-paramétrique sont concordants quant au rejet du MÉDAF au seuil de significativité de 1%. Au tableau 4.5, la valeur de la statistique  $\tilde{J}_0$  construite sur l'approche semi-paramétrique est significativement plus élevée que  $J_0$ , obtenue par les modèles paramétriques, même avec la correction de Zhou (1993). Le MÉDAF est donc rejeté plus fortement avec l'approche adaptative.

Tableau 4.4 Estimation semi-paramétrique du MÉDAF.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

Version	Titre	$\tilde{\alpha}$			$\tilde{\beta}$		
		valeur	$\tilde{\sigma}_{\tilde{\alpha}}$	valeur-P	valeur	$\tilde{\sigma}_{\tilde{\beta}}$	valeur-P
MODÈLE AD1	1	-0,040	0,005	< 0,01	-0,054	0,006	< 0,01
	2	-0,074	0,007	< 0,01	-0,077	0,008	< 0,01
	3	-0,116	0,007	< 0,01	-0,056	0,006	< 0,01
MODÈLE AD2	1	-0,052	0,009	< 0,01	0,015	0,010	< 0,01
	2	-0,090	0,010	< 0,01	0,060	0,011	< 0,01
	3	-0,094	0,010	< 0,01	-0,014	0,009	< 0,01

Sur la base des résultats du tableau 4.4 ci-haut, on note que les coefficients  $\alpha$  estimés par le *MODÈLE AD2* sont aussi mesurés avec un signe négatif mais avec une valeur en moyenne 7,19% plus faible que ceux du *MODÈLE3*. De plus, les  $\beta$  estimés sont en moyenne 11,97% plus élevés et significatifs aux seuils de 5%. Ces résultats appuient toujours le constat que le modèle de Sharpe-Lintner-Mossin a surestimé la rentabilité des titres boursiers canadiens sur la période. En termes de convergence, 18 journées de l'échantillon, soit 1,5% des observations ont dû être écartées de la procédure de calcul du *Score*, en vertu des conditions d'élimination. En termes de précision des estimateurs, les écart-types mesurés des coefficients  $\hat{\alpha}$  et  $\hat{\beta}$  du *MODÈLE AD2* sont respectivement 60,5% et 56,8% plus faibles que

ceux de la version paramétrique correspondante. Ces résultats témoignent de la meilleure performance des modèles adaptatifs lorsqu'ils sont purgés de l'hétéroscédasticité conditionnelle et exempts d'erreur de spécification.

Tableau 4.5 Tests paramétriques vs semi-paramétriques.

Période : 4 Janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

Version	Statistique	valeur-p
$J_0$		
<i>MODÈLE1</i>	9151,115	< 0,01
<i>MODÈLE 2</i>	1606,066	< 0,01
<i>MODÈLE 3</i>	2143,430	< 0,01
$\tilde{J}_0$		
<i>MODÈLE1</i>	9266,028*	< 0,01
<i>MODÈLE AD1</i>	18718,787	< 0,01
<i>MODÈLE AD2</i>	10079,863	< 0,01

\* La statistique  $J$  associée à *MODÈLE1* correspond celle ajustée avec la correction de Zhou (1993)

#### 4.4.2 Sensibilité de l'efficacité semi-paramétrique.

Le tableau B.6 à l'appendice B, illustre la sensibilité des écart-types mesurés des coefficients du *MODÈLE AD3* au nombre d'équations introduites dans l'estimation du MÉDAF, la constitution des portefeuilles en terme de nombre de titres et la taille de l'échantillon. Dans chaque scénario, des changements sont effectués, *caeteris paribus* (par rapport aux résultats présentés au tableau 4.4). On remarque que l'augmentation du nombre d'équations et du nombre de titres par portefeuille provoquent toutes un accroissement de la précision des

coefficients  $\hat{\alpha}$  et  $\hat{\beta}$  estimés (mesurée par la diminution de leur écart-type estimé). De plus, l'élargissement de l'échantillon a contribué à l'efficacité de la procédure.

## CONCLUSION

L'objectif du présent travail était de montrer qu'en présence de rendements distribués selon la famille à symétrie elliptique, un test du MÉDAF reposant sur une approche semi-paramétrique, adaptative est plus efficace qu'un test basé sur approche strictement paramétrique. Les principaux fondements statistiques et économétriques pertinents au sujet ont été exposés. Leur importance et applications pratiques en économie financière ont été dégagées. La méthodologie de référence de Hodgson, Linton et Vorkink (2002) a été exposée et discutée. L'augmentation du nombre d'observations, du nombre de portefeuilles ainsi que de nombre d'actifs par portefeuille améliore la performance de cette procédure. Un test de symétrie elliptique facile de mise en œuvre a été présenté et appliqué. Les données de l'échantillon ne rejettent pas la symétrie elliptique avec le test de Li, Fang et Zhu (1997).

L'approche semi-paramétrique, s'est révélée empiriquement plus performante que les estimateurs paramétriques. Le meilleur modèle adaptatif estime des coefficients  $\hat{\beta}$  et  $\hat{\alpha}$  en moyenne 7,19% plus faibles et 11,97% plus élevés respectivement, avec des écart-types mesurés, respectivement 60,5% et 56,8% plus faibles que le meilleur modèle paramétrique, en moyenne. Avec des  $\hat{\alpha}$  négatifs et significatifs, les modèles indiquent que le MÉDAF de Sharpe-Lintner-Mossin a surévalué la rentabilité des titres canadiens sur la période. Aussi, les deux approches du test GRS le rejettent fortement. Néanmoins, l'approche semi-paramétrique est en principe, plus apte à rejeter le modèle car des statistiques de Wald significativement plus élevées sont obtenues avec elle.

L'optimalité de ces résultats doit toutefois être prise avec réserve. Le présent travail comporte certaines limitations qu'il convient de souligner. D'une part, la robustesse des estimateurs non paramétriques du noyau de densité utilisés n'a pas été investiguée en profondeur. Les techniques utilisées pour la sélection de largeur de bande optimale s'inspirant de la méthode de Silverman (1986) reposent sur le choix arbitraire d'une densité de référence. Une plus

grande efficacité dans l'estimation non paramétrique pourrait être obtenue en appliquant des techniques de Validation-Croisée ou un noyau de densité et une largeur de bande variables qui s'adaptent aux données, selon une approche similaire à celle de Chen (2000). Ils pourraient s'appliquer aussi à l'estimation non paramétrique de la variance conditionnelle des résidus, qui traiterait plus parcimonieusement le *MODÈLE 2* pour l'hétéroscédasticité. En somme, intuitivement il existe des raisons de penser que la performance de la procédure adaptative puisse être améliorée en faisant des innovations dans cette direction.

Finalement, dans la littérature récente de ces 20 dernières années, il y a de plus en plus d'intérêt pour la famille de lois elliptiques asymétriques plutôt que la classe à symétrie elliptique, qui est relativement plus restrictive. L'asymétrie est un des faits stylisés en finance appliquée. Élargir la portée des modèles d'évaluation d'actifs financiers, ainsi que la modélisation semi-paramétrique adaptative à cette famille de lois pourrait être une piste de recherche alternative et porteuse d'avancées significatives en économie financière.

## APPENDICE A

### STATISTIQUES DESCRIPTIVES

Tableau A.1 Statistiques descriptives et  $\beta$  des actifs.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

Description	$\hat{\beta}$	$\hat{\mu}$	$\hat{\sigma}$	Kurtosis	Stat. de J-B	Valeur-p
Bond du trésor à échéance 3 mois	S/O	0,000	0,000	2,026	202,0	< 0,01
Indice Composite du S&P TSX	1,000	0,000	0,011	5,536	339,0	< 0,01
BlackBerry Lte.	0,200	-0,001	0,027	8,148	1 642,4	< 0,01
Kinross Gold Corporation	0,170	-0,001	0,018	6,333	715,1	< 0,01
Linamar Corporation	0,113	0,000	0,034	7,333	1 212,2	< 0,01
Cameco Corporation	0,149	0,000	0,028	11,117	4 252,9	< 0,01
Potash Corp./Saskatchewan Inc.	0,112	-0,001	0,028	9,678	2 958,4	< 0,01
Methanex Corporation	0,063	0,000	0,055	26,991	36 470,0	< 0,01
Algonquin Power & Utilities Corp.	0,092	-0,001	0,012	5,606	426,8	< 0,01
Goldcorp Inc.	0,113	-0,001	0,019	6,888	936,0	< 0,01
AGF Management Limited	0,090	0,001	0,036	9,494	2 827,2	< 0,01
Eldorado Gold Corp (USA).	0,122	-0,001	0,021	5,921	533,9	< 0,01
Encana Corporation	0,073	-0,002	0,025	6,710	891,0	< 0,01
Agrium Inc.	0,080	0,000	0,038	5,787	521,1	< 0,01
Allied Properties Real Estate Investment	0,069	0,000	0,030	6,255	691,4	< 0,01
Treasury Metals Inc	0,050	-0,001	0,019	10,449	3 437,8	< 0,01
Westjet Airline Vtg	0,033	-0,001	0,019	6,161	657,0	< 0,01
Bombardier, Inc.	0,022	0,000	0,029	5,562	451,7	< 0,01
SNC-Lavalin Group Inc	0,047	0,000	0,029	7,651	1 411,4	< 0,01
Toromont Industries Ltd	0,013	-0,001	0,016	7,316	1 188,5	< 0,01
North West Company Inc	0,033	-0,001	0,027	10,670	3 642,9	< 0,01

Canfor Corporation	0,016	-0,001	0,015	6,446	782,3	< 0,01
Aecon Group Inc	0,025	-0,001	0,024	13,290	6 808,8	< 0,01
Calfrac Well Services Ltd	0,048	-0,001	0,033	7,405	1 226,4	< 0,01
Banque Laurentienne du Canada	0,015	0,000	0,045	95,409	533 396,2	< 0,01
Alimentation Couche-Tard Inc Classe B	0,007	0,000	0,043	14,231	7 915,7	< 0,01
Rona Inc.	0,033	0,001	0,043	4,665	191,2	< 0,01
Imperial Oil	0,022	-0,001	0,014	5,156	291,4	< 0,01
Maple Leaf Foods Inc.	0,011	0,001	0,027	4,924	242,5	< 0,01
Penn West Petroleum Ltd	0,063	0,000	0,027	14,181	8 138,5	< 0,01
Dominion Diamond Corp.	0,010	-0,001	0,020	6,882	969,8	< 0,01
Enerplus Corp.	0,023	-0,001	0,015	10,585	3 557,4	< 0,01
Canadian Tire Corporation Ltd Class A	0,000	-0,001	0,014	5,488	389,7	< 0,01
CAE, Inc.	-0,012	0,001	0,035	5,808	522,0	< 0,01
Le Groupe Jean Coutu PJC Inc	-0,006	-0,001	0,023	14,171	7 777,7	< 0,01
Boardwalk REIT	-0,003	0,000	0,020	5,722	492,9	< 0,01
Canadian Oil Sands Ltd.	-0,004	-0,001	0,038	5,730	461,9	< 0,01
Agnico-Eagle Mines	0,057	0,003	0,041	14,913	9 109,0	< 0,01
Suncor Energy Inc.	0,010	-0,002	0,039	14,343	7 970,7	< 0,01
Loblaw Companies Ltd	-0,006	-0,002	0,031	10,232	3 271,8	< 0,01
Fairfax Financial Holdings Ltd	-0,014	-0,001	0,023	8,207	1 684,2	< 0,01
Toronto-Dominion Bank	-0,015	-0,001	0,032	7,125	1 052,7	< 0,01
Telus Corp.	-0,018	-0,002	0,017	10,221	3 237,8	< 0,01
Quebecor, Inc. Class B	-0,020	-0,001	0,033	21,142	20 908,4	< 0,01
Canadian Western Bank	-0,020	0,000	0,039	10,158	3 205,7	< 0,01
National Bank of Canada	-0,023	-0,001	0,022	7,574	1 293,8	< 0,01
Power Corporation of Canada	-0,028	0,001	0,039	4,643	197,7	< 0,01
Enbridge Inc	-0,034	0,001	0,033	15,556	10 505,2	< 0,01
TransCanada Corporation	-0,032	0,000	0,030	16,942	12 215,5	< 0,01
Fortis Inc.	-0,025	-0,001	0,023	10,591	3 625,4	< 0,01
Power Financial Corp.	-0,033	-0,001	0,022	13,572	7 042,4	< 0,01
ATCO Ltd. Class I	-0,015	0,001	0,043	10,353	3 649,6	< 0,01
ONEX Corporation	-0,034	-0,001	0,017	6,915	1 005,2	< 0,01
TransAlta Corporation	-0,026	-0,001	0,019	10,541	3 532,9	< 0,01
Canadian REIT	-0,033	0,000	0,027	49,333	134 868,8	< 0,01
Thomson Reuters Corporation (USA)	-0,037	0,001	0,043	13,090	6 470,7	< 0,01
Newalta Corporation	-0,036	0,000	0,031	12,078	5 340,9	< 0,01
Interfor Corp	0,008	0,001	0,031	8,780	2 205,0	< 0,01
Great-West Lifeco Inc.	-0,045	0,000	0,044	9,482	2 629,6	< 0,01
Canadian National Railway Company	-0,054	-0,001	0,018	5,112	283,0	< 0,01
Pan American Silver Corp.	-0,011	-0,001	0,021	6,024	574,8	< 0,01

Royal Bank of Canada	-0,037	-0,001	0,018	10,258	3 270,5	< 0,01
Bank of Montreal	-0,043	0,000	0,019	8,849	2 154,5	< 0,01
Norbord Inc	-0,051	-0,001	0,014	6,548	869,9	< 0,01
Sherritt International Corporation	0,000	-0,002	0,032	28,733	41 327,9	< 0,01
Canadian Utilities Limited Classe A	-0,034	-0,002	0,022	6,955	967,2	< 0,01
First Capital Realty Inc	-0,037	-0,002	0,022	7,097	1 120,2	< 0,01
BCE Inc.	-0,040	-0,002	0,017	8,148	1 637,9	< 0,01
RioCan Real Estate Investment Trust	-0,044	0,000	0,028	9,116	2 437,0	< 0,01
Valeant Pharmaceuticals Intl Inc.	-0,050	0,002	0,039	9,569	2 789,9	< 0,01
ARC Resources Ltd	-0,070	-0,001	0,018	11,663	4 637,4	< 0,01
Industrial Alliance Ass. & Serv. Fin.	-0,055	-0,002	0,017	6,367	700,7	< 0,01
Canadian Natural Resources Limited	-0,064	-0,001	0,023	9,006	2 364,1	< 0,01
Precision Drilling Corporation	-0,066	-0,001	0,022	9,777	2 910,9	< 0,01
Magna International Inc.	-0,075	-0,001	0,017	8,299	1 735,3	< 0,01
CCL Industries Inc. Class B	-0,060	-0,001	0,026	6,050	580,2	< 0,01
CGI Technologies and Solutions Inc	-0,034	-0,001	0,017	17,711	13 777,6	< 0,01
Cott Corporation	-0,089	0,000	0,023	11,042	4 014,8	< 0,01
ATS Automation Tooling Systems Inc.	-0,111	-0,001	0,017	11,151	4 127,5	< 0,01
Teck Resources Ltd	-0,139	-0,001	0,026	8,358	1 827,7	< 0,01
Air Canada	-0,189	-0,001	0,019	9,628	2 766,2	< 0,01

## APPENDICE B

### TESTS D'ADÉQUATION STATISTIQUE ET ANALYSES

Tableau B.1 Test de normalité sur les résidus des modèles paramétriques.

Période : 4 janvier 2009 au 29 octobre 2014

Nombre d'observations : 1227

MODÈLE	Équation	Test d'asymétrie			Test du kurtosis			Test univarié de Jarque Béra		Test joint de Jarque-Béra			
		valeur	statistique*	val. P	valeur	statistique*	val. P	statistique	ddl	val. P	statistique	ddl	val. P
1	1	0,485	48,016	< 0,01	5,951	445,152	< 0,01	493,167	2	< 0,01			< 0,01
	2	-0,006	0,008	0,931	4,726	152,359	< 0,01	152,367	2	< 0,01	714,558	6	< 0,01
	3	0,06	0,758	0,384	4,156	68,266	< 0,01	69,024	2	< 0,01			< 0,01
2	1	0,890	161,998	0,000	13,233	5353,011	< 0,01	5515,009	2	< 0,01			< 0,01
	2	0,054	0,589	0,443	13,765	5924,667	< 0,01	5925,256	2	< 0,01	11568,39	6	< 0,01
	3	-0,107	2,333	0,127	4,569	125,795	< 0,01	128,128	2	< 0,01			< 0,01
3	1	0,235	11,244	0,001	3,860	37,762	< 0,01	49,006	2	< 0,01			< 0,01
	2	-0,164	5,525	0,019	4,030	54,241	< 0,01	59,765	2	< 0,01	115,490	6	< 0,01
	3	-0,020	0,086	0,770	3,360	6,633	0,010	6,719	2	0,0348			< 0,01

\* Cette statistique est distribuée selon une loi de chi-deux à 1 degré de liberté.

Tableau B.2 Tests d'hétéroscédasticité sur les modèles paramétriques.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

MODÈLE	Équation	Test de White		Test ARCH(1)		Test ARCH(3)		Test ARCH(5)
		Statistique*	valeur. P	Statistique**	valeur. P	Statistique**	valeur. P	Statistique**
1	1	0,666	0,414	26,488	<0,01	111,141	<0,01	130,064
	2	0,305	0,581	72,665	<0,01	197,435	<0,01	219,091
	3	0,915	0,339	34,420	<0,01	160,818	<0,01	188,799
2	1	5,251	0,022	23,930	<0,01	189,374	<0,01	189,394
	2	88,183	<0,01	98,228	<0,01	198,633	<0,01	224,002
	3	10,213	<0,01	23,077	<0,01	120,633	<0,01	145,761
3	1	0,226	0,635	8,939	<0,01	10,265	0,016	18,326
	2	0,440	0,507	1,513	0,219	4,367	0,225	7,504
	3	3,186	0,074	0,488	0,485	2,103	0,551	3,768

\* Ici la statistique du test de White, basée sur le multiplicateur de Lagrange, est distribuée selon une loi de chi-deux à 1 degré de liberté.

\*\* La statistique du test ARCH (p) suit une loi de chi-deux à p degrés de liberté.

Tableau B.3 Tests d'autocorrélation sur les modèles paramétriques.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

MODÈLE	Équation	Test de Breusch-Godfrey						Test de Ljung-Box	
		à 1 retard		à 3 retards		à 5 retards		7 retards*	
		Statistique*	valeur-P	Statistique	valeur-P	Statistique	valeur-P	Statistique	valeur-P
1	1	11,567	<0,01	12,513	<0,01	14,613	0,012	10,288	0,173
	2	1,921	0,166	9,008	0,029	11,010	0,051	19,786	<0,01
	3	5,310	0,021	8,132	0,043	13,576	0,019	9,299	0,232
2	1	17,094	<0,01	18,136	<0,01	18,748	<0,01	23,827	<0,01
	2	1,427	0,232	2,394	0,495	4,458	0,486	8,936	0,257
	3	10,528	<0,01	18,014	<0,01	28,862	<0,01	2,879	0,896
3	1	9,966	<0,01	9,995	0,019	11,252	0,047	12,158	0,095
	2	1,550	0,213	5,582	0,134	8,158	0,148	5,724	0,572
	3	8,937	<0,01	10,766	0,013	15,264	<0,01	7,426	0,386

\* Ici la statistique du test est distribuée selon une loi de chi-deux à p degrés de liberté où p désigne le nombre de retards.

\*\* Le test est basé sur la statistique de Box-Pierce qui suit une loi de chi-deux à p degrés de liberté où p est le nombre de retards incorporés dans la régression auxiliaire du test. Le nombre optimal de retards est déterminé par selon un critère d'information Hannan et Quinn.

Tableau B.4 Tests de RESET sur les modèles paramétriques.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

MODÈLE	Équation	Test de RESET			
		Puissance* 2 et 3		Puissance 2 à 4	
		Statistique	valeur-p	Statistique**	valeur-p
1	1	1,773	0,170	1,231	0,297
	2	1,910	0,149	1,352	0,256
	3	0,306	0,736	1,692	0,167
2	1	2,516	0,081	1,685	0,168
	2	13,055	< 0,01	18,559	< 0,01
	3	3,488	0,031	2,815	0,038
3	1	0,274	0,761	1,001	0,392
	2	-	-	-	-
	3	3,020	0,049	2,067	0,103

\* Dans la régression auxiliaire utilisée pour la construction du test de RESET, la série prédite de la variable endogène est élevée aux puissances indiquées ( par exemple 2 et 3 ) et ajoutée comme régresseur additionnel. Ici la statistique du test suit une loi de Fisher à 2 et 1223 degrés de liberté. Ici la statistique suit une loi de Fisher à 3 et 1222 degrés de liberté.

\*\*Ici la statistique du test de RESET suit une loi de Fisher à 3 et 1222 degrés de liberté.

Tableau B.5 Test de stabilité de Quandt-Andrews.

Période : 4 janvier 2010 au 29 octobre 2014

Nombre d'observations : 1227

MODÈLE	Équation	Test de Quandt-Andrews **		
		Statistique**	valeur-p	Période
1	1	20,955	< 0,01	01/01/2014
	2	12,836	0,029	10/14/2013
	3	4,079	0,713	01/01/2014
2	1	6,483	0,351	09/26/2013
	2	7,059	0,288	09/27/2013
	3	14,320	0,015	02/28/2011
3	1	6,960	0,298	09/26/2013
	2	4,446	0,651	10/11/2013
	3	6,683	0,328	10/18/2011

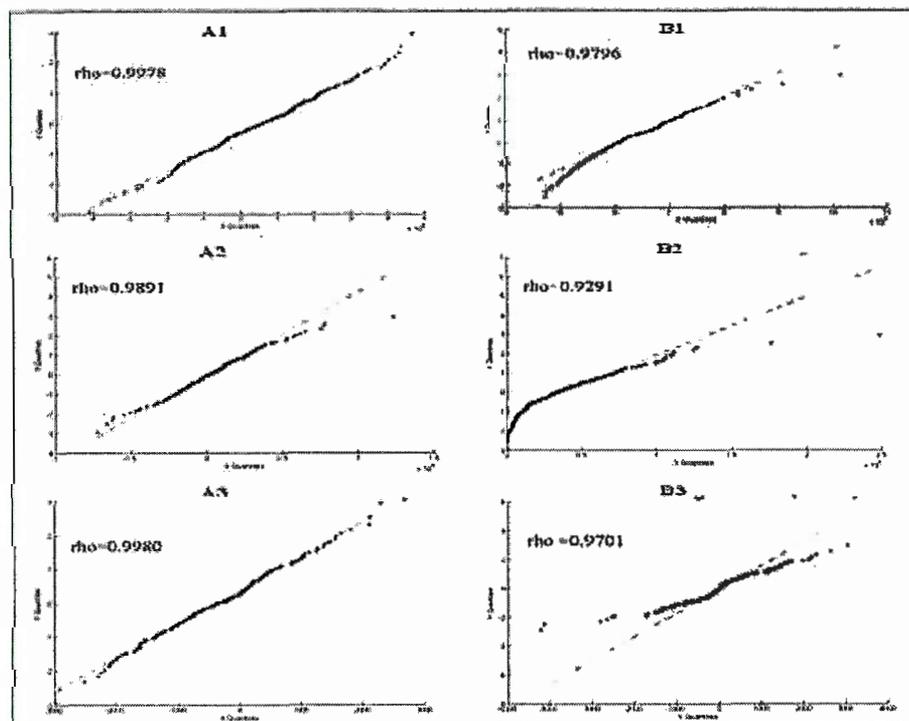
\*\*Ce test est basé sur la valeur maximale de la statistique de Wald atteinte sur la période échantillonnale après suppression de 15 % des observations.

Tableau B.6 Analyse de sensibilité de l'estimation semi-paramétrique.

Nb de portefeuilles*	2	-1,249	11,279
	3	-60,485	-56,828
	5	-97,736	-96,746
Nombre de titres par portefeuille**	8	-60,485	-56,828
	10	-66,873	-61,641
	16	-95,224	-94,204
Nombre d'observations	250	-57,353	-54,917
	1227	-60,485	-56,828

\* Il s'agit du nombre de portefeuilles contenant 8 titres

\*\* Il s'agit du nombre de titres dans chacun des 3 portefeuilles du système



La figure ci-dessus rapporte les résultats de tests de symétrie elliptique appliqués sur un échantillon de 200 observations d'un vecteur de 50 variables aléatoires. Celles de la colonne de gauche proviennent de lois elliptiques symétriques telles que une loi normale multivariée en A1, une t de Student à 5 degrés de liberté en A2 et une combinaison d'échelle de distributions normales de moyennes et variances identiques en A3. Sur la colonne de droite sont des variables à distributions exclues à cette classe. En B1, on a une loi uniforme multivariée définie sur  $a = [-0.002, b = 0.2]$ , en B2 on a une loi de Fischer à  $\nu_1 = 10$  et  $\nu_2 = 2$  degrés de liberté. Finalement en B3, on a une combinaison d'une loi de Cauchy et d'une loi de Student à 5 degrés de liberté. Les valeurs des statistiques robustes ainsi que les coefficients de corrélation rapportés sur chaque figure sont obtenues par des simulations de Monte Carlo avec 10,000 essais.

Figure B.1 Test de Li, Fang et Zhu (1997), simulations.

## BIBLIOGRAPHIE

- Abdul-Hamid, H. Nolan, J.P. 1998. « *Multivariate Stable Densities as Functions of One Dimensional Projections* ». *Journal of Multivariate Analysis*, 67(1): 80-89.
- Abramson, I.S. 1982. « *On Bandwidth Variation in Kernel Estimates-A Square Root Law* ». *The Annals of Statistic*, 10(4): 1217-1223.
- Ahn, S.K. 1992. « *F-Probability Plot and its Application to Multivariate Normality* ». *Communications in Statistical Theory and Methods*, 21(4): 997-1023.
- Andrews, D. 1993. « *Test for Parameter Instability and Structural Change with Unknown Change Point* ». *Econometrica*, 61(4): 821-856
- Arrow, K.:J. 1965. « *Aspects of The Theory of Risk-Bearing* ». Helsinki, Yrjo Jahnsson Saatio.
- Banz, R. 1981. « *The Relationship Between Return and Market Value of Common Stocks* ». *Journal of Financial Economics* 9(1): 3-18.
- Basu, S. 1977. « *The Investment Performance of Common Stocks in Relation to Their Price to Earnings Ratios: A Test of The Efficient Market Hypothesis* ». *The Journal of Finance* 32(3): 663-682.
- Bawa, V.S. 1975. « *Optimal Rules for Ordering Uncertain Prospects* ». *Journal of Financial Economics* 2(1): 95-121.
- Beaulieu, M.-C., Dufour, J.-M. et Khalaf, L. 2002. « *Testing Black's CAPM with Possibly Non-Gaussian Errors: An Exact Identification-Robust Simulation-Based Approach* ». Technical Report, CIRANO et CIREQ, Université de Montréal et Département d'Économique, Université Laval.
- Bera, A. K. and Jarque, C. M. 1981. « *An Efficient Large-Sample Test for Normality of Observations and Regression Residuals* », Australian National University Working Papers in Econometrics 40, Canberra
- Béran, R. 1974. « *Asymptotically Efficient Adaptive Rank Estimates in Location Models* », *Annals of Statistics*, 2(1): 63-74
- Béran, R. 1979. « *Testing for Ellipsoidal Symmetry of A Multivariate Density* ». *Annals of Statistics* 7(1):150-162.
- Beck, J. B. 1997. « *Necessary Conditions for The CAPM* ». *Journal of Economic Theory*, 73(1): 245-257

- Bhandari, L. C. 1988. « *Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence* ». The Journal of Finance 43(2): 507-528.
- Bickel, P.J. 1982. « *On Adaptive Estimation* ». The annals of statistics, 10(3), 647-671.
- Bingham, N.H, et Kiesel, R. 2001. *Modelling Asset Return with Hyperbolic Distributions*, p. 1-20, de Return Distributions in Finance. (ed. J. Knight and Satchell), Butterworth-Heinemann.
- Black, F. 1972. « *Capital Market Equilibrium with Restricted Borrowing* ». The Journal of Business 45(3): 444-455.
- Black, F., Jensen, M. et Scholes, M. 1972. « *The Capital Asset Pricing Model: Some Empirical Tests* ». Studies in the Theory of Capital Markets, M. Jensen, Paeger, NY.
- Blattberg, R.C., Gonedes, N.J. 1974 « *A Comparison of the Stable and Student Distributions as Statistical Models* ». The Journal of Business 47(2): 244-280.
- Blume, M., and Friend I., 1970. « *Measurement of Portfolio Performance Under Uncertainty* ». American Economic Review, 1970, 60(4): 561-575.
- Bodnar, T. et Schmid, W. 2008. « *The Distribution of The Sample Variance of The Global Minimum Variance Portfolio in Elliptical Models* ». Metrica, 67(2): 127-143.
- Bollerslev, T. 1986. « *Generalized Autoregressive Conditional Heteroskedasticity* ». Journal of Econometrics, 31(3): 307-327.
- Bollerslev, T. et Wooldridge, J.M. 1992. « *Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances* ». Econometric Review, 11(2): 143-172
- Box, G. et Cox, D. 1964. « *An Analysis of Transformations* ». Journal of the Royal Statistical Society, Series B: 211-264.
- Breiman, L., Meisel, W., Purcell, E., 1977. « *Variable Kernel Estimates of Multivariate Densities* ». Technometrics 19 (2): 135-144.
- Brennan, M. J. 1970. « *Taxes, Market Valuation and Corporate Financial Policy* ». National Tax Journal, 23: 417-427.
- Brennan, M. J. 1971. « *Capital Market Equilibrium with Divergent Borrowing and Lending Rates* ». Journal of Financial and Quantitative Analysis, 6(5): 1997-1205.
- Breeden, D.T. 1979. « *An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities* ». Journal of Financial Economics 7(3): 265-296.

- Breusch, T. 1978, « *Testing for Autocorrelation in Dynamic Linear Models* », Australian Economic Papers, 17: 334–355.
- Byczkowski, T., Nolan, J.P. et Rajput, B. 1993 « *Approximation of Multidimensional Stable Densities* ». Journal of Multivariate Analysis 46:13-31.
- Campbell, J.Y., Lo, A.W. et MacKinlay, A.C. 1997. *The Econometrics of Financial Markets*. Princeton University Press
- Chamberlain, G. 1983. « *A Characterization of The Distributions that Imply Mean-Variance Utility Functions* ». Journal of Economic Theory, 29: 185-201.
- Casella, G. et Berger, R.L. 1990. *Statistical Inference*. Belmont. CA. Duxbury Press.
- Chen S. X. 2000. « *Probability Density Function Estimation Using Gamma Kernels* », Annals of the Institute of Statistical Mathematics, 52(3): 471-480
- Chicken, E. et Cai, T. 2005. « *Block Thresholding for Density Estimation: Local and Global Adaptivity* ». Journal of Multivariate Analysis, 95, 76–106 MR2164124
- Chu, K. C. 1973. « *Estimation and Decision for Linear Systems with Elliptical Random Processes* », IEEE Trans. Automat. Control, 18: 499—505.
- Clark, P. K. 1973, « *A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices* ». Econometrica 41: 135-156.
- Cochrane J. H. 1991, « *Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations* ». The Journal of Finance, 46(1): 209-237
- Cont, R. 2001. « *Empirical Properties of Asset Returns Stylized Facts and Statistical Issues* ». Quantitative Finance 1: 223-236.
- Cox, D.R. et Small, N.J.H 1978. « *Testing Multivariate Normality* », Biometrika, 65(2): 263-272.
- Davidson, R., MacKinnon, J., 2004. *Econometric Theory and Methods*. Oxford University Press.
- Doornik, J. A. and Hansen, H. 1994. « *An Omnibus Test for Univariate and Multivariate Normality* », Economics Working Papers W4 & 91, Nuffield College.
- Eaton, M. L., 1981. « *On the Projection of Isotropic Distributions* », Ann. of Statist. 9: 391-400.
- Eberlein, E. 2001. « *Applications of Generalized Hyperbolic Lévy Motions to Finance* ». In Barndorff-Nielsen, O.E. et al. (eds.), Lévy Processes. Theory and Applications,

- Birkhäuser, 319-336.
- Eberlein, E. et v. Hammerstein, E.A. 2004. « *Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes* », Seminar on Stochastic Analysis, Random Fields and Applications IV Progress in Probability 58: 221-264.
- Eberlein, E. et Keller, U. 1995. « *Hyperbolic Distributions in Finance* », Bernoulli, 1(3): 281-299.
- Elton, E.J. et Gruber, M.J. 1997. *Modern Portfolio Theory and Investment Analysis*. Wiley, 5ième édition.
- Engle, R.F. 1982. « *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation* ». *Econometrica*, 50: 987-1007.
- Eye, A. 2005. « *Comparing Tests of Multinormality - A Monte Carlo Study* ». InterStat, <http://interstat.statjournals.net/> (October, 2005).
- Fama, E. 1965. « *The Behavior of Stock Market Prices* ». *Journal of Business*, 38: 34-105.
- Fama, E. et French, K. R. 1992. « *The Cross-Section of Expected Stock Returns* ». *The Journal of Finance*, 47(2): 427.
- Fama, E. et French, K. R. 1993. « *Common Risk Factors in the Returns on Stocks and Bonds* ». *Journal of Financial Economics*, 33: 3-56.
- Fama, E., et J. MacBeth. 1973. « *Risk, Return, and Equilibrium: Empirical Tests* ». *Journal of Political Economy*, 71: 607-636.
- Fama, E., French, K. 2004, « *The Capital Asset Pricing Model: Theory and Evidence* », *Journal of Economic Perspectives*, 18(3): 25-48.
- Fan, K.-T., Kotz, S. et Ng, K.-W. 1990. *Symmetric Multivariate and Related Distributions*, Chapman & Hall, New York.
- Faraway, J. J. 1992. « *Smoothing in Adaptive Estimation* », *The Annals of Statistics*, 20(1): 414-427.
- Fielitz, B. D., et Rozelle, J.P. 1983. « *Stable Distributions and the Mixtures of Distributions Hypotheses for Common Stock Returns* », *Journal of American Statistical Association*, 78(381): 28-36.
- Friend, I., Landskroner, Y. et Losq, E. 1976. « *The Demand for Risky Assets Under Uncertain Inflation* ». *The Journal of Finance*, 31: 1287-1297.
- Gibbons, M. R., Ross, S.-A. et Shanken, J. 1989. « *A Test of The Efficiency of A Given*

- Portfolio* ». *Econometrica*, 57: 1121-1152.
- Godfrey, L. 1978. « *Testing Against General Autoregressive and Moving Average Error Models when the Regressors Include Lagged Dependent Variables* », *Econometrica*, 46: 1293–1302.
- Good, I. J. et Gaskins, R. A. 1971. « *Nonparametric Roughness Penalties for Probability Densities* ». *Biometrika*, 58(2): 255-277.
- Gray, B.J. et French, D.W. 1990. « *Empirical Comparisons of Distributional Models for Stock Index Returns* ». *Journal of Business Finance & Accounting* 17: 451-459.
- Gupta, A. K. et Varga, T. 1993. « *Elliptically Contoured Models in Statistics* », Kluwer Academic Publishers, Dordrecht.
- Hájek, J. 1962. « *Asymptotically Most Powerful Rank Order Tests* », *Annals of Mathematical Statistics*. 33: 1124-1147
- Hájek, J. 1972. « *Local Asymptotic Normality Minimax and Inadmissibility in Estimation* », *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1: 175 – 194.
- Hannan, E.J. et Quinn, B.G. 1979. « *The Determination of the Order of an Autoregression* ». *Journal of the Royal Statistical Society, Series B*, 41(2) 190-195.
- Härdle, W. et Linton, O.B. 1994. « *Applied Nonparametric Methods* ». In D.F. McFadden and R.F. Engle III (eds.), *The Handbook of Econometrics*, 4: 2295-2339, North Holland.
- Hearly, M.J.R. 1968. « *Multivariate Normal Plotting* », *Journal of Applied Statistics*, 17: 157-161.
- Hodgson DJ. 1998a. « *Adaptive Estimation of Cointegrating Regressions with ARMA Errors* ». *Journal of Econometrics* 85: 231–268.
- Hodgson, D., Linton, O. et Vorkinck, K. 2002. « *Testing the Capital Asset Pricing Model under Elliptical Symmetry : A Semiparametric Approach* ». *Journal of Applied Econometrics*, 17(6): 617-639.
- Huffer, F.W. Park, C. 2002. « *The Limiting Distribution of A Test for Multivariate Structure* », *Journal of Multivariate Statistic. Plann. Inference* 105: 417–431.
- Ingersoll, J. 1987. *Theory of Financial Decision Making*. Rowman & Littlefield.
- Jobson, J. et Korkie, R. 1982. « *Potential Performance and Tests of Portfolio Efficiency* ».

- Journal of Financial Economics, 10, 433-466.
- Kelker, D. 1970. « *Distribution Theory of Spherical Distributions and Location-Scale Parameter Generalization* ». Sankhya, 32A: 419-430.
- Kraus, A. et Litzenberger, R.H. 1976. « *Skewness Preference and the Valuation of Risk Assets* ». The Journal of Finance, 31: 1085-1100.
- Landsman, Z. et Nešlehová, J. 2008. « *Stein's Lemma for Elliptical Random Vectors* », Journal of Multivariate Analysis, 99: 912 – 927.
- Le Cam, L. 1969. « *Théorie Asymptotique de la Décision Statistique* », Les Presses de l'Université de Montréal.
- Lindenberg, E. 1979. « *Capital Market Equilibrium with Price Affecting Institutional Investors* » dans Elton Gruber Portfolio Theory 25 Years Later, Amsterdam, North-Holland
- Lintner, J. 1965. « *The valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets* ». Review of Economics and Statistics, 47(1): 13-37.
- Lintner, J. 1969. « *The Aggregation of Investor's Diverse Judgments and Preferences in Purely Competitive Markets* ». Journal of Financial and Quantitative Analysis, December 1969: 347-400.
- Ljung, G. M. and Box, G. E. P. 1978. « *On a Measure of Lack of Fit in Time Series Models* », Biometrika, 65(2), 297-303.
- Loftsgaarden, D.O. et Quesenberry, C.P. 1965. « *A Nonparametric Estimate of a Multivariate Density Function* ». Annals of Mathematical Statistics 36(3): 1049-1051.
- Lucas R. Jr, 1978. « *Asset Prices in an Exchange Economy* ». Econometrica, 46(6):1429- 45.
- MacKinley, A. C. 1985. *An Analysis of Multivariate Financial Tests*. Ph.D. Dissertation, Graduate School of Business, University of Chicago, Chicago, IL.
- MacKinlay, A.C. et Richardson, M. P. 1991. « *Using Generalized of Moments to Test Mean-Variance Efficiency* ». The Journal of Finance, 46: 511-527.
- Mandelbrot, B. 1963. « *The Variation of Certain Speculative Prices* ». Journal of Business, 36: 394-149.
- Manzotti, A., Pérez, F.J. et Quiroz, A.J. 2002. « *A Statistic for Testing the Null Hypothesis of Elliptical Symmetry* », Journal of Multivariate Analysis. 81 (2002) 274–285.
- Markovitz, H. 1959. « *Portfolio Selection: Efficient Diversification of Investments* ». John

Wiley & Sons, Inc., New York, Chapman & Hall, Limited, London.

- Mardia, K.V. 1970. « *Measures of Multivariate Skewness and Kurtosis with Applications* ». *Biometrika*, 57 (3): 519-530.
- Markowitz, H. 1952. « *Portfolio Selection* ». *The Journal of Finance*, 7: 77-91.
- Markowitz, H. 1991. « *Foundations of Portfolio Theory* ». *The Journal of Finance*, 46: 469-77.
- Mayers, D. 1972. « *Nonmarketable Assets and Capital Market Equilibrium Under Uncertainty* ». *Studies in the Theory of Capital Markets*. Michael C. Jensen New York: Praeger: 223-248.
- Merton, R.C. 1973. « *An Intertemporal Capital Asset Pricing Model* ». *Econometrica*, 41: 867-887.
- Mossin, J. 1966. « *Equilibrium in A Capital Asset Market* ». *Econometrica*, 34: 768-783.
- Muirhead, R. 1983. « *Aspects of Multivariate Statistical Theory* ». John Wiley and Sons, New York.
- Nadaraya, E.A. 1964. « *On Estimating Regression* ». *Theory of Probability and its Applications*, 10: 186-190.
- Newey, W. K. 2004 . « *Efficient Semiparametric Estimation via Moment Restrictions* », *Econometrica*, 72(6): 1877-1897.
- Newey, W. K. and West, K. D. 1987. « *A Simple Positive-Definite Heteroskedasticity and Autocorrelation-Consistent Covariance Matrix* », *Econometrica* 55: 703-708.
- Osborne, M. F. M. 1959. « *Brownian Motion in the Stock Market* ». *Operations Research*, 7(2): 145-173.
- Owen, J. et Rabinovitch, R. 1983. « *On the Class of Elliptical Distributions and Their Applications to The Theory of Portfolio Choice* ». *The Journal of Finance*, 38: 745-752.
- Pagan, A. et Ullah, A. 1999. *Nonparametric Econometrics*. Cambridge University Press
- Parzen, E. 1962. « *On Estimation of a Probability Density Function and Mode* ». *The Annals of Mathematical Statistics*, 33: 1065-1076.
- Pratt, J. W. 1964. « *Risk Aversion in the Small and in the Large* ». *Econometrica*, 32: 122-136.
- Prause, K. 1997. *Modelling Financial Data Using Generalized Hyperbolic Distributions*. FDM Preprint 48, University of Freiburg.

- Prause, K. 2000. *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, Ph.D. Dissertation, University of Freiburg.
- Quandt, R. 1960. «*Tests of the Hypothesis that a Linear Regression System Obeys Two Different Regimes*», *Journal of the American Statistical Association* 55: 324-330.
- Rachev, S.T. et Mittnik, S. 2000. «*Stable Paretian Models in Finance* ». Wiley, New York.
- Ramsey, J. B. 1969. «*Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis* », *Journal of the Royal Statistical Society B*, 31(2): 350-371.
- Roll, R., 1969. «*Bias in Fitting the Sharpe Model to Time Series Data* ». *Journal of Financial and Quantitative Analysis*, 4: 271-89.
- Romano, J. 1988. «*A Bootstrap Revival of Some Nonparametric Distance Test* ». *Journal of the American Statistical Association*, 83: 698-708.
- Rosenblatt, M. 1956. «*Remarks on Some Nonparametric Estimates of Density Function* ». *The Annals of Mathematical Statistics*, 27: 832-837.
- Ross, S. A. 1978. «*Mutual Fund Separation in Financial Theory: The Separating Distributions* ». *Journal of Economic Theory*, 17: 254-286.
- Rothenberg, T.J. et Leenders, C.T. 1964. «*Efficient Estimation of Simultaneous Equations Systems*». *Econometrica*, 32: 57-76.
- Rubinstein, M. 1976. «*The Valuation of Uncertain Income Stream and the Pricing of Options* ». *The Bell Journal of Economics*, 7: 407-425.
- Sentana, E., 2008 «*The Econometrics of Mean-Variance Efficiency Tests : A Survey* ». CEMFI working papers, Working Paper No. 0807.
- Sentana, E., et Amengual, D. 2008. «*A Comparison of Mean-Variance Efficiency Tests* » CEMFI, Working Paper No. 0806.
- Shanken, J., 1987 «*Multivariate Proxies and Asset Pricing Relations : Living with the Roll Critique*». *The Journal of Financial Economics*, 18: 91-110.
- Sharpe, W.F. 1964. «*Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk*». *The Journal of Finance*, 19: 425-442.
- Schuster E.F. 1985. «*Incorporating Support Constraints into Nonparametric Estimators of Densities*». *Communications in Statistics—Theory and Methods* 14: 1123–1136.
- Silverman, B.W. 1986. *Density Estimation for Statistics and Data Analysis*. London: à

Chapman & Hall.

- Stambaugh, R.F. 1982. « *On the Exclusion of Assets from Tests of The Two-Parameter Model: A Sensitivity Analysis* ». *Journal of Financial Economics*, 10: 237-268.
- Stein, C. 1956. « *Efficient Nonparametric Testing and Estimation* », dans J. Neyman ed. *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, University of California Press, Berkeley, CA, 187-195.
- Tobin, J. 1958. « *Liquidity Preference as Behavior Towards Risk* », *Review of Economic Studies*, 25: 65-86.
- Treynor, J. L. 1962. « *Toward a Theory of Market Value of Risky Assets* », Manuscrit non publié datant d'automne 1962 par Mr. Treynor. Une version finale a été publiée en 1999 dans *Asset Pricing and Portfolio Performance* de l'éditeur Robert A. Korajczyk London : Risk Books.
- Vasicek, O. 1971. « *Capital Asset Pricing Model with No Riskless Borrowing* ». Wells Fargo Bank.
- Von Neumann, A. et Morgenstern, O. 1953. « *Theory of Games and Economic behaviour* ». Princeton University Press.
- Watson, G.S. 1964. « *Smooth Regression Analysis* ». *Shankya, Série A*, 26: 359 -372.
- White, H. 1982. « *Maximum Likelihood Estimation of Misspecified Models* ». *Econometrica*, 50(1): 1 -25.
- Zellner, A. 1962. « *An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias* ». *Journal of the American Statistical*, 57(298): 348- 368.
- Zhou, G. 1993. « *Asset Pricing Tests under Alternative Distributions* ». *The Journal of Finance*, 48: 1927-1942.