HEC MONTRÉAL

Calcul de VaR et CVaR: doit-on adopter une modélisation univariée ou multivariée?

par

Alain-Philippe Fortin

Sciences de la gestion
(Option économie financière appliquée)

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

Mai 2017

© Alain-Philippe Fortin, 2017

Résumé

Ce mémoire compare une modélisation multivariée à une modélisation univariée pour la prévision de VaRs et CVaRs de portefeuille sur une semaine. Nous effectuons une extension de l'étude de Christoffersen et Langlois (2013) en utilisant leurs données et leurs modèles. Au total, deux modèles univariés sont comparés à huit modèles multivariés impliquant des copules et des corrélations dynamiques. La précision des différents modèles hors échantillon de 1983 à 2010 est comparée en utilisant le test CPA de Giacomini et White (2006). Nous trouvons qu'il n'y a pas de différences significatives entre la précision des modèles multivariés avec corrélations dynamiques et les modèles univariés. À l'inverse, les modèles univariés performent mieux que les modèles multivariés sans corrélations dynamiques.

Abstract

This thesis compares a multivariate approach to a univariate approach for forecasting one week portfolio VaRs and CVaRs. We do an extension of the study of Christoffersen and Langlois (2013) by borrowing their data and models. In total, two univariate models are compared to eight multivariate models involving copulas and dynamic correlations. The forecasting precision of the different models is compared out of sample from 1982 to 2010 using the CPA test of Giacomini and White (2006). We find no significant differences in forecasting precision between multivariate models with dynamic correlations and univariate models. Conversely, we find that univariate models outperform multivariate models without dynamic correlations.

Table des matières

Résumé **i**

Abstract **iii**

Liste des tableaux **vii**

Liste des figures **ix**

Introduction **1**

1 La mesure du risque **7**

1.1 La VaR 9

1.2 La CVaR 10

1.3 La VaR et la CVaR par simulation 11

2 La modélisation univariée du risque **15**

2.1 Un modèle 15

2.2 Estimation du modèle 16

2.3 Simulation du modèle 18

3 La modélisation multivariée du risque **19**

3.1 L'approche multivariée 19

3.2 Les copules 21

3.3 Estimation avec une copule 22

3.4 Simulation avec une copule 22

4 Le *backtesting* **25**

4.1 Le *backtest* de la VaR 25

4.1.1 Test de couverture non conditionnelle 27

4.1.2 Test d'indépendance 28

4.1.3 Test de couverture conditionnelle 29

4.2 Le *backtest* de la CVaR 30

4.2.1 Test 1 30

4.2.2 Test 2 32

5 L'élicitabilité **35**

5.1 Sélection de modèle de VaR 36

5.2 Sélection de modèle de VaR/CVaR 37

6 Analyse empirique **39**

6.1 Rendements des facteurs 40

6.2 Dynamique des facteurs 42

6.3 Dépendance des facteurs 44

6.4 Dynamique de portefeuille 47

6.5 Calculs de VaRs et CVaRs 47

6.6 Le *backtesting* 49

6.7 Comparaison des modèles 51

Conclusion **55**

Bibliographie **57**

Annexe A – Copule t asymétrique de Demarta et McNeil (2005) **i**

Annexe B – Standardisation des fractiles et ciblage de corrélation dans le DCC **v**

Annexe C – Fonction MATLAB pour calculer la log-vraisemblance de la copule t asymétrique dynamique **vii**

Liste des tableaux

1 Statistiques descriptives des rendements hebdomadaires des facteurs, 1963-2010 xxv

2 Modèles AR-GARCH univariés des rendements de portefeuille, 1963-2010 . xxvi

3 Modèles AR-GARCH univariés des rendements des facteurs, 1963-2010 . . xxvii

4 Estimation des modèles de dépendance entre les facteurs, 1963-2010 xxviii

5 *Backtest* des VaRs et CVaRs des modèles, 1983-2010 xxix

6 Comparaison de la moyenne des scores des modèles, 1983-2010 xxx

7 Test CPA de Giacomini et White (2006) des VaRs 1% pour chaque paire de modèles, 1983-2010 xxxi

8 Test CPA de Giacomini et White (2006) des VaRs 5% pour chaque paire de modèles, 1983-2010 xxxii

9 Test CPA de Giacomini et White (2006) des VaRs/CVaRs 1% pour chaque paire de modèles, 1983-2010 xxxiii

10 Test CPA de Giacomini et White (2006) des VaRs/CVaRs 5% pour chaque paire de modèles, 1983-2010 xxxiv

Liste des figures

1 Séries chronologiques des rendements des facteurs xii

2 Diagrammes Quantile-Quantile des rendements des facteurs, 1963-2010 . . . xiii

3 Diagrammes de dispersion et corrélations de seuil pour les paires de facteurs incluant le facteur *Market* xiv

4 Diagrammes de dispersion et corrélations de seuil pour les paires de facteurs excluant le facteur *Market* xv

5 Autocorrélations des rendements et des rendements absolus xvi

6 Autocorrélations des résidus et des résidus absolus xvii

7 Diagrammes Quantile-Quantile pour les résidus du modèle AR(3)-NGARCH(1,1) avec la distribution de Hansen xviii

8 Corrélations dynamiques de la copule *t* asymétrique, 2006-2010 xix

9 Prévisions de CVaRs hors échantillon pour un portefeuille équipondéré des quatre facteurs xx

10 Dépassements de VaRs 1% pour les modèles non dynamiques, 1983-2010 . . xxi

11 Dépassements de VaRs 1% pour les modèles dynamiques, 1983-2010 xxii

12 Dépassements de VaRs 5% pour les modèles non dynamiques, 1983-2010 . . xxiii

13 Dépassements de VaRs 5% pour les modèles dynamiques, 1983-2010 xxiv

Introduction

La Valeur à Risque (VaR) est la mesure de risque la plus populaire en finance [1]. Toutefois, la récente crise financière a soulevé plusieurs questions sur la gestion des risques financiers et la pertinence de la VaR pour mesurer l'exposition au risque. La Valeur à Risque Conditionnelle (CVaR), aussi appelée *Expected Shortfall*, est une mesure alternative qui a gagné en popularité au cours des dernières années. Contrairement à la VaR, elle est une mesure de risque cohérente [2] qui prend en compte l'ensemble des pertes dans la queue de gauche de la distribution.

Lorsqu'on calcule la VaR ou la CVaR d'un portefeuille, la première décision à prendre concerne la dimension de la modélisation à adopter. En d'autres mots, doit-on modéliser la distribution multivariée des rendements de tous les titres du portefeuille ou la distribution univariée des rendements du portefeuille ? Par exemple, considérons un portefeuille de deux actifs dont les rendements sont observés sur un certain horizon de temps. On peut choisir de modéliser la distribution jointe du vecteur de rendements des actifs à l'aide de ces observations. On doit alors caractériser la distribution marginale du rendement de chaque actif ainsi que les dépendances entre eux. Les VaRs ou CVaRs peuvent ensuite être calculées à l'aide des poids de portefeuille et la distribution multivariée des rendements. L'alternative consiste plutôt à calculer une série univariée de pseudo rendements de portefeuille en appliquant les poids actuels des actifs sur leurs séries de rendements historiques. Avec cette série chronologique univariée construite, nous pouvons modéliser

1. Voir Duffie et Pan (1997) pour un résumé des différentes applications de la VaR en finance.
2. Une mesure de risque est dite cohérente si elle respecte les propriétés d'homogénéité positive, de sous-additivé, de monotonicité, et d'invariance par translation. Voir Artzner et collab. (1999).

la distribution univariée des pseudo rendements de portefeuille pour ensuite calculer les VaRs et CVaRs du portefeuille actuel.

Ce mémoire vise à comparer ces deux approches afin de déterminer si l'une se démarque significativement de l'autre dans la prévision de VaRs et de CVaRs de portefeuille.

Cette question est importante, car la VaR et la CVaR sont les deux mesures de risque les plus utilisées par les institutions financières afin d'évaluer le risque de marché. Elles permettent de contrôler le niveau acceptable de risque et servent à établir le capital règlementaire pour couvrir les pertes potentielles. Nous voulons donc les calculer le plus précisément possible.

Il n'est pas clair que l'approche multivariée soit préférable à l'approche univariée (ou vice-versa) pour la prévision de ces mesures de risque. D'un côté, l'approche univariée a l'avantage d'être relativement simple pour la modélisation et l'estimation. Ceci est un point important pour les institutions financières qui ont besoin de méthodes flexibles et rapides afin de mesurer le risque de manière routinière. Cette approche est toutefois conditionnelle aux poids de portefeuille de sorte qu'à chaque fois que le portefeuille est rebalancé le modèle doit être réestimé. De l'autre côté, l'approche multivariée utilise davantage d'information dans sa modélisation en prenant directement en compte les dépendances entre les actifs du portefeuille. Toutes choses étant égales par ailleurs, l'utilisation d'information supplémentaire par rapport à la première approche devrait permettre une meilleure prévision du risque. L'approche multivariée peut également être utilisée pour l'optimisation de portefeuille étant donné qu'elle n'est pas conditionnelle aux poids d'actifs. Les désavantages de cette approche sont plutôt d'ordre pratique. D'une part, à mesure que la dimension du portefeuille augmente, le grand nombre de paramètres des modèles multivariés alourdit et complique l'estimation pouvant ainsi compromettre la précision des résultats. D'autre part, le temps et les ressources nécessaires pour la modélisation et les calculs limitent la flexibilitéde cette approche.

La question a été directement abordée dans Santos et collab. (2013) pour la prévision de VaRs et des portefeuilles largement diversifiés. Ils comparent une collection de modèles univariés et multivariés à l'aide du test statistique CPA (*Comparative Predictive*

Ability) de Giacomini et White (2006). Leur conclusion est que, bien que dans certains cas il n'y a pas de différence significative entre les deux approches, lorsqu'il y a une différence entre les deux elle est en faveur de l'approche multivariée. À l'inverse, Christoffersen (2009) compare une modélisation FHS [3] univariée et multivariée et soutient que l'approche univariée est suffisante pour la mesure du risque tandis que l'approche multivariée serait seulement nécessaire pour l'optimisation de portefeuille où l'on doit évaluer l'effet d'un changement des poids d'actifs sur le risque. Ultimement, ce problème est une question avantage-coût. Autrement dit, est-ce que la précision gagnée en utilisant un modèle multivarié, si elle existe, fait plus que compenser les efforts supplémentaires nécessaires à son implémentation par rapport à un modèle univarié ? La réponse de Brooks et Persand (2003) lorsqu'ils comparent des modèles GARCH univariés et multivariés pour la prévision de volatilité est négative.

Ce mémoire contribue à la littérature en comparant ces deux approches à la fois pour la VaR et la CVaR. Il se distingue également sur plusieurs autres points. Premièrement, notre analyse ne se limite pas aux résultats du *backtest* [4]. Nous utilisons le test CPA où les fonctions de score sont basées sur les propriétés d'élicitabilité de la VaR et d'élicitabilité jointe du couple VaR/CVaR. Deuxièmement, notre application utilise les trois facteurs de Fama et French (1993) et le facteur de Carhart (1997), des facteurs largement utilisés dans la littérature et par les intervenants en finance. Troisièmement, nous employons des modèles plus sophistiqués de la distribution conditionnelle jointe des rendements impliquant des copules et des corrélations dynamiques.

Les modèles et les données utilisés proviennent de l'étude de Christoffersen et Langlois (2013). Dans cet article les auteurs modélisent la dynamique de la distribution jointe des trois facteurs de Fama et French (1993) et du facteur de Carhart (1997). Ils montrent que la prise en compte des dépendances non linéaires et dynamiques entre les facteurs

3. L'approche FHS ou *Filtered Historical Simulation* est une méthode semi-paramétrique qui utilise la distribution empirique des résidus standardisés plutôt qu'une distribution spécifique pour le calcul des mesures de risque.

4. Le *backtest* est l'utilisation de tests statistiques afin de vérifier la validité des prédictions de VaR/CVaR du modèle. Ces tests ne peuvent cependant pas servir à comparer les modèles entre eux.

entraîne des avantages économiques significatifs mesurés par les gains d'utilité d'un investisseur ayant des préférences CRRA. Ils montrent également que les CVaRs hebdomadaires calculées avec leurs modèles dépassent de 20% à 50% celles calculées en supposant une distribution normale multivariée. Dans ce mémoire nous complétons leur article en étudiant une question différente, c'est-à-dire déterminer s'il y a une préférence entre une modélisation multivariée et univariée pour la prévision de VaRs et de CVaRs de portefeuille.

Le mémoire est structuré comme suit :

— Chapitre 2 — Ce chapitre présente l'approche probabiliste employée dans le texte pour modéliser le risque. Il comprend une description des mesures de risque VaR et CVaR. Il montre comment nous pouvons calculer ces deux mesures de risque par simulation selon l'approche univariée et multivariée.

— Chapitre 3 — Ce chapitre présente le cadre utilisé pour l'approche univariée dans notre étude. Il inclut des sections sur l'estimation et la simulation du modèle.

— Chapitre 4 — Ce chapitre présente les problèmes associés à l'extension multivariée du cadre du chapitre 3. Il présente les copules comme alternative flexible et explique comment estimer et simuler une copule.

— Chapitre 5 — Ce chapitre explique comment effectuer le *backtest* de nos modèles. Il comprend les tests les plus utilisés pour valider les modèles de VaRs. Il présente aussi deux tests pour valider des modèles de CVaRs.

— Chapitre 6 — Ce chapitre présente la propriété d'élicitabilité d'une mesure de risque. Il explique comment nous pouvons exploiter cette propriété pour comparer la précision de différents modèles de risque.

— Chapitre 7 — Ce chapitre présente l'application empirique du mémoire où nous comparons les modèles univariés et multivariés employés dans Christoffersen et Langlois (2013).

— Chapitre 8 — Ce chapitre résume les conclusions de notre étude et explique la divergence de nos résultats avec ceux de Santos et collab. (2013).

Chapitre 1

La mesure du risque

La mesure du risque est un objectif central de la gestion des risques. Mathématiquement, une mesure de risque est une fonction qui assigne à une variable aléatoire un nombre réel. Idéalement, cette mesure doit bien quantifier le risque en plus d'être facile à comprendre et à interpréter. Les deux mesures de risque les plus populaires actuellement sont la Valeur à Risque (VaR) et la Valeur à Risque Conditionnelle (CVaR), communément appelée *Expected Shortfall* [1]. Elles seront présentées dans les prochaines sections.

Une variable aléatoire continue X est décrite par sa distribution de probabilité. Une fois cette distribution de probabilité connue, des mesures de risque peuvent être calculées. C'est pourquoi la majorité du travail dans le calcul d'une mesure de risque consiste à modéliser adéquatement la distribution de la variable aléatoire X. Il s'agit ici d'une distribution univariée puisque la variable aléatoire est unique. Une distribution de probabilité est caractérisée par la fonction de répartition

$$F(x) = \Pr(X \leq x).$$

Elle représente la probabilité que la variable aléatoire X soit inférieure ou égale à un certain nombre x. La fonction de densité de la variable aléatoire continue X est donnée

1. Certains auteurs font une distinction entre la *Conditionnal Value at Risk* et l'*Expected Shortfall* mais pour des distributions continues leurs définitions se confondent.

par

$$f(x) = \frac{dF(x)}{dx}.$$

La fonction de densité permet de calculer la probabilité que la variable aléatoire X soit dans l'intervalle $[a,b]$ avec

$$\Pr(a \leq X \leq b) = \int_a^b f(x)dx.$$

Nous pouvons également considérer un vecteur de variables aléatoires continues

$$X = (X_1, X_2, \ldots, X_N)'.$$

Ce vecteur aléatoire est décrit par une distribution multivariée. Cette distribution contient l'information sur les distributions marginales des variables ainsi que sur les dépendances entre ces variables. Une distribution multivariée est caractérisée par la fonction de répartition multivariée

$$F(x_1, x_2, \ldots, x_N) = \Pr(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_N \leq x_N).$$

La fonction de densité multivariée du vecteur aléatoire est donnée par

$$f(x_1, x_2, \ldots, x_N) = \frac{\partial^N F(x_1, x_2, \ldots, x_N)}{\partial x_1 \partial x_2 \ldots \partial x_N}.$$

La fonction de densité multivariée permet de calculer la probabilité que les variables aléatoires continues du vecteur soient dans des intervalles donnés . Par exemple, pour deux variables nous avons

$$\Pr(a \leq X_1 \leq b, c \leq X_2 \leq d) = \int_c^d \int_a^b f(x_1, x_2) dx_1 dx_2.$$

La variable aléatoire X peut représenter n'importe quelle position risquée dont la valeur future est incertaine, mais dans le cadre de ce mémoire nous nous intéressons aux rendements d'un portefeuille d'actifs financiers au travers du temps r_t, c'est-à-dire au risque de marché. De même, le vecteur aléatoire $r_t = (r_{1t}, r_{2t}, \ldots, r_{Nt})'$ représentera les rendements des N actifs du portefeuille au temps t. Toutes les variables aléatoires et les distributions qui seront présentées sont continues.

On note habituellement une variable aléatoire par une lettre majuscule et sa réalisation par une lettre minuscule, mais cette convention ne sera pas suivie rigoureusement dans le texte afin d'alléger la notation. Ainsi, le rendement r_t représentera parfois une variable aléatoire et parfois une valeur spécifique.

1.1 La VaR

La VaR en rendement pour la prochaine période est définie comme étant la perte telle que

$$\Pr(r_{t+1} < -VaR_{t+1}^{p}) = p \tag{1.1}$$

où $p \in [0,1]$ est une probabilité généralement fixée à 0,05 ou 0,01.

En mots, la VaR représente la perte en rendement qui sera dépassée dans une proportion $p \cdot 100\%$ des cas sur la prochaine période [2]. Elle est normalisée pour être un nombre positif. Si nous connaissons la distribution de probabilité du rendement pour la prochaine période r_{t+1}, alors la VaR dans (1.1) peut être exprimée par

$$VaR_{t+1}^{p} = -F_{r_{t+1}}^{-1}(p)$$

où $F_{r_{t+1}}^{-1}(p)$ représente la fonction de répartition inverse, aussi appelée fonction quantile, de r_{t+1} . Cette expression montre que la VaR correspond au quantile p de la queue de gauche de la distribution de r_{t+1}, normalisé pour être positif .

À partir de la variable aléatoire r_{t+1} nous pouvons définir une nouvelle variable aléatoire de rendement standardisée z_{t+1} d'espérance nulle et de variance unitaire par la transformation de localisation et de mise à l'échelle suivante :

$$z_{t+1} = \frac{r_{t+1} - \mu_{t+1}}{\sigma_{t+1}} \tag{1.2}$$

où r_{t+1} et σ_{t+1} sont respectivement l'espérance et l'écart-type de r_{t+1}. L'équation (1.2) peut être réécrite afin d'exprimer r_{t+1} en fonction de z_{t+1} :

$$r_{t+1} = \mu_{t+1} + \sigma_{t+1} z_{t+1}. \tag{1.3}$$

2. La proportion des cas où la VaR ne sera pas dépassée est donnée par $(1-p) \cdot 100\%$. Ce nombre est souvent désigné comme le niveau de confiance.

Cette expression sera utile afin de modéliser individuellement les dynamiques de μ_{t+1} et σ_{t+1}.

On peut obtenir une représentation moyenne-variance de la VaR en utilisant (1.1) et (1.3). On obtient alors

$$VaR_{t+1}^{p} = -\mu_{t+1} - \sigma_{t+1} F_{z_{t+1}}^{-1}(p). \tag{1.4}$$

Ainsi, la VaR correspond également au quantile p de la queue de gauche de la distribution de z_{t+1} auquel on applique la transformation (1.4).

1.2 La CVaR

Bien que la VaR soit une mesure de risque populaire et intuitive, elle a deux limites importantes [3] :

1. Elle n'est pas une mesure de risque cohérente. Spécifiquement, elle ne respecte pas toujours la propriété de sous-additivité ou le principe de diversification.
2. Elle ne dit rien par rapport à l'étendue des pertes au-delà de la VaR dans $p \cdot 100\%$ des cas. Autrement dit, la VaR est insensible à la portion de la distribution à sa gauche.

La CVaR permet d'éviter ces deux limitations. Elle est définie de la manière suivante :

$$CVaR_{t+1}^{p} = -E\left[r_{t+1} \mid r_{t+1} < -VaR_{t+1}^{p}\right]. \tag{1.5}$$

En mots, la CVaR correspond à la moyenne des pertes à la gauche de la VaR. Tout comme la VaR, elle est normalisée pour être positive.

Si nous connaissons la distribution de probabilité du rendement pour la prochaine période, alors la CVaR peut être exprimée en utilisant la fonction de répartition et la fonction de densité de r_{t+1} [4] :

$$CVaR_{t+1}^{p} = -\frac{1}{p}\int_{-\infty}^{-VaR_{t+1}^{p}} r_{t+1} \cdot f(r_{t+1})\, dr_{t+1}. \tag{1.6}$$

3. Voir Artzner et collab. (1999).
4. Voir l'annexe A dans Simonato (2011) pour la dérivation.

De même, tout comme la VaR, nous pouvons exprimer la CVaR en fonction de l'espérance et l'écart type de r_{t+1} en utilisant (1.3). On obtient alors

$$CVaR_{t+1}^{p} = -\mu_{t+1} - \sigma_{t+1}\frac{1}{p}\int_{-\infty}^{F_{z_{t+1}}^{-1}(p)} z_{t+1} \cdot f(z_{t+1})\, dz_{t+1}. \tag{1.7}$$

Les équations (1.4) et (1.7) permettent de calculer des VaRs et des CVaRs non conditionnelles puisqu'elles supposent une distribution non conditionnelle pour r_{t+1}. Toutefois, il est toujours avantageux d'utiliser l'information disponible afin d'améliorer nos prévisions du risque. C'est pourquoi nous utiliserons plutôt la distribution conditionnelle de r_{t+1}, où l'on conditionne par rapport à l'ensemble des rendements passés. Plus précisément, nous supposons que seuls les deux premiers moments de la distribution conditionnelle de r_{t+1} varient dans le temps alors que les moments d'ordre supérieur sont constants [5]. Cela ne change pas la forme des équations (1.4) et (1.7), à l'exception qu'il faut interpréter μ_{t+1} et σ_{t+1} comme étant des moments conditionnels aux rendements passés. La distribution de z_{t+1} demeure une distribution non conditionnelle par hypothèse. De même, toutes les VaRs et les CVaRs qui seront calculées seront conditionnelles aux rendements passés.

1.3 La VaR et la CVaR par simulation

Les expressions (1.4) et (1.7) sont utiles pour calculer des VaRs et des CVaRs de portefeuille avec une approche univariée, mais elles ne peuvent pas être utilisées avec une approche multivariée.

Avec l'approche univariée, on construit une série de pseudo rendements de portefeuille pour $t = 1$ à T en utilisant les poids actuels des actifs :

$$\{r_t\}_{t=1}^{T} = \{w_T^{\top} r_t\}_{t=1}^{T}$$

où $w_T = (w_{1T}, w_{2T}, \ldots, w_{NT})^{\top}$ est le vecteur des poids d'actifs au temps T et $r_t = (r_{1t}, r_{2t}, \ldots, r_{Nt})^{\top}$ est le vecteur des rendements d'actifs au temps t. Une fois cette série construite, nous pou-

5. Autrement dit, nous supposons que les moments d'ordre supérieur sont équivalents aux moments de la distribution non conditionnelle que l'on suppose stationnaire. Voir Hansen (1994) pour plus de détails.

vons modéliser et estimer les paramètres de la distribution conditionnelle de r_{T+1} avec les observations $\{r_t\}_{t=1}^{T}$. Comme μ_{T+1} et σ_{T+1} sont fonction des paramètres estimés et des rendements passés jusqu'au temps T, ils peuvent être calculés au temps T. Les expressions analytiques (1.4) et (1.7) nous donnent alors directement la VaR_{T+1}^{p} et la $CVaR_{T+1}^{p}$.

Avec l'approche multivariée, on utilise les rendements d'actifs $\{r_t\}_{t=1}^{T}$ afin de modéliser la distribution conditionnelle multivariée du vecteur r_{T+1}. Le problème est que la distribution univariée conditionnelle du portefeuille pour la prochaine période $r_{T+1} = w_{T+1}^{\top} r_{T+1}$ n'est pas connue dans la majorité des cas. En effet, à l'exception de la distribution normale multivariée, il n'y a pas de solution analytique pour la fonction de densité d'une combinaison linéaire de variables aléatoires conjointement distribuées . Nous utiliserons donc des simulations, autant pour l'approche univariée que multivariée, pour la prévision de VaRs et de CVaRs.

Lorsque nous utilisons l'approche univariée, nous simulons directement les rendements aléatoires de portefeuille pour la prochaine période. On obtient alors la série suivante :

$$\{\check{r}_{i,T+1}\}_{i=1}^{K}$$

où $\check{r}_{i,T+1}$ représente le rendement pour la prochaine période de la simulation i et K représente le nombre de simulations.

Lorsque nous utilisons l'approche multivariée, nous simulons plutôt le vecteur de rendements d'actifs r_t du portefeuille. On obtient alors la série suivante :

$$\{\check{r}_{i,T+1}\}_{i=1}^{K}.$$

Pour obtenir une série de rendements de portefeuille, il suffit d'utiliser les poids des actifs au temps T :

$$\{\check{r}_{i,T+1}\}_{i=1}^{K} = \{w_T^{\top} \check{r}_{i,T+1}\}_{i=1}^{K}.$$

Dans les deux cas, les expressions pour le calcul de VaR et CVaR sur la prochaine période sont données dans Christoffersen (2012) par :

$$VaR_{T+1}^{p} = -Percentile\left\{\{\check{r}_{i,T+1}\}_{i=1}^{K}, 100p\right\} \tag{1.8}$$

$$CVaR_{T+1}^{p} = -\frac{1}{p \cdot K} \sum_{i=1}^{K} \check{r}_{i,T+1} \cdot 1_{\{\check{r}_{i,T+1} < -VaR_{T+1}^{p}\}} \tag{1.9}$$

où $1_{\{\cdot\}}$ est une fonction indicatrice qui vaut 1 si l'argument est vrai et 0 autrement. La VaR correspond au $100p^{\text{e}}$ centile des rendements de la simulation alors que la CVaR correspond à la moyenne des rendements simulés inférieurs à la VaR (négative). Toutes les VaRs et les CVaRs qui seront calculées dans le texte utilisent les formules (1.8) et (1.9).

Chapitre 2

La modélisation univariée du risque

2.1 Un modèle

Pour prévoir les rendements de la prochaine période, Christoffersen (2012) utilise un modèle général dont la forme est la suivante :

$$r_{t+1} = \mu_{t+1} + \sigma_{t+1} z_{t+1} \quad \text{avec} \quad z_{t+1} \overset{i.i.d}{\sim} D(0,1) \tag{2.1}$$

où μ_{t+1} et σ_{t+1} représentent respectivement l'espérance et l'écart type de la distribution conditionnelle de r_{t+1}. Le terme d'innovation z_{t+1} suit une distribution centrée réduite D non conditionnelle. Il est indépendant et identiquement distribué au travers du temps.

Cette modélisation est flexible puisqu'elle permet que la moyenne et la volatilité des rendements soient dynamiques et que la distribution conditionnelle des rendements soit non gaussienne. Toutefois, ce modèle suppose que l'asymétrie et le kurtosis de la distribution conditionnelle des rendements sont constants dans le temps.

La spécification du modèle se fait en trois étapes :

1. Choisir la distribution univariée D
2. Choisir la forme de μ_{t+1}
3. Choisir la forme de σ^2_{t+1}

L'espérance conditionnelle μ_{t+1} peut prendre n'importe qu'elle forme et n'est pas nécessairement linéaire dans les paramètres. Le but est alors d'obtenir une approximation en choisissant une certaine spécification pour μ_t . Par exemple, on peut supposer une certaine persistance dans le niveau des rendements en choisissant un modèle autorégressif (AR) pour μ_t :

$$\mu_{t+1} = \phi_0 + \phi_1 r_t + \ldots + \phi_k r_{t-k}.$$

De même, la variance σ^2_{t+1} est un paramètre de la distribution conditionnelle pouvant prendre la forme de n'importe quelle fonction positive des rendements passés. Le but est d'en faire une approximation en choisissant une spécification pour σ^2_{t+1}. Une approche possible est de supposer que σ^2_{t+1} réagit positivement à l'amplitude des innovations passées du modèle. Ceci donne lieu au modèle ARCH de Engle (1982). Une autre serait de supposer que la variance conditionnelle affiche également une certaine persistance. Nous avons alors le modèle GARCH de Bollerslev (1986) dont la forme la plus simple est donnée par

$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 z_t^2 + \beta_1 \sigma_t^2$$

avec $\alpha_0 > 0$ $\alpha_1, \beta_1 \geq 0$ et $\alpha_1 + \beta_1 < 1$.

2.2 Estimation du modèle

Une fois que le modèle univarié est bien spécifié nous devons estimer conjointement les paramètres de la distribution D, de μ_{t+1} et de σ^2_{t+1} à partir d'un échantillon de données $\{r_t\}_{t=1}^T$. Comme σ^2_{t+1} comprend typiquement des termes d'erreurs retardés non observables, la méthodologie privilégiée est l'estimation numérique par maximum de vraisemblance. La procédure comprend deux étapes :

1. Construire la fonction de vraisemblance.

2. Trouver les paramètres qui maximisent cette fonction.

La fonction de vraisemblance est une réécriture de la densité conditionnelle jointe d'observer un certain échantillon de rendements compte tenu d'un ensemble de paramètres θ :

$$L(\theta \mid r_1, r_2, ...r_T) = f(r_1, r_2, ...r_T \mid \theta).$$

Cette densité jointe peut être décomposée de la manière suivante :

$$L(\theta \mid r_1, r_2, ...r_T) = \prod_{t=1}^{T} f(r_t \mid \mathscr{F}_{t-1}, \theta)$$

avec $\mathscr{F}_{t-1} = \{r_{t-1}, ..., r_1\}$ et $\mathscr{F}_0 = \emptyset$.

Il est généralement préférable de maximiser le log de la fonction de vraisemblance :

$$\log\ L(\theta \mid r_1, r_2, ...r_T) = \sum_{t=1}^{T} \log f(r_t \mid \mathscr{F}_{t-1}, \theta). \tag{2.2}$$

Il est souvent impossible de calculer l'espérance et la variance conditionnelle des premières observations étant donné que les rendements avant $t = 1$ sont non observables. Si l'échantillon est suffisamment grand nous pouvons ignorer ces observations et maximiser uniquement le log de la fonction de vraisemblance conditionnelle. Nous pouvons aussi fixer l'espérance et la variance conditionnelle de ces observations à l'espérance et la variance non conditionnelle des rendements qui sont estimées respectivement par la moyenne et la variance de l'échantillon.

Évidemment, la forme de (2.2) dépend de la distribution conditionnelle choisie. Par exemple, si l'on suppose que les rendements suivent une distribution conditionnelle normale, le log de la fonction de vraisemblance s'écrit :

$$\log\ L(\theta \mid r_1, r_2, ...r_T) = \sum_{t=1}^{T} \left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{(r_t - \mu_t)^2}{\sigma_t^2} \right]. \tag{2.3}$$

Une fois que nous avons construit le log de la fonction de vraisemblance nous devons calculer une solution numérique $\hat{\theta}$ donnée par :

$$\hat{\theta} = \arg\max_{\theta} \left[\ \log\ L(\theta \mid r_1, r_2, ...r_T)\ \right]$$

La procédure est la suivante :

1. On initialise les paramètres à des valeurs de départ θ_0 et on évalue le log de la fonction de vraisemblance en ce point : $\log L(\theta_0 \mid r_1, r_2, ...r_T)$.
2. Un algorithme est utilisé afin de générer des nouvelles valeurs de paramètres θ_1 et réévaluer le log de la fonction de vraisemblance : $\log L(\theta_1 \mid r_1, r_2, ...r_T)$.
3. Ces itérations sont effectuées tant et aussi longtemps que $\log L(\theta_i \mid r_1, r_2, ...r_T) > \log L(\theta_{i-1} \mid r_1, r_2, ...r_T)$.

2.3 Simulation du modèle

Une fois que nous avons estimé les paramètres de D, μ_{t+1} et σ^2_{t+1} à l'aide des observations $\{r_t\}_{t=1}^T$ nous pouvons simuler le modèle (2.1) afin de calculer la VaR^p_{T+1} et la $CVaR^p_{T+1}$. La procédure est la suivante :

1. On génère K termes d'erreur aléatoires suivant la distribution $D(0,1)$:

$$\check{z}_{i,T+1} \quad i = 1, \ldots, K$$

2. On calcule $\check{\mu}_{T+1}$ et $\check{\sigma}^2_{T+1}$ à l'aide des rendements et des termes d'erreur passés.
3. On calcule les rendements simulés pour la prochaine période :

$$\check{r}_{i,T+1} = \check{\mu}_{T+1} + \check{\sigma}_{T+1}\check{z}_{i,T+1} \quad i = 1, \ldots, K$$

Les formules (1.8) et (1.9) nous donnent alors directement la VaR^p_{T+1} et la $CVaR^p_{T+1}$.

Chapitre 3

La modélisation multivariée du risque

3.1 L'approche multivariée

Le modèle du chapitre précédent peut être généralisé à un vecteur aléatoire de rendements d'actifs de dimension N :

$$r_{t+1} = \mu_{t+1} + \Sigma_{t+1}^{1/2} z_{t+1} \quad \text{avec} \quad z_{t+1} \overset{i.i.d}{\sim} D(0, I_N) \tag{3.1}$$

où

$$\mu_{t+1} = \begin{pmatrix} \mu_{1,t+1} \\ \mu_{2,t+1} \\ \vdots \\ \mu_{N,t+1} \end{pmatrix},$$

$$\Sigma_{t+1} = \begin{pmatrix} \sigma_{11,t+1} & \sigma_{12,t+1} & \dots & \sigma_{1N,t+1} \\ \sigma_{21,t+1} & \sigma_{22,t+1} & \dots & \sigma_{2N,t+1} \\ \vdots & & & \\ \sigma_{N1,t+1} & \sigma_{N2,t+1} & \dots & \sigma_{NN,t+1} \end{pmatrix}$$

et

$$z_{t+1} = \begin{pmatrix} z_{1,t+1} \\ z_{2,t+1} \\ \vdots \\ z_{N,t+1} \end{pmatrix}$$

représentent respectivement le vecteur d'espérance conditionnelle , la matrice variance-covariance conditionnelle (supposée définie positive) et le vecteur des termes d'innovation. En mots, le modèle (3.1) implique que le vecteur de rendements r_{t+1} suit une distribution conditionnelle multivariée D d'espérance μ_{t+1} et de variance Σ_{t+1} .

Un problème avec les modèles multivariés comme en (3.1) est qu'ils contiennent naturellement beaucoup de paramètres. On doit donc imposer des restrictions fortes sur μ_{t+1} et Σ_{t+1} afin de pouvoir utiliser cette approche. Un autre désavantage de cette approche est que les choix de distributions multivariées pour z_{t+1} sont limités puisqu'il nous faut une densité analytique afin d'estimer le modèle par maximum de vraisemblance. De plus, les distributions multivariées non gaussiennes, comme la distribution de Student multivariée, imposent des restrictions fortes sur les distributions marginales des actifs [1]. Enfin, le plus grand problème, du point de vue de la gestion des risques, d'utiliser des distributions multivariées analytiques traditionnelles est qu'elles sous-estiment les dépendances dans les queues des distributions des actifs puisqu'elles ne prennent pas en compte les dépendances non-linéaires.

Les copules permettent de répondre à ces problèmes. Elles seront présentées dans la prochaine section.

1. Par exemple, la distribution de Student multivariée impose que chaque distribution marginale ait le même nombre de degrés de liberté. Voir chapitre 9 dans Christoffersen (2012) pour plus de détails.

3.2 Les copules

Une copule est une fonction de répartition multivariée d'un vecteur aléatoire U dont les distributions marginales sont uniformes :

$$C(u) = P(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_N \leq u_N).$$

Il s'agit ici d'une copule non-conditionnelle. Par le théorème de Sklar (1959), cette copule non-conditionnelle permet de modéliser, de manière unique, une distribution multivariée non-conditionnelle continue à partir des distributions marginales des variables aléatoires. Plus précisément, le théorème de Sklar (1959) nous permet d'exprimer la fonction de répartition multivariée de r_t en fonction des fonctions de répartition univariées des termes d'erreur $F_i(z_{it})$:

$$F(r_t) = C(F_1(z_{1,t}), F_2(z_{2,t}), \ldots, F_N(z_{N,t})) = C(u_{1,t}, u_{2,t}, \ldots, u_{N,t}).$$

Comme chaque fonction de répartition marginale contient l'ensemble de l'information univariée sur la variable aléatoire $r_{i,t}$, la copule contient l'ensemble de l'information reliée à la dépendance entre les variables aléatoires $r_{i,t}$.

Si la fonction de répartition multivariée $F(r_t)$ est différentiable N fois, alors la densité de la distribution multivariée de r_t est donnée par :

$$f(r_t) = \frac{\partial^N}{\partial r_{1t} \partial r_{2t} \ldots \partial r_{Nt}} F(r_t) = \prod_{i=1}^{N} f(r_{i,t}) \cdot c(u_{1,t}, u_{2,t}, \ldots, u_{N,t}) \tag{3.2}$$

où c représente la fonction de densité de la copule.

Comme il est toujours avantageux d'utiliser l'information disponible pour prévoir le risque, nous voulons conditionner la distribution par rapport aux rendements passés. Patton (2006) montre que le théorème de Sklar (1959) s'applique pour les distributions conditionnelles.

3.3 Estimation avec une copule

Les copules offrent une nouvelle approche pour estimer les paramètres d'une distribution multivariée conditionnelle. La densité (3.2) implique que le log de la vraisemblance de l'échantillon $\{r_t\}_{t=1}^T$ est donné par

$$\log\ f(r_t) = \sum_{t=1}^{T}\sum_{i=1}^{N} \log\ f(r_{i,t}) + \sum_{t=1}^{T} \log\ c(u_{1,t}, u_{2,t}, \ldots, u_{N,t}) \tag{3.3}$$

où le conditionnement par rapport aux paramètres et aux rendements passés est implicite.

Si les paramètres des densités marginales sont différents, la maximisation du premier terme de (3.3) revient à maximiser individuellement la vraisemblance de chaque actif, c'est-à-dire faire une estimation univariée en supposant une distribution marginale différente pour chaque actif. Par la suite, la maximisation de (3.3) est obtenue en maximisant la vraisemblance de la copule. Cette estimation en deux étapes offre beaucoup de flexibilité pour modéliser une distribution conditionnelle multivariée.

La procédure d'estimation est la suivante :

1. À l'aide des observations $\{r_{i,t}\}_{t=1}^T$, on estime les paramètres de D_i, $\mu_{i,t}$ et $\sigma_{i,t}^2$ pour $i = 1, \ldots, N$ par maximum de vraisemblance comme dans la section 2.2.
2. On calcule les termes d'erreur $z_{i,t} = (r_{i,t} - \mu_{i,t})/\sigma_{i,t}$ pour $i = 1, \ldots, N$ et $t = 1, \ldots, T$.
3. On calcule les probabilités $u_{i,t} = F_i(z_{i,t})$ pour $i = 1, \ldots, N$ et $t = 1, \ldots, T$.
4. On estime les paramètres de la copule par maximum de vraisemblance, c'est-à-dire en maximisant $\sum_{t=1}^T \log\ c(u_{1,t}, u_{2,t}, \ldots, u_{N,t})$.

3.4 Simulation avec une copule

Les VaRs et les CVaRs multivariées pour la prochaine période peuvent être calculées en simulant la densité de la copule dans (3.2) . La procédure est la suivante :

1. On simule les probabilités $(\breve{u}_{1,T+1}, \breve{u}_{2,T+1}, ..., \breve{u}_{N,T+1})$ à partir de la densité de la copule.

2. On calcule les termes d'erreur à l'aide des fonctions de répartition inverses marginales : $\check{z}_{j,T+1} = F_j^{-1}(\check{u}_{j,T+1})$ pour $j = 1, \ldots, N$.
3. On calcule $\check{\mu}_{j,T+1}$ et $\check{\sigma}^2_{j,T+1}$ pour $j = 1, \ldots, N$ à l'aide des rendements et des termes d'erreur passés.
4. On calcule les rendements simulés pour la prochaine période :

$$\check{r}_{j,T+1} = \check{\mu}_{j,T+1} + \check{\sigma}_{j,T+1}\check{z}_{j,T+1} \quad j = 1, \ldots, N.$$

5. On répète K fois.

On obtient ainsi une série multivariée $\{\check{r}_{i,T+1}\}_{i=1}^K$ à laquelle on applique les poids de portefeuille au temps T pour obtenir $\{\check{r}_{i,T+1}\}_{i=1}^K = \{w_T^\top \check{r}_{i,T+1}\}_{i=1}^K$. Finalement, les expressions (1.8) et (1.9) nous donnent la VaR^p_{T+1} et la $CVaR^p_{T+1}$.

Chapitre 4

Le *backtesting*

Le *backtesting* est un diagnostic permettant d'accepter ou de rejeter un modèle de risque. Dans ce mémoire, le *backtesting* consiste à comparer les mesures de risque VaRs et CVaRs calculées ex-ante aux rendements de portefeuille réalisés ex-post. Nous utiliserons une approche statistique afin de tester la validité de nos modèles.

4.1 Le *backtest* de la VaR

Nous utiliserons les tests de Kupiec (1995) et Christoffersen (1998). À partir d'une série chronologique de prévisions de VaRs ex-ante et de rendements de portefeuille réalisés ex-post nous pouvons définir la séquence de *hits* $\{I_{t+1}\}_{t=1}^{T}$ où

$$I_{t+1} = \begin{cases} 1 \text{ si } r_{p,t+1} < -VaR_{t+1}^{p} \\ 0 \text{ si } r_{p,t+1} \geq -VaR_{t+1}^{p} \end{cases} .$$

Autrement dit, la séquence de *hits* affiche 1 lorsque le rendement de portefeuille a été inférieur à $-VaR$ sur la période et 0 sinon. Lorsque $I_{t+1} = 1$, on dit qu'il y a eu dépassement de la VaR. Christoffersen (1998) explique que pour qu'un modèle de VaR soit valide la séquence de *hits* doit respecter deux propriétés :

1. Propriété de couverture non conditionnelle :

La fréquence des dépassements de VaR ne doit pas être statistiquement différente de la probabilité théorique p du modèle. Si la fréquence de dépassement observée est supérieure à p alors notre modèle sous-estime le risque. Si la fréquence de dépassement observée est inférieure à p alors notre modèle surestime le risque. Mathématiquement, cette propriété s'exprime

$$\Pr(I_{t+1} = 1) = E(I_{t+1}) = p.$$

2. Propriété d'indépendance :

   Les *hits* doivent être indépendants dans le temps. Intuitivement, la séquence passée de *hits* $\{\ldots, I_{t-1}, I_t\}$ ne doit pas révéler d'information sur la probabilité de dépassement de la VaR pour la prochaine période. Autrement, la VaR serait surestimée ou sous-estimée. Par exemple, supposons que les *hits* arrivent toujours en paires de sorte qu'un dépassement de VaR est toujours suivi d'un autre dépassement. Dans ce cas, une fois que le premier *hit* est noté la probabilité de dépassement n'est plus de p, mais de 100%. La VaR est alors sous-estimée. Mathématiquement, la propriété d'indépendance s'exprime

$$\Pr(I_{t+1} = 1 \mid \mathscr{F}_t) = E(I_{t+1} \mid \mathscr{F}_t) = \Pr(I_{t+1} = 1)$$

   où $\mathscr{F}_t = \{\ldots, I_{t-1}, I_t\}$.

Un modèle de VaR qui respecte ces deux propriétés est valide. On dit alors qu'il respecte la propriété de couverture conditionnelle définie par

$$\Pr(I_{t+1} = 1 \mid \mathscr{F}_t) = E(I_{t+1} \mid \mathscr{F}_t) = p.$$

Notons que la couverture conditionnelle implique à la fois la couverture non conditionnelle et l'indépendance. La propriété de couverture conditionnelle implique également que la séquence $\{I_{t+1}\}_{t=1}^{T}$ est identiquement et indépendamment distribuée selon une distribution de Bernoulli de paramètre p :

$$I_{t+1} \overset{iid}{\sim} Bernoulli(p).$$

4.1.1 Test de couverture non conditionnelle [1]

Nous voulons d'abord tester si la probabilité de dépassement non conditionnelle du modèle, notée π, est statistiquement différente la probabilité théorique choisie p. L'hypothèse nulle est

$$H_0 : \pi = p$$

et l'hypothèse alternative est

$$H_1 : \pi \neq p.$$

La vraisemblance de l'échantillon sous l'hypothèse nulle et en supposant l'indépendance est donnée par

$$L(p \mid I_1, I_2, \ldots, I_T) = \prod_{t=1}^{T} (1-p)^{T_0} p^{T_1}$$

où T_0 et T_1 représentent respectivement le nombre de 0 et de 1 dans la séquence des *hits*. La vraisemblance sous l'alternative en supposant l'indépendance est donnée par

$$L(\pi \mid I_1, I_2, \ldots, I_T) = \prod_{t=1}^{T} (1-\pi)^{T_0} \pi^{T_1}.$$

Évidemment la probabilité π n'est pas connue, mais nous pouvons l'estimer par maximum de vraisemblance. L'estimateur est

$$\hat{\pi} = \frac{T_1}{T_0 + T_1}.$$

L'hypothèse nulle peut ensuite être testée avec un test du rapport de vraisemblance standard :

$$LR_{uc} = -2 \log \left[\frac{L(p \mid I_1, I_2, \ldots, I_T)}{L(\hat{\pi} \mid I_1, I_2, \ldots, I_T)} \right] \overset{a}{\sim} \chi^2(1).$$

Plus la statistique LR_{uc} est grande, moins l'hypothèse nulle de couverture non conditionnelle est probable. Sous l'hypothèse nulle la statistique suit asymptotiquement une distribution χ^2 avec un degré de liberté correspondant à une restriction ($\pi = p$). La p-valeur associée à la statistique est donnée par

$$\text{p-valeur} = 1 - F_{\chi^2(1)}(LR_{uc})$$

où $F_{\chi^2(1)}(\cdot)$ est la fonction de répartition de la distribution χ^2 avec un degré de liberté.

1. Ce test est développé dans Kupiec (1995).

4.1.2 Test d'indépendance [2]

Pour tester l'indépendance des dépassements de VaR nous allons supposer sous l'hypothèse alternative que la séquence des *hits* suit un processus markovien de premier ordre décrit par la matrice de transition suivante :

$$\Pi_1 = \begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}$$

où

$$\begin{aligned} \pi_{01} &= \Pr(I_{t+1} = 1 \mid I_t = 0) \\ \pi_{11} &= \Pr(I_{t+1} = 1 \mid I_t = 1) \\ 1-\pi_{01} &= \Pr(I_{t+1} = 0 \mid I_t = 0) \\ 1-\pi_{11} &= \Pr(I_{t+1} = 0 \mid I_t = 1). \end{aligned}$$

Avec ce processus, la vraisemblance de l'échantillon sous l'hypothèse alternative est donnée par

$$L(\Pi_1 \mid I_1, I_2, \ldots, I_T) = (1-\pi_{01})^{T_{00}} \pi_{01}^{T_{01}} (1-\pi_{11})^{T_{10}} \pi_{11}^{T_{11}}.$$

Les probabilités π_{01} et π_{11} ne sont pas connues, mais nous pouvons les estimer par maximum de vraisemblance. Les estimateurs sont donnés par

$$\hat{\pi}_{01} = \frac{T_{01}}{T_{00}+T_{01}} \qquad \text{et} \qquad \hat{\pi}_{11} = \frac{T_{11}}{T_{10}+T_{11}}$$

où T_{ij} $i, j = 0, 1$ est le nombre d'observations avec un j suivant un i dans la séquence des *hits*. Maintenant, l'hypothèse nulle d'indépendance des dépassements sous ce processus se résume à

$$H_0 : \pi_{01} = \pi_{11} = \pi$$

et la matrice de transition sous l'hypothèse nulle devient

$$\Pi = \begin{pmatrix} 1-\pi & \pi \\ 1-\pi & \pi \end{pmatrix}.$$

2. Ce test est développé dans Christoffersen (1998).

La vraisemblance de l'échantillon sous l'hypothèse nulle est donc

$$L(\Pi \mid I_1, I_2, \ldots, I_T) = \prod_{t=1}^{T} (1-\pi)^{(T_{00}+T_{10})} \pi^{(T_{01}+T_{11})}.$$

Notons que $L(\Pi \mid I_1, I_2, \ldots, I_T)$ correspond à la vraisemblance sous l'hypothèse alternative dans le test de couverture non conditionnelle. De même, nous utilisons l'estimateur

$$\hat{\pi} = \frac{T_{01}+T_{11}}{T_{00}+T_{10}+T_{01}+T_{11}} = \frac{T_1}{T_0+T_1}$$

pour effectuer le test. L'hypothèse nulle d'indépendance peut ensuite être testée avec un test du rapport de vraisemblance :

$$LR_{ind} = -2\log\left[\frac{L(\hat{\Pi} \mid I_1, I_2, \ldots, I_T)}{L(\hat{\Pi}_1 \mid I_1, I_2, \ldots, I_T)}\right] \overset{a}{\sim} \chi^2(1).$$

Plus la statistique LR_{ind} est grande, moins l'hypothèse nulle d'indépendance est probable. Notons que ce test est insensible à la fréquence des dépassements de VaR. Sous l'hypothèse nulle, la statistique suit asymptotiquement une distribution χ^2 avec un degré de liberté correspondant à une restriction ($\pi_{01} = \pi_{11}$). La p-valeur associée à la statistique est donnée par

$$\text{p-valeur} = 1 - F_{\chi^2(1)}(LR_{ind}).$$

4.1.3 Test de couverture conditionnelle [3]

Nous allons maintenant tester conjointement la couverture non conditionnelle et l'indépendance de la séquence des *hits*. Pour ce faire, nous allons tester l'hypothèse nulle du test de couverture non conditionnelle contre l'hypothèse alternative du test d'indépendance. Le test du rapport de vraisemblance est donc

$$LR_{cc} = -2\log\left[\frac{L(p \mid I_1, I_2, \ldots, I_T)}{L(\hat{\Pi}_1 \mid I_1, I_2, \ldots, I_T)}\right] \overset{a}{\sim} \chi^2(2).$$

Plus la statistique LR_{cc} est grande, moins l'hypothèse nulle de couverture conditionnelle est probable. Notons que la statistique suit asymptotiquement une distribution χ^2 avec

3. Ce test est développé dans Christoffersen (1998).

deux degrés de liberté correspondant à deux restrictions ($\pi_{01} = \pi_{11} = p$). La p-valeur associée à la statistique est donnée par

$$\text{p-valeur} = 1 - F_{\chi^2(2)}(LR_{cc}).$$

Les trois tests sont liés par la relation suivante :

$$LR_{cc} = LR_{uc} + LR_{ind}.$$

L'approche de Christoffersen (1998) permet d'évaluer si l'échec du test de couverture conditionnelle provient d'une couverture inadéquate ou du regroupement des *hits*.

4.2 Le *backtest* de la CVaR

Nous utiliserons les deux premiers tests proposés par Acerbi et Szekely (2014). Ces tests ne font aucune hypothèse sur la distribution des rendements de portefeuille à chaque période et ne requièrent aucune estimation de paramètres.

4.2.1 Test 1

Rappelons que la CVaR est définie par

$$CVaR_{t+1}^{p} = -E\left[r_{t+1} \mid r_{t+1} < -VaR_{t+1}^{p}\right].$$

En faisant un peu d'algèbre on obtient

$$E\left[\frac{r_{t+1}}{CVaR_{t+1}^{p}} \mid r_{t+1} + VaR_{t+1}^{p} < 0\right] + 1 = 0. \tag{4.1}$$

Maintenant, définissons comme précédemment la variable indicatrice I_{t+1} telle que

$$I_{t+1} = \begin{cases} 1 \text{ si } r_{p,t+1} < -VaR_{t+1}^{p} \\ 0 \text{ si } r_{p,t+1} \geq -VaR_{t+1}^{p} \end{cases}.$$

L'estimateur d'échantillon analogue à 4.1 peut alors être écrit comme

$$Z_1 = \frac{\sum_{t=1}^{T} \frac{r_{t+1} I_{t+1}}{CVaR_{t+1}^{p}}}{N_T} + 1$$

où $N_T = \sum_{t=1}^{T} I_{t+1}$ est supposé positif. C'est la première statistique de test proposée par Acerbi et Szekely (2014). L'hypothèse nulle pour le test est

H_0 : Les queues de gauche des distributions du modèle sont égales aux queues des vraies distributions à chaque période. Cette hypothèse implique que les CVaRs estimées sont égales aux vraies CVaRs, mais pas vice-versa.

L'hypothèse alternative est

H_1 :

1. Les CVaRs estimées sous-estiment le risque.
2. Les VaRs estimées sont égales aux vraies VaRs.

On note que sous l'alternative on rejette le modèle pour la CVaR, mais pas pour la VaR. En effet, ce test exige d'avoir préalablement accepté le modèle pour la VaR et est insensible à la fréquence des dépassements. Aussi, le test détecte seulement la sous-estimation du risque pour la CVaR.

Acerbi et Szekely (2014) montrent qu'en supposant l'indépendance des rendements r_{t+1} dans le temps la statistique Z_1 a une espérance nulle sous l'hypothèse nulle et une espérance négative sous l'hypothèse alternative. Ainsi, plus la statistique Z_1 est négative moins l'hypothèse nulle est probable.

La distribution de la statistique Z_1 sous l'hypothèse nulle n'est pas connue, mais nous pouvons la simuler. Voici les étapes :

1. Simuler M rendements aléatoires à partir de la distribution du modèle à chaque période t. Si les VaRs et les CVaRs ont déjà été calculées par simulation, il suffit de conserver la distribution simulée à chaque période.
2. Calculer pour chaque simulation $i = 1, \ldots, M$ une statistique Z_1^i.

La p-valeur du test peut finalement être estimée avec

$$\text{p-valeur} = \frac{\sum_{i=1}^{M} 1_{(Z_1^i < Z_1)}}{M}.$$

4.2.2 Test 2

À partir de la définition de la CVaR et en faisant un peu d'algèbre on obtient

$$\frac{E[r_{t+1}I_{t+1}]}{CVaR_{t+1}^{p}\ p} + 1 = 0. \tag{4.2}$$

Ceci suggère la statistique suivante

$$Z_2 = \sum_{t=1}^{T} \frac{r_{t+1}I_{t+1}}{T\ \ p\ CVaR_{t+1}^{p}} + 1.$$

C'est la deuxième statistique de test proposée par Acerbi et Szekely (2014). L'hypothèse nulle pour le test est

H_0 : Les queues de gauche des distributions du modèle sont égales aux queues des vraies distributions à chaque période. Cette hypothèse implique que les CVaRs estimées sont égales aux vraies CVaRs, mais pas vice-versa.

L'hypothèse alternative est

H_1 :
1. Les CVaRs estimées sous-estiment le risque.
2. Les VaRs estimées sous-estiment le risque.

Notons que ce test évalue conjointement la fréquence et la grandeur des dépassements. Autrement dit, il teste simultanément la VaR et la CVaR. Acerbi et Szekely (2014) montrent que la statistique Z_2 a une espérance nulle sous l'hypothèse nulle et une espérance négative sous l'hypothèse alternative. Ainsi, plus la statistique Z_2 est négative moins l'hypothèse nulle est probable. Contrairement au premier test, ces résultats ne requièrent pas l'indépendance des rendements dans le temps.

La distribution de la statistique Z_2 sous l'hypothèse nulle n'est pas connue, mais nous pouvons la simuler. Voici les étapes :

1. Simuler M rendements aléatoires à partir de la distribution du modèle à chaque période t. Si les VaRs et les CVaRs ont déjà été calculées par simulation, il suffit de conserver la distribution simulée à chaque période.
2. Calculer pour chaque simulation $i = 1, \ldots, M$ une statistique Z_2^i.

La p-valeur du test peut finalement être estimée avec

$$\text{p-valeur} = \frac{\sum_{i=1}^{M} 1_{(Z_2^i < Z_2)}}{M}.$$

Les deux tests sont liés par la relation suivante

$$Z_2 = 1 - (1 - Z_1)\frac{N_T}{Tp}.$$

Chapitre 5

L'élicitabilité

Bien que le *backtesting* permette de valider un modèle de manière isolée, il ne permet pas la comparaison entre plusieurs modèles. L'enjeu de ce mémoire concerne la sélection de modèle. Nous voulons déterminer quel modèle offre la meilleure précision lors la prévision de VaRs et CVaRs de portefeuille sur une période. L'approche traditionnelle pour répondre à cette question est de comparer la prévision de la variable d'intérêt, effectuée *ex-ante*, à la réalisation *ex-post* de cette même variable, le but étant d'évaluer si la prévision est relativement « près » de la vraie valeur. Ceci est généralement fait à l'aide d'une fonction de score, aussi appelée fonction de perte, $S(y,x)$ où y correspond à la prévision et x correspond à la réalisation de la variable. Des exemples de fonctions de score très utilisées sont l'erreur quadratique et l'erreur absolue, données respectivement par $(y-x)^2$ et $|y-x|$. Pour une série de prévisions et de réalisations, la moyenne des scores peut être calculée et le modèle affichant le plus petit résultat est alors le plus précis.

Le problème lors de la prévision de VaRs et CVaRs est que la réalisation *ex-post* de la mesure de risque n'est jamais observée. Autrement dit, il est impossible de comparer les prévisions aux réalisations de ces mesures de risque avec des critères tels que l'erreur quadratique moyenne (MSE) ou l'erreur absolue moyenne. Nous n'observons pas non plus la distribution *ex-post* des rendements à partir de laquelle les mesures de risque sont calculées. Nous observons seulement un scénario possible de la distribution des rende-

ments. Ainsi, si nous voulons comparer la précision relative de nos modèles, il nous faut une fonction de score qui compare nos prévisions *ex-ante* de VaRs et CVaRs aux réalisations *ex-post* des rendements de portefeuille plutôt qu'aux réalisations *ex-post* de VaRs et CVaRs qui ne se matérialisent jamais.

Pour comparer les modèles univariés et multivariés, nous exploiterons la propriété d'élicitabilité (*elicitability* en anglais). Cette propriété a été introduite par Osband (1985).

Une mesure de risque $\psi(X)$ est dite élicitable si elle minimise l'espérance d'une fonction de score S :

$$\psi(X) = \arg\min_y E\left[S(y,X))\right].$$

Par exemple, la moyenne solutionne le problème de minimisation suivant :

$$E(X) = \arg\min_y E\left[(X-y)^2\right].$$

Un autre exemple est la médiane qui minimise l'erreur absolue moyenne $E(|X-y|)$.

L'élicitabilité offre un critère de sélection afin de comparer plusieurs modèles de risque entre eux. En comparant la série des prévisions de risque *ex-ante* y_{t+1} aux réalisations *ex-post* de la variable aléatoire x_{t+1} pour $t = 1,\ldots,T$, nous pouvons calculer la moyenne d'échantillon des scores

$$\bar{S} = \frac{1}{T}\sum_{t=1}^{T} S(y_{t+1}, x_{t+1})$$

pour chaque modèle. Le modèle privilégié est alors celui qui affiche le plus petit $\bar{S}$.

5.1 Sélection de modèle de VaR

La VaR, qui correspond au quantile (normalisé positif) de la queue de gauche de la distribution de X, est aussi élicitable :

$$VaR = \arg\min_y E\left[(X+y)(p - 1_{\{X+y<0\}})\right]$$

où p est la probabilité de dépassement associée à la VaR.

Dans ce mémoire, la variable aléatoire X correspond au rendement *ex-post* d'un portefeuille et la variable y correspond à une prévision *ex-ante* de VaR. Notre critère de sélection pour les modèles de VaR devient donc

$$\bar{S} = \frac{1}{T}\sum_{t=1}^{T}\left[(r_{t+1} + VaR_{t+1})\left(p - 1_{\{r_{t+1}+VaR_{t+1}<0\}}\right)\right]$$

Le modèle univarié ou multivarié qui affiche le plus petit score est ensuite sélectionné pour le calcul de VaRs.

5.2 Sélection de modèle de VaR/CVaR

Malheureusement, il a été prouvé dans Gneiting (2011) que la CVaR n'est pas une mesure de risque élicitable. Ceci empêche la comparaison de modèles pour la CVaR. Cependant, Acerbi et Szekely (2014) ont montré que la VaR et la CVaR sont conjointement élicitables. Ils suggèrent la fonction de score suivante :

$$S(y,z,X) = \frac{p}{2}z^2 + \frac{Wp}{2}y^2 - pzy + \left[z(y+X) + \frac{W}{2}(X^2 - y^2)\right]1_{\{X+y<0\}}$$

où W est un nombre réel. Ils montrent que le couple $\{VaR, CVaR\}$ solutionne le problème de minimisation suivant :

$$\{VaR, CVaR\} = \arg\min_{y,z} E\left[S(y,z,X)\right]$$

sous la condition que $VaR \cdot W > ES$. La fonction de score des mesures de risque VaR et CVaR *ex-ante* au temps t suite à la réalisation du rendement de portefeuille *ex-post* est donnée par

$$\begin{aligned} S(VaR_{t+1}, CVaR_{t+1}, r_{t+1}) &= \frac{p}{2}CVaR_{t+1}^2 + \frac{Wp}{2}VaR_{t+1}^2 - p\,CVaR_{t+1}VaR_{t+1} \\ &+ \left[CVaR_{t+1}(VaR_{t+1} + r_{t+1}) + \frac{W}{2}(r_{t+1}^2 - VaR_{t+1}^2)\right]1_{\{r_{t+1}+VaR_{t+1}<0\}}. \end{aligned}$$

Enfin, le critère de sélection pour les modèles de VaR et CVaR devient

$$\bar{S} = \frac{1}{T}\sum_{t=1}^{T} S(VaR_{t+1}, CVaR_{t+1}, r_{t+1}).$$

Le modèle qui affiche le plus petit score est privilégié pour le calcul conjoint de VaRs et de CVaRs.

Chapitre 6

Analyse empirique

Suivant Christoffersen et Langlois (2013), nous utilisons les trois facteurs de Fama et French (1993) et le facteur de Carhart (1997) pour notre application empirique. Nous les appellerons respectivement le facteur *Market*, le facteur *Value*, le facteur *Size* et le facteur *Momentum*. Ces quatre facteurs sont très utilisés dans la littérature financière entre autres parce qu'ils expliquent une part importante de la variation des rendements d'action. Ils ont aussi l'avantage d'être empiriquement orthogonaux, c'est-à-dire faiblement corrélés, ce qui produit des coefficients plus précis lors de l'estimation de régressions.

Le facteur *Market* représente le rendement moyen (pondéré par le cours) de toutes les actions cotées sur le NYSE, le AMEX et le NASDAQ, moins le taux des bons du Trésor d'un mois. À chaque fin du mois de juin, toutes les actions cotées sur ces trois indices boursiers sont séparées en six portefeuilles selon la valeur au marché de l'équité (MVE) médiane et les centiles du ratio *Book-to-Market* (BTM) dans le NYSE :

	MVE médiane	
70e centile BTM	Small Value	Big Value
30e centile BTM	Small Neutral	Big Neutral
	Small Growth	Big Growth

.

Le facteur *Size* représente la différence entre la moyenne des rendements des trois portefeuilles de la première colonne et celle des trois portefeuilles de la deuxième colonne.

Le facteur *Value* représente la différence entre la moyenne des rendements des deux portefeuilles de la première ligne et celle des deux portefeuilles de la dernière ligne.

Chaque mois, les actions cotées sur les trois indices sont séparées en six portefeuilles selon la valeur au marché de l'équité (MVE) médiane et les centiles des rendements de $t-12$ à $t-2$ dans le NYSE :

	MVE médiane	
70e centile rendements (2-12)	Small Up	Big Up
30e centile rendements (2-12)	Small Medium	Big Medium
	Small Down	Big Down

Le facteur *Momentum* représente la différence entre la moyenne des rendements des deux portefeuilles de la première ligne et celle des deux portefeuilles de la dernière ligne.

Comme Christoffersen et Langlois, nous considérons un portefeuille équipondéré des quatre facteurs. Nous cherchons à calculer des VaRs et CVaRs sur ce portefeuille avec les deux approches : univariée et multivariée. Leurs données sont constituées des rendements hebdomadaires, en pourcentage, des quatre facteurs allant du 5 juillet 1963 au 31 décembre 2010 [1]. Les rendements sont convertis en log-rendements avec la formule $r_{\log} = \ln(1 + \frac{r_{\text{hebdo}}}{100})$. Pour la suite, le terme rendement signifiera log-rendement hebdomadaire.

6.1 Rendements des facteurs [2]

Les quatre premiers graphiques de la figure 1 montrent les séries chronologiques des rendements hebdomadaires des facteurs. On note visuellement que tous les facteurs affichent des régimes de forte et de faible volatilité ce qui est caractéristique des actifs financiers. Les séries semblent aussi stationnaires puisqu'elles oscillent dans un intervalle fixe. Les deux graphiques du bas montrent les rendements cumulés des facteurs pour dif-

1. Toutes les données sont disponibles publiquement sur le site de Kenneth French.
2. Cette section reprend l'analyse des sections II. A et B dans Christoffersen et Langlois (2013).

férentes périodes. Les rendements cumulés de 1963 à 2010 pour le facteur *Momentum* se distinguent particulièrement des autres.

La table 1 présente les statistiques descriptives des facteurs. La variance domine largement la moyenne pour tous les facteurs. Les facteurs *Market*, *Size* et *Momentum* affichent une asymétrie négative considérable tandis que le facteur *Value* est davantage symétrique. Tous les facteurs montrent un kurtosis normalisé très important ce qui implique la présence de queues de distribution beaucoup plus épaisses que la distribution normale. Une partie de ce kurtosis excédentaire provient de la dynamique de la distribution conditionnelle des facteurs. Les corrélations entre les facteurs sont très faibles ou négatives. Toutes les corrélations sont non-significatives à un seuil de 1% ce qui montre l'orthogonalité des facteurs entre eux.

La non-normalité univariée des facteurs peut aussi être visualisée à l'aide des diagrammes Quantile-Quantile à la figure 2. Les quantiles empiriques des rendements sont comparés à ceux de la distribution normale de sorte que les déviations par rapport à la ligne de 45 degrés indiquent la non-normalité. On note que les fréquences d'observations dans les queues de droite et de gauche des distributions empiriques sont plus grandes que celles impliquées par la distribution normale.

Il y a également une non-normalité importante au niveau de la distribution jointe des rendements des facteurs. Dans le cas d'une distribution normale multivariée, les corrélations suffisent pour caractériser les dépendances entre les variables. Nous avons alors des dépendances symétriques et linéaires. Pour des distributions non-normales, les corrélations ne caractérisent pas entièrement les relations entre les variables puisqu'il peut y avoir des dépendances non-linéaires et asymétriques. On peut vérifier la non-normalité multivariée entre deux séries chronologiques i et j à l'aide des corrélations de seuil (*Threshold Correlations*) définies dans Christoffersen (2012) comme suit :

$$\bar{\rho}_{ij}(u) = \begin{cases} Corr(r_{it}, r_{jt} \mid r_{it} \leq F_i^{-1}(u) \text{ et } r_{jt} \leq F_j^{-1}(u)) & \text{si } u \leq 0,5 \\ Corr(r_{it}, r_{jt} \mid r_{it} > F_i^{-1}(u) \text{ et } r_{jt} > F_j^{-1}(u)) & \text{si } u > 0,5 \end{cases}$$

où u est une probabilité entre 0 et 1 et $F_i^{-1}(u)$ et $F_j^{-1}(u)$ correspondent aux quantiles

empiriques des séries chronologiques. En mots, il s'agit de calculer des coefficients de corrélation en se limitant à une portion de la queue de gauche ou de droite de la distribution.

Les graphiques de gauche des figures 3 et 4 montrent les diagrammes de dispersion et les corrélations de seuil entre les paires de facteurs. On note le regroupement de plusieurs observations dans les coins des diagrammes de dispersion, et ce même si les coefficients de corrélation sont faibles. Ceci suggère la présence des dépendances non linéaires. Les graphiques de droite des figures 3 et 4 comparent les corrélations de seuil empiriques entre les paires de facteurs à celle qu'implique la distribution normale bivariée avec le même coefficient de corrélation que dans les données. Les différences sont frappantes. Pour toutes les paires de facteurs, les corrélations de seuil sous l'hypothèse de normalité sont près de zéro à tous les quantiles tandis que les corrélations de seuil empiriques sont fortement positives (autour de 0,5 en moyenne). Les corrélations empiriques pour $u \leq 0,5$ sont aussi plus élevées que pour $u > 0,5$ ce qui montre l'asymétrie dans les dépendances entre les rendements des facteurs. Le coefficient de corrélation ne capture pas les fortes dépendances dans les queues des distributions.

6.2 Dynamique des facteurs [3]

Les coefficients d'autocorrélation de la table 1 montrent que les séries chronologiques des facteurs affichent une certaine persistance sur un horizon de trois semaines. La figure 5 illustre la fonction d'autocorrélation empirique des rendements en niveau et des rendements absolus, ainsi que les bornes d'un intervalle de confiance de 95% autour de zéro. On note que pour les rendements en niveau les autocorrélations sont significatives pour les premières semaines puis non significatives par la suite. Pour les rendements absolus, les autocorrélations sont beaucoup plus importantes et persistantes dans le temps. Ceci est typique des actifs financiers dont l'amplitude des rendements, pouvant être mesurée par

3. Cette section reprend l'analyse de la section II. C dans Christoffersen et Langlois (2013).

le rendement au carré ou en absolu, affiche beaucoup plus persistante que les rendements en niveau.

Pour l'approche multivariée, nous utilisons la modélisation employée dans Christoffersen et Langlois (2013). Nous modélisons la distribution conditionnelle des log-rendements hebdomadaires de chacun des quatre facteurs. Le modèle utilisé pour le facteur j est de la forme

$$r_{j,t} = \mu_{j,t} + \sigma_{j,t} z_{j,t} \quad \text{avec} \quad z_{j,t} \overset{i.i.d}{\sim} D(0,1)$$

où $r_{j,t}$ est le rendement du facteur j au temps t et $D(0,1)$ est une distribution centrée réduite. L'espérance conditionnelle des rendements des facteurs est modélisée avec un processus AR(3) :

$$\mu_{j,t} = \phi_{0,j} + \phi_{1,j} r_{j,t-1} + \phi_{2,j} r_{j,t-2} + \phi_{3,j} r_{j,t-3}$$

La variance conditionnelle est modélisée avec un processus NGARCH(1,1) suggéré dans Engle et Ng (1993) :

$$\sigma_{j,t}^2 = \omega_j + \beta_j \sigma_{j,t-1}^2 + \alpha_j (z_{j,t-1} - \theta_j)^2$$

Le paramètre θ_j(supposé positif) introduit un effet de levier en permettant que les chocs négatifs ($z_{j,t-1} < 0$) aient un plus grand impact que les chocs positifs sur la variance conditionnelle de la prochaine période.

Les paramètres des quatre modèles sont estimés par maximum de vraisemblance. Les résultats pour la distribution normale sont présentés dans la première partie de la table 3. Le diagnostic indique que la variance conditionnelle est très persistante pour tous les facteurs. Le paramètre de levier θ_j est fortement positif et significatif pour le facteur *Market* et fortement négatif et significatif pour le facteur *Momentum*. Ainsi, un choc positif ($z_{j,t-1} > 0$) a un plus grand impact sur la variance conditionnelle de la prochaine période qu'un choc négatif pour le facteur *Momentum*. Pour les facteurs *Size* et *Value* l'effet levier est beaucoup moins important.

La figure 6 montre la fonction d'autocorrélation pour les résidus des modèles avec la distribution normale. On remarque que la modélisation des dynamiques de premier et de

second ordre a éliminé l'autocorrélation en niveau et en absolu. Ceci est confirmé par un test de Ljung-Box sur les 20 premières autocorrélations dans les résidus en niveau et en absolu. Les p-valeurs sont données dans le diagnostic de la table 3.

L'hypothèse de normalité pour les termes d'erreur implique que l'asymétrie et le kurtosis excédentaire des résidus devraient être nuls. Le diagnostic de la table 3 indique que l'asymétrie empirique des résidus est négative pour les facteurs *Market*, *Size* et *Momentum* et légèrement positive pour le facteur *Value*. Le kurtosis excédentaire empirique est positif pour tous les facteurs. Les modèles avec la distribution normale ont capturé une partie de l'asymétrie et de la kurtose excédentaire dans les rendements des facteurs, mais il en reste encore.

Comme Christoffersen et Langlois, nous considérons la distribution t asymétrique de Hansen (1994) pour les termes d'erreur. Cette distribution univariée centrée réduite est caractérisée par deux paramètres ν et κ. Le paramètre κ est relié à l'asymétrie et le paramètre ν est relié au kurtosis. La densité et les moments de la distribution t asymétrique de Hansen sont donnés dans l'annexe de Christoffersen et Langlois (2013). Les résultats de l'estimation des modèles AR(3)-NGARCH(1,1) avec la distribution t asymétrique sont donnés dans la deuxième partie de la table 3. L'asymétrie et le kurtosis excédentaire des résidus sont maintenant beaucoup plus près de ceux du modèle. De plus, toutes les log-vraisemblances sont supérieures aux modèles sous la normale. La figure 7 montre les diagrammes Quantile-Quantile qui comparent les quantiles empiriques des résidus à ceux de la distribution de Hansen. Les observations sont beaucoup plus près de la ligne de 45 degrés ce qui témoigne du meilleur *fit* du modèle t asymétrique. Pour la suite nous considérons seulement les modèles AR(3)-NARCH(1,1) avec la distribution t asymétrique.

6.3 Dépendance des facteurs [4]

Nous avons vu dans la dernière section que les distributions marginales des facteurs ne suivent pas la loi normale. Les corrélations de seuil des figures 3 et 4 indiquent éga-

4. Cette section reprend l'analyse des sections III. A et B dans Christoffersen et Langlois (2013).

lement qu'il y a de la non-normalité au niveau de la distribution jointe des facteurs. Pour modéliser cette distribution jointe, nous utiliserons une copule. Les copules offrent une approche flexible pour modéliser une distribution conditionnelle multivariée non-normale en permettant de lier les distributions marginales des facteurs (voir section sur les copules). Patton (2006) montre que le théorème de Sklar (1959) s'applique aux distributions conditionnelles. Nous pouvons donc décomposer la distribution conditionnelle jointe de N facteurs comme suit :

$$f_t(r_{1,t},\ldots,r_{N,t}) = c_t(u_{1,t},\ldots,u_{N,t})\prod_{j=1}^{N} f_{j,t}(r_{j,t})$$

où $c_t(u_{1,t},\ldots,u_{N,t})$ est la fonction de densité conditionnelle de la copule, $u_{j,t} = F_{j,t}(z_{j,t})$ est la probabilité marginale du facteur j et $f_{j,t}(r_{j,t})$ est la fonction de densité marginale conditionnelle du facteur j. Comme nous avons déjà modélisé préalablement les densités marginales conditionnelles $f_{j,t}(r_{j,t})$ de chaque facteur, il reste à choisir une forme fonctionnelle pour la copule $c_t(u_{1,t},\ldots,u_{N,t})$. L'asymétrie dans les corrélations de seuil suggère l'utilisation d'une copule asymétrique.

Suivant Christoffersen et Langlois nous utilisons une copule basée sur la distribution t asymétrique multivariée de Demarta et McNeil (2005). Cette distribution est caractérisée par un vecteur λ de dimension N (le nombre de facteurs) et un scalaire ν_c. Dans la distribution de Demarta et McNeil le vecteur λ capture à la fois l'asymétrie univariée et multivariée. Toutefois, dans la copule t asymétrique l'asymétrie univariée est déjà prise en compte par la modélisation des distributions marginales de sorte que le vecteur λ capture seulement l'asymétrie multivariée. Le paramètre ν_c indique le nombre de degrés de liberté. La densité conditionnelle de la copule t asymétrique est notée $c_t(u_{1t},u_{2t},\ldots,u_{Nt};\lambda,\nu_c,\Psi)$ où Ψ représente la matrice de corrélation de la copule [5]. Les paramètres de la copule sont estimés en maximisant la log-vraisemblance $\sum_{t=1}^{T}\log\ c_t(u_{1,t},u_{2,t},\ldots,u_{N,t};\lambda,\nu_c,\Psi)$.

La première partie de la table 4 présente les résultats de l'estimation de trois copules non-dynamiques : la copule t asymétrique, la copule Student symétrique comme cas spécial où $\lambda_j \to 0$ pour tous les facteurs et la copule normale avec en plus $\nu_c \to \infty$. De

5. Pour plus de détails sur la copule t asymétrique, voir l'annexe A.

gauche à droite, on remarque une augmentation de la log-vraisemblance des modèles, le plus grand gain provenant de l'utilisation de la copule t par rapport à la copule normale. Deux des paramètres d'asymétrie sont significatifs dans la copule t asymétrique.

Comme Christoffersen et Langlois nous permettons à la matrice de corrélation Ψ de la copule t asymétrique d'évoluer au travers du temps selon le modèle DCC (*Dynamic Conditional Correlation*) de Engle (2002). Dans ce modèle, la matrice de corrélation évolue selon la spécification suivante :

$$Q_t = Q(1-\beta_c-\alpha_c)+\beta_c Q_{t-1}+\alpha_c z_{t-1} z_{t-1}^{\top}$$

Dans le modèle de Engle le vecteur z_{t-1} contient les termes d'erreurs standardisés. Cependant ici ce vecteur contient les fractiles standardisés $F_j^{-1}(u_{j,t})$ puisque nous modélisons la matrice de corrélation de la copule. $F_j^{-1}(\cdot)$ est la fonction de répartition inverse marginale spécifique à la copule [6]. La matrice Q est définie comme la matrice de corrélation de z_t dans l'échantillon [7]. La matrice de corrélation conditionnelle est obtenue en normalisant les éléments dans la matrice Q_t :

$$\Psi_t = diag(dg(Q_t))^{-0,5} Q_t diag(dg(Q_t))^{-0,5}$$

où $dg(A)$ est un opérateur qui renvoie un vecteur contenant les éléments de la diagonale de la matrice A.

La deuxième partie de la table 4 montre les résultats de l'estimation des copules dynamiques. On note que les log-vraisemblances de ces modèles sont significativement supérieures à celles des copules non-dynamiques. La figure 8 montre l'évolution des corrélations dynamiques de la copule t asymétrique entre 2006 et 2010. On note d'importantes variations ce qui montre les dangers du supposer des corrélations constantes.

6. Voir l'annexe B pour les détails sur la standardisation des fractiles.
7. Nous utilisons le modèle cDCC dans Aielli (2013), mais les différences entre ce modèle et le modèle original de Engle sont négligeables. Voir les annexes B et C pour les détails de l'estimation du DCC.

6.4 Dynamique de portefeuille

Pour l'approche univariée nous construisons la série chronologique pour $t = 1, \ldots, T$ des pseudo rendements du portefeuille équipondéré $r_{p,t}$ en appliquant les poids aux rendements passés des facteurs. Nous modélisons la distribution conditionnelle du portefeuille de la même façon que pour les distributions marginales des facteurs :

$$r_{p,t} = \mu_{p,t} + \sigma_{p,t} z_{p,t} \quad \text{avec} \quad z_{p,t} \overset{i.i.d}{\sim} D(0,1).$$

L'espérance conditionnelle du rendement de portefeuille est modélisée avec un processus AR(3) :

$$\mu_{p,t} = \phi_{0,p} + \phi_{1,p} r_{p,t-1} + \phi_{2,p} r_{p,t-2} + \phi_{3,p} r_{p,t-3}.$$

La variance conditionnelle est modélisée avec un processus NGARCH(1,1) :

$$\sigma^2_{p,t} = \omega_p + \beta_p \sigma^2_{p,t-1} + \alpha_p (z_{p,t-1} - \theta_p)^2.$$

Les paramètres du modèle sont estimés par maximum de vraisemblance. Les résultats pour la distribution normale et la distribution t asymétrique de Hansen sont présentés à la table 2.

6.5 Calculs de VaRs et CVaRs

Pour calculer les VaRs et les CVaRs hebdomadaires du portefeuille équipondéré selon les deux approches nous utilisons la procédure de prévision hors échantillon de Christoffersen et Langlois (2013).

Pour l'approche multivariée, nous commençons par estimer le modèle AR(3)-NARCH(1,1) avec la distribution t asymétrique de Hansen pour chaque facteur en utilisant les rendements allant de 1963 à 1983, c'est-à-dire 20 ans de données. Nous estimons ensuite les copules de la table 4 à l'aide des résidus des modèles AR(3)-NARCH(1,1). Nous estimons aussi deux modèles de référence avec la distribution normale multivariée, un avec

et l'autre sans corrélations dynamiques [8]. Nous réestimons les modèles une fois par année en utilisant toutes les données disponibles jusqu'à la date d'estimation. Même si les paramètres sont actualisés seulement une fois par année, les espérances, les variances et les corrélations conditionnelles sont mises à jour chaque semaine. Une fois que la distribution jointe conditionnelle pour la prochaine semaine $f_t(r_{t+1})$ est construite, nous simulons 100 000 vecteurs de rendements r_{t+1} de cette distribution. Les VaRs et les CVaRs du portefeuille équipondéré selon l'approche multivariée sont obtenues en appliquant les poids de portefeuille aux rendements des facteurs et en utilisant les formules 1.8 et 1.9 de la section 1.3.

La méthode d'estimation pour l'approche univariée est analogue à celle de l'approche multivariée. Nous commençons par estimer le modèle AR(3)-NARCH(1,1) avec la distribution t asymétrique de Hansen sur les pseudo rendements du portefeuille préalablement construits pour la période allant de 1963 à 1983. Nous estimons aussi un modèle référence avec la distribution normale univariée. Nous réestimons les modèles une fois par année en utilisant toutes les données disponibles jusqu'à la date d'estimation. Même si les paramètres sont actualisés seulement une fois par année l'espérance et la variance conditionnelle du portefeuille sont mises à jour chaque semaine. Une fois que la distribution conditionnelle du portefeuille pour la prochaine semaine $f_t(r_{p,t})$ est construite nous simulons 100 000 rendements de portefeuille $r_{p,t}$ de cette distribution. Les VaRs et les CVaRs du portefeuille équipondéré selon l'approche univariée sont obtenues en utilisant les formules 1.8 et 1.9 de la section 1.3.

À la suite de ces deux procédures nous obtenons 1 436 prévisions de VaRs et de CVaRs hebdomadaires pour chacun des 10 modèles (2 modèles univariés et 8 modèles multivariés). Ces prévisions hors échantillon couvrent la période allant de la première semaine de juillet 1983 jusqu'à la fin de décembre 2010. Par exemple, la figure 9 montre les CVaRs calculées pour la distribution normale multivariée avec et sans corrélations dynamiques. On note qu'à l'automne 2008 les CVaRs calculées avec des corrélations constantes sont

8. Ces deux modèles permettent tout de même des espérances et des variances dynamiques pour les facteurs.

plus faibles que celles obtenues avec des corrélations dynamiques. Cela est dû aux corrélations dynamiques plus faibles entre le facteur *Momentum* et les autres facteurs durant cette période. Dans la prochaine section nous effectuerons le *backtest* de ces VaRs et CVaRs.

6.6 Le *backtesting*

Nous effectuons le *backtesting* des prévisions de VaRs de nos modèles avec les tests de Kupiec (1995) et Christoffersen (1998). Pour ce faire nous définissons la séquence de *hits* $\{I_{t+1}\}_{t=1}^{T}$ où

$$I_{t+1} = \begin{cases} 1 \text{ si } r_{p,t+1} < -VaR_{t+1}^{p} \\ 0 \text{ si } r_{p,t+1} \geq -VaR_{t+1}^{p} \end{cases} .$$

Idéalement, les dépassements de VaRs hebdomadaires devraient avoir une fréquence proche de p et être indépendants. Ces deux propriétés sont testées respectivement avec le test de couverture non conditionnelle et le test d'indépendance dans Christoffersen (1998). Nous effectuons aussi le test conjoint de couverture conditionnelle. Les p-valeurs [9] des trois tests sont données dans les trois premières colonnes à la table 5 pour des VaRs à 1% et 5%. On note que pour les VaRs 1% seul le modèle de la copule t asymétrique dynamique affiche une fréquence de dépassement acceptable à un seuil de 10% [10]. Pour les VaRs 5%, tous les modèles respectent la fréquence de dépassement fixée. Ceci souligne l'importance d'utiliser des modèles plus sophistiqués, capable de mieux modéliser les queues des distributions, à mesure que la probabilité choisie p diminue. À l'exception de la copule t asymétrique, tous les modèles affichent de la dépendance markovienne de premier ordre pour les VaRs 1% tandis qu'à 5% seuls les modèles multivariés affichent ce type de dépendance. Comme le test de dépendance Christoffersen (1998) est très spécifique, il est intéressant de vérifier visuellement les dépassements de VaRs. Les figures 10 et 11

9. Les p-valeurs des tests de VaRs sont obtenues en simulant 999 statistiques de test plutôt qu'en utilisant la distribution asymptotique. Voir chapitre 13 dans Christoffersen (2012).

10. Tous les résultats des tests dans cette section sont interprétés à un seuil 10%. Ceci permet de minimiser la probabilité d'erreur de type II. Voir chapitre 13 dans Christoffersen (2012).

montrent la séquence des *hits* pour les VaRs 1% tandis que les figures 12 et 13 montrent la séquence des *hits* pour les VaRs 5%. On note que les dépassements de VaRs sont regroupés de manière similaire entre les modèles. Autrement dit, l'utilisation de modèles plus sophistiqués ne change pas beaucoup la structure de dépendance entre les hits. Au niveau du test conjoint de couverture conditionnelle, tous les modèles sont rejetés pour les VaRs 1% alors que pour les VaRs 5% aucun modèle n'est rejeté à l'exception de la distribution normale multivariée. Somme toute, ils semblent que pour des VaRs 5% l'utilisation de modèles univariés soit adéquat. Pour des VaRs 1% il est nécessaire d'utiliser des modèles pouvant mieux capturer les queues de distribution des rendements de portefeuille, comme la copule t asymétrique, afin d'obtenir une fréquence de *hits* convenable.

Les CVaRs de portefeuille sont *backtestées* avec les deux premiers tests de Acerbi et Szekely (2014). Dans les deux cas, l'hypothèse nulle est que la queue de gauche de la distribution du modèle est identique à celle de la vraie distribution à chaque période. Pour le premier test, l'hypothèse alternative est que les CVaRs sont sous-estimées tandis que les VaRs estimées sont égales aux VaRs de la vraie distribution à chaque période. Autrement dit, ce test suppose d'avoir préalablement accepté le modèle pour la VaR. Pour le deuxième test, l'hypothèse alternative est que les CVaRs et les VaRs estimées surestiment le risque à chaque période. Il teste donc conjointement la fréquence et l'amplitude des dépassements. Pour les deux tests, les p-valeurs sont obtenues avec 100 000 simulations de la statistique de test sous l'hypothèse nulle. Les résultats sont présentés dans les deux dernières colonnes à la table 5 pour les CVaRs à 1% et 5%. Comme aucun des modèles n'est accepté pour des VaRs 1% le deuxième est plus pertinent pour le calcul de CVaR 1%. À l'inverse, les deux tests sont pertinents pour le calcul de CVaRs 5%. À un seuil de 10% on note qu'aucun des modèles n'est retenu pour des CVaRs 1% . Pour les CVaRs 5% les résultats des deux tests sont contradictoires. Par exemple, le modèle AR(3)-NGARCH(1,1) avec la distribution t asymétrique de Hansen est le seul accepté avec le premier test , mais il est rejeté avec le deuxième test. Ces différences peuvent être dues au fait que le premier test suppose que les rendements de portefeuille sont indépendants au travers du temps ,

contrairement au deuxième test [11].

6.7 Comparaison des modèles

L'objectif de ce mémoire est de déterminer si l'utilisation de modèles multivariés est préférable à l'utilisation de modèles univariés pour la prévision de VaRs et de CVaRs. Bien que le *backtesting* permette de valider un modèle en absolu, il ne permet d'évaluer la précision relative entre plusieurs modèles. Pour comparer la précision des modèles univariés et multivariés, il nous faut une fonction de score (fonction de perte) permettant de comparer les mesures de risque estimées *ex-ante*, VaRs ou CVaRs, aux réalisations des rendements de portefeuille *ex-post*. Les fonctions de score utilisées ici sont basées sur la propriété d'élicitabilité de la VaR et la propriété d'élicitabilité jointe de la VaR et la CVaR :

$$VaR = \arg\min_{y} E\left[(X+y)(p-1_{\{X+y<0\}})\right]$$

$$\{VaR, ES\} = \arg\min_{y,z} E\left[S(y,z,X)\right]$$

avec

$$S(y,z,X) = \frac{p}{2}z^2 + \frac{Wp}{2}y^2 - pzy + \left[z(y+X) + \frac{W}{2}(X^2-y^2)\right]1_{\{X+y<0\}}.$$

En mots, la VaR et le couple VaR/CVaR minimisent l'espérance d'une fonction ayant comme arguments la ou les mesures de risques ainsi que la variable aléatoire X, ici des rendements de portefeuille. Les équivalents d'échantillon de ces deux dernières expressions sont respectivement

$$\bar{S} = \frac{1}{T}\sum_{t=1}^{T}\left[(r_{t+1}+VaR_{t+1})\left(p-1_{\{r_{t+1}+VaR_{t+1}<0\}}\right)\right]$$

et

$$\bar{S} = \frac{1}{T}\sum_{t=1}^{T} S(VaR_{t+1}, CVaR_{t+1}, r_{t+1}).$$

Pour calculer ces critères, il nous faut trois séries chronologiques : une série de VaRs, une série de CVaRs et une série de rendements. Les deux premières séries sont estimées

11. Voir Acerbi et Szekely (2014).

ex-ante tandis que la troisième est réalisée *ex-post*. L'avantage de ces deux critères est qu'ils permettent de classer les modèles entre eux. Plus petit est le score, mieux est la précision du modèle. Plus important, un modèle ayant un petit score sera préféré à un modèle ayant un score plus grand. Le premier critère permet de déterminer quel modèle offre la meilleure précision pour la prévision de VaRs tandis que le deuxième permet d'évaluer quel modèle est préférable pour la prévision jointe de VaRs et CVaRs. Les résultats sont donnés à la table 6 [12]. Pour la prévision à 1%, le modèle ayant le plus petit score à la fois pour la VaR et le couple VaR/CVaR est le modèle AR(3)-NGARCH(1,1) univarié avec la distribution t asymétrique de Hansen. À 5%, les copules Student et t asymétriques dynamiques sont préférées pour le calcul de VaRs , mais pour le calcul joint de VaRs et CVaRs le modèle de Hansen est encore le plus précis. En général, l'utilisation de copules plus sophistiquées améliore la précision des modèles par rapport à la distribution normale multivariée, mais sans l'utilisation de corrélations dynamiques les modèles multivariés sont moins précis que la distribution normale univariée.

Afin de déterminer s'il y a une différence statistique entre la précision des prévisions de chaque paire de modèles, nous utilisons le test CPA (*conditional predictive ability*) développé dans Giacomini et White (2006) [13]. L'hypothèse nulle de ce test est que, conditionnellement à l'ensemble d'information disponible à la période courante, la différence de score entre les deux modèles pour la prochaine période est nulle en moyenne :

$$H_0 : E(\Delta S_{t+1} = 0 \mid \mathscr{F}_t)$$

où ΔS_{t+1} est la différence de score entre les deux modèles à la prochaine période et $\mathscr{F}_t$ est l'ensemble d'information disponible au temps t. La statistique du test est donnée par

$$CPA = T \left(T^{-1} \sum_{t=1}^{T} h_t \, \Delta S_{t+1} \right)' \hat{\Omega}^{-1} \left(T^{-1} \sum_{t=1}^{T} h_t \, \Delta S_{t+1} \right)$$

où h_t est un vecteur $q \times 1$ contenant des variables observables au temps t et $\hat{\Omega}$ est un estimateur de la variance de $h_t \, \Delta S_{t+1}$. Suivant Giacomini et White (2006), nous posons

12. Pour le score VaR/CVaR la constante W est fixée à 2 ce qui assure que $VaR \cdot W > CVaR$ à chaque période et pour tous les modèles.

13. Ce test est aussi celui utilisé dans Santos et collab. (2013) pour comparer leurs modèles univariés et multivariés.

$h_t = (1, \Delta S_t)'$. Sous l'hypothèse nulle, cette statistique de test suit asymptotiquement une distribution χ^2_q. Les p-valeurs du test pour chaque paire de modèles et pour les différents scores à 1% et 5% sont données dans les tables 7,8,9 et 10. Les nombres entre parenthèses indiquent la fraction de périodes hors échantillon où le modèle colonne performe mieux que le modèle ligne. Pour obtenir cette fraction, nous utilisons l'approche suggérée dans Giacomini et White (2006) [14]. Une flèche vers le haut (vers la gauche) indique que l'hypothèse nulle est rejetée et que le modèle colonne (ligne) performe mieux que l'autre plus de 50% du temps. Pour la prévision de VaRs 1%, il n'y a pas de différence significative entre la distribution normale univariée et l'ensemble des modèles multivariés. Pour les VaRs 5%, la distribution normale univariée fait mieux que tous les modèles multivariés non dynamiques tandis qu'il n'y pas de différence significative avec les modèles dynamiques. Pour la prévision de VaRs et de CVaRs 1%, il n'y a pas de différence significative entre la distribution normale univariée et chacun des autres modèles. À 5%, les conclusions sont similaires. Dans tous les cas l'utilisation de la distribution t asymétrique de Hansen améliore significativement la prévision du risque par rapport à la distribution normale univariée et aux modèles multivariés non dynamiques.

14. On régresse ΔS_{t+1} sur h_t avec les observations hors échantillon puis on calcule la fraction avec $T^{-1}\sum_{t=1}^{T} 1_{\{\hat{\delta}' h_t > 0\}}$ où $\hat{\delta}$ est le vecteur contenant les coefficients estimés de la régression.

Conclusion

Ce mémoire vise à comparer une approche univariée à une approche multivariée pour la prévision de VaRs et de CVaRs de portefeuille sur une période. Nous effectuons une extension de l'étude de Christoffersen et Langlois (2013) en utilisant leur modélisation de la dynamique jointe des trois facteurs de Fama et French (1993) et du facteur de Carhart (1997). Au total, deux modèles univariés sont comparés à huit modèles multivariés impliquant des copules et des corrélations dynamiques. Par des simulations, nous générons 1 436 VaRs et CVaRs hors échantillon allant de 1983 à 2010. Nous effectuons le backtesting des VaRs et CVaRs de chacun des modèles en comparant ces mesures de risques *ex-ante* aux réalisations de rendements de portefeuille *ex-post*. Les modèles sont comparés avec le test CPA de Giacomini et White (2006) en exploitant l'élicitabilité de la VaR et l'élicitabilité jointe du couple VaR/CVaR.

Nous trouvons qu'il n'y a pas de différences significatives entre la précision des modèles univariés et celle des modèles multivariés avec DCC pour la prévision de VaRs et de CVaRs hors échantillon. À l'inverse, nous trouvons qu'il y a des différences de précision significatives entre les modèles univariés et les modèles multivariés sans DCC. Ces différences sont toutes en faveur des modèles univariés. Aussi, l'utilisation de corrélations dynamiques améliore significativement la précision des modèles multivariés.

Les deux premiers résultats sont en opposition avec les conclusions de Santos et collab. (2013). Plusieurs raisons peuvent expliquer cette divergence. Premièrement, notre application se limite à quatre actifs tandis que Santos et collab. (2013) utilisent des portefeuilles composés de plusieurs dizaines d'actifs. Un plus grand nombre d'actifs com-

plique l'estimation d'un modèle multivarié, mais prend en compte un plus grand nombre de dépendances. Une extension possible du mémoire serait de modéliser la distribution jointe d'un plus grand portefeuille à partir de l'exposition des actifs aux quatre facteurs. Deuxièmement, nous utilisons des rendements hebdomadaires tandis que Santos et collab. (2013) utilisent des rendements quotidiens. Il serait intéressant de vérifier si nos conclusions tiennent pour d'autres fréquences. Troisièmement, nos poids de portefeuille sont fixés ce qui peut favoriser l'approche univariée étant donné que les pseudo rendements coïncident avec les rendements réalisés ex-post. Quatrièmement, notre processus de prévision est différent et couvre une période différente. En effet, nous utilisons une fenêtre d'estimation en expansion alors que Santos et collab. (2013) utilisent une fenêtre d'estimation roulante. Notre période hors échantillon va de 1983 à 2010 tandis que la leur va de 2006 à 2010. Cinquièmement, nos modèles sont différents. Nos résultats sont toutefois en accord avec Santos et collab. (2013) au niveau de l'amélioration des modèles multivariés lorsqu'on utilise des corrélations dynamiques.

Bibliographie

Acerbi, C. et B. Szekely. 2014, «Backtesting expected shortfall», *Risk*, p. 76.

Aielli, G. P. 2013, «Dynamic conditional correlation : on properties and estimation», *Journal of Business & Economic Statistics*, vol. 31, n° 3, p. 282–299.

Artzner, P., F. Delbaen, J.-M. Eber et D. Heath. 1999, «Coherent measures of risk», *Mathematical finance*, vol. 9, n° 3, p. 203–228.

Bollerslev, T. 1986, «Generalized autoregressive conditional heteroskedasticity», *Journal of Econometrics*, vol. 31, n° 3, p. 307–327.

Brooks, C. et G. Persand. 2003, «Volatility forecasting for risk management», *Journal of forecasting*, vol. 22, n° 1, p. 1–22.

Carhart, M. M. 1997, «On persistence in mutual fund performance», *The Journal of Finance*, vol. 52, n° 1, p. 57–82.

Christoffersen, P. 2009, «Value–at–risk models», dans *Handbook of Financial Time Series*, Springer, p. 753–766.

Christoffersen, P. et H. Langlois. 2013, «The joint dynamics of equity market factors», *Journal of Financial and Quantitative Analysis*, vol. 48, n° 05, p. 1371–1404.

Christoffersen, P. F. 1998, «Evaluating interval forecasts», *International Economic Review*, p. 841–862.

Christoffersen, P. F. 2012, *Elements of Financial Risk Management*, Academic Press.

Demarta, S. et A. J. McNeil. 2005, «The t copula and related copulas», *International Statistical Review/Revue Internationale de Statistique*, vol. 73, n° 1, p. 111–129.

Duffie, D. et J. Pan. 1997, «An overview of value at risk», *The Journal of Derivatives*, vol. 4, n° 3, p. 7–49.

Engle, R. 2002, «Dynamic conditional correlation : A simple class of multivariate generalized autoregressive conditional heteroskedasticity models», *Journal of Business & Economic Statistics*, vol. 20, n° 3, p. 339–350.

Engle, R. F. 1982, «Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation», *Econometrica*, vol. 50, n° 4, p. 987–1007.

Engle, R. F. et V. K. Ng. 1993, «Measuring and testing the impact of news on volatility», *The Journal of Finance*, vol. 48, n° 5, p. 1749–1778.

Fama, E. F. et K. R. French. 1993, «Common risk factors in the returns on stocks and bonds», *Journal of Financial Economics*, vol. 33, n° 1, p. 3–56.

Giacomini, R. et H. White. 2006, «Tests of conditional predictive ability», *Econometrica*, vol. 74, n° 6, p. 1545–1578.

Gneiting, T. 2011, «Making and evaluating point forecasts», *Journal of the American Statistical Association*, vol. 106, n° 494, p. 746–762.

Hansen, B. E. 1994, «Autoregressive conditional density estimation», *International Economic Review*, vol. 35, n° 3, p. 705–730.

Kupiec, P. H. 1995, «Techniques for verifying the accuracy of risk measurement models», *The Journal of Derivatives*, vol. 3, n° 2, p. 73–84.

McNeil, A. J., R. Frey et P. Embrechts. 2015, *Quantitative Risk Management : Concepts, Techniques and Tools*, Princeton University Press.

Osband, K. H. 1985, *Providing Incentives for Better Cost Forecasting*, thèse de doctorat, University of California, Berkeley.

Patton, A. J. 2006, «Modelling asymmetric exchange rate dependence», *International Economic Review*, vol. 47, n° 2, p. 527–556.

Santos, A. A., F. J. Nogales et E. Ruiz. 2013, «Comparing univariate and multivariate models to forecast portfolio value-at-risk», *Journal of Financial Econometrics*, vol. 11, n° 2, p. 400–441.

Simonato, J.-G. 2011, «The performance of johnson distributions for computing value at risk and expected shortfall», *The Journal of Derivatives*, vol. 19, n° 1, p. 7–24.

Sklar, M. 1959, *Fonctions de répartition à n dimensions et leurs marges*, Université Paris 8.

Annexe A – Copule t asymétrique de Demarta et McNeil (2005)

La copule t asymétrique de Demarta et McNeil (2005) est basée sur une paramétrisation de la distribution t asymétrique multivariée dont la densité est donnée par

$$f(x) = c\frac{K_{\frac{v_c+N}{2}}\left(\sqrt{\left(v_c+(x-\mu)^\top\Sigma^{-1}(x-\mu)\right)\lambda^\top\Sigma^{-1}\lambda}\right)\exp\left(x-\mu)^\top\Sigma^{-1}\lambda\right)}{\left(\sqrt{\left(v_c+(x-\mu)^\top\Sigma^{-1}(x-\mu)\right)\lambda^\top\Sigma^{-1}\lambda}\right)^{-\frac{v_c+N}{2}}\left(1+\frac{(x-\mu)^\top\Sigma^{-1}(x-\mu)}{v_c}\right)^{\frac{v_c+N}{2}}}$$

où $K_\gamma(\cdot)$ est la fonction Bessel modifiée de troisième espèce et c est une constante définie par

$$c = \frac{2^{\frac{2-(v_c+N)}{2}}}{\Gamma\left(\frac{v_c}{2}\right)(\pi v_c)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}}.$$

Aussi, λ est un vecteur $N\times 1$ contenant des paramètres d'asymétrie, v_c est un paramètre scalaire de degrés de liberté, μ est un vecteur $N\times 1$ contenant des paramètres de location et Σ est une matrice $N\times N$ définie positive contenant des paramètres de dispersion.

Les deux premiers moments de x sont donnés par

$$E(x) = \mu + \frac{v_c}{v_c-2}\lambda$$

$$Cov(x) = \frac{v_c}{v_c-2}\Sigma + \frac{2v_c^2}{(v_c-2)^2(v_c-4)}\lambda^\top\lambda.$$

La distribution t asymétrique peut être simulée avec la représentation stochastique suivante :

$$X = \mu + \sqrt{W}Y + \lambda W$$

où W ou une variable inverse-gamma, $W \sim IG(\frac{v_c}{2}, \frac{v_c}{2})$, Y est un vecteur normal $N \times 1$, $Y \sim N(0, \Sigma)$, et Y et W sont indépendants.

Christoffersen et Langlois (2013) utilisent une version standardisée de cette distribution en posant $\mu = 0$ et $\Sigma = \Upsilon$ une matrice de corrélation. La densité de la distribution devient donc

$$f(z) = c \frac{K_{\frac{v_c+N}{2}}\left(\sqrt{\left(v_c + z^\top \Upsilon^{-1} z\right) \lambda^\top \Sigma^{-1} \lambda}\right) \exp\left(z^\top \Upsilon^{-1} \lambda\right)}{\left(\sqrt{\left(v_c + z^\top \Upsilon^{-1} z\right) \lambda^\top \Upsilon^{-1} \lambda}\right)^{-\frac{v_c+N}{2}} \left(1 + \frac{z^\top \Upsilon^{-1} z}{v_c}\right)^{\frac{v_c+N}{2}}}$$

avec

$$E(z) = \frac{v_c}{v_c - 2} \lambda \tag{1}$$

$$Cov(z) = \frac{v_c}{v_c - 2} \Upsilon + \frac{2v_c^2}{(v_c-2)^2(v_c-4)} \lambda^\top \lambda. \tag{2}$$

La distribution t asymétrique standardisée peut être simulée avec la représentation stochastique suivante :

$$Z = \sqrt{W} Y + \lambda W \tag{3}$$

avec $W \sim IG(\frac{v_c}{2}, \frac{v_c}{2})$, $Y \sim N(0, \Upsilon)$, et Y et W indépendants.

Rappelons qu'une copule est une fonction de répartition multivariée d'un vecteur aléatoire U de dimensions $N \times 1$ dont les distributions marginales sont uniformes :

$$C(u) = P(U_1 \leq u_1, \ldots, U_N \leq u_N).$$

Maintenant, nous pouvons construire une copule à partir d'une fonction de répartition multivariée $F(\cdot)$ à l'aide de la relation suivante [15] :

$$C(u) = F\left(F_1^{-1}(u_1), \ldots, F_N^{-1}(u_N)\right)$$

où $F_i^{-1}(\cdot)$ $i = 1, \ldots, N$ sont les fonctions de répartition inverses marginales . La densité de la copule est ensuite donnée par

$$c(u) = \frac{\partial^N}{\partial u_1 \ldots \partial u_N} C(u) = \frac{f\left(F_1^{-1}(u_1), \ldots, F_N^{-1}(u_N)\right)}{\prod_{i=1}^N f_i\left(F_i^{-1}(u_i)\right)} \tag{4}$$

15. Voir McNeil et collab. (2015).

où $f(\cdot)$ est la fonction de densité multivariée et $f_i(\cdot)$ $i = 1,\ldots,N$ sont les fonctions de densité marginales. En remplaçant la fonction de densité multivariée et les fonctions de densité marginales de la distribution t asymétrique standardisée dans (4) on obtient [16]

$$
\begin{aligned}
c(u) = & \frac{2^{\frac{(v_c-2)(N-1)}{2}} K_{\frac{v_c+N}{2}}\left(\sqrt{\left(v_c + z^\top \Upsilon^{-1} z\right)\lambda^\top \Upsilon^{-1}\lambda}\right)\exp\left(z^\top \Upsilon^{-1}\lambda\right)}{\Gamma\left(\frac{v_c}{2}\right)^{1-N}|\Upsilon|^{\frac{1}{2}}\left(\sqrt{\left(v_c + z^\top \Upsilon^{-1} z\right)\lambda^\top \Upsilon^{-1}\lambda}\right)^{-\frac{v_c+N}{2}}\left(1+\frac{z^\top \Upsilon^{-1} z}{v_c}\right)^{\frac{v_c+N}{2}}} \\
& \times \prod_{i=1}^{N} \frac{\left(\sqrt{\left(v_c + z_i^2\right)\lambda_i^2}\right)^{-\frac{v_c+1}{2}}\left(1+\frac{z_i^2}{v_c}\right)^{\frac{v_c+1}{2}}}{K_{\frac{v_c+1}{2}}\left(\sqrt{\left(v_c + z_i^2\right)\lambda_i}\right)\exp\left(z_i\lambda_i\right)}
\end{aligned}
\tag{5}
$$

où nous définissons les fractiles $z_i = F_i^{-1}(u_i)$ $i = 1,\ldots,N$ avec $F_i^{-1}(\cdot)$ la fonction de répartition inverse marginale de la distribution t asymétrique. Comme il n'y a pas de solution analytique pour $F_i^{-1}(\cdot)$, nous procédons par simulation pour le calcul des fractiles. Nous simulons 100 000 fractiles de la distribution t asymétrique marginale i à l'aide la représentation stochastique (3) puis nous calculons les fractiles par interpolation linéaire.

La densité (5) est celle utilisée pour l'estimation par maximum de vraisemblance des paramètres de la copule t asymétrique. Les densités des copules normales et Student sont obtenues de façon similaire, c'est-à-dire en utilisant la relation (4) [17].

16. La copule sera identique, peut importe si l'on utilise la distribution standardisée ou non standardisée. En effet, une copule est invariante à la standardisation des distributions marginales. Voir Demarta et McNeil (2005).

17. Voir chapitre 9 dans Christoffersen (2012).

Annexe B – Standardisation des fractiles et ciblage de corrélation dans le DCC

Dans Christoffersen et Langlois (2013), la copule conditionnelle est faite dynamique en modélisant la matrice de corrélation conditionnelle des fractiles Ψ_t avec le processus DCC de Engle (2002). Plus précisément, ils utilisent une version modifiée du DCC développée dans Aielli (2013) :

$$Q_t = Q(1-\beta_c-\alpha_c)+\beta_c Q_{t-1}+\alpha_c\{Q_{t-1}^{*\frac{1}{2}} z_{t-1} z_{t-1}^{\top} Q_{t-1}^{*\frac{1}{2}}\} \tag{6}$$

où $Q_t^* = diag(dg(Q_t))$ [18] et z_{t-1} le vecteur $N \times 1$ contenant les fractiles standardisés, c'est-à-dire de moyenne nulle et de variance unitaire [19]. La matrice de corrélation conditionnelle est ensuite obtenue en normalisant les éléments dans la matrice Q_t :

$$\Psi_t = diag(dg(Q_t))^{-\frac{1}{2}} Q_t diag(dg(Q_t))^{-\frac{1}{2}}$$

Pour réduire le nombre de paramètres à estimer, nous pouvons cibler la matrice de corrélation non conditionnelle Q avec l'estimateur d'échantillon $\frac{1}{T}\sum_{t=1}^{T} Q_t^{*\frac{1}{2}} z_t z_t^{\top} Q_t^{*\frac{1}{2}}$. Le problème est que Q est nécessaire pour calculer la série chronologique Q_t^*. Pour contourner cette circularité Aielli (2013) propose de calculer la série chronologique (6) pour les

18. L'opérateur $dg(A)$ renvoie un vecteur contenant les éléments de la diagonale de la matrice A.
19. Pour standardiser les z_t, il suffit de soustraire le vecteur d'espérance (1) puis de diviser (élément par élément) par la racine des variances sur la diagonale dans (2).

éléments de la diagonale seulement :

$$q_{ii,t} = (1-\beta_c-\alpha_c)+\beta_c q_{ii,t-1}+\alpha_c z_{i,t-1}^2 q_{ii,t-1}$$

pour $i = 1,\ldots,N$ et $t = 1,\ldots,T$. Ceci nous permet de calculer $Q_t^* = diag(q_{11,t},\ldots,q_{NN,t})$ et de cibler Q.

Pour la copule t asymétrique, la matrice de corrélation conditionnelle des fractiles Ψ_t n'est pas égale à la matrice de corrélation de la copule Υ_t. Le lien entre les deux est donné dans (2) :

$$D\Psi_t D = \frac{v_c}{v_c-2}\Upsilon_t + \frac{2v_c^2}{(v_c-2)^2(v_c-4)}\lambda^\top\lambda.$$

où D est une matrice diagonale contenant les écarts-types des fractiles. En isolant Υ_t on obtient

$$\Upsilon_t = \frac{v_c-2}{v_c}\left(D\Psi_t D - \frac{2v_c^2}{(v_c-2)^2(v_c-4)}\lambda^\top\lambda\right).$$

C'est cette matrice de corrélation conditionnelle qui entre dans la vraisemblance (5) lors de l'estimation. L'annexe C donne un exemple de fonction MATLAB pour le calcul de la log-vraisemblance de la copule t asymétrique dynamique.

Annexe C – Fonction MATLAB pour calculer la log-vraisemblance de la copule t asymétrique dynamique

```
function [fv]=f_copule_tasym(param,Z,Support,u)

% Param est un vecteur (N+3) x 1 contenant les paramètres
de la copule t asymétrique dynamique.
% Z est un vecteur 100 000 x 1 contenant des réalisations de
variables aléatoires normales standards.
% Support est un vecteur 100 002 x 1 allant de 0 à 1 par
incréments de 0.5/100 000.
% u est une matrice T x N où T est le nombre d'observations
et N est le nombre d'actifs.

%Extraction des paramètres

[T,N]=size(u);              % Dimensions de u
nu=abs(param(1));           % Paramètre de degrés de liberté
lambda=param(2:N+1)';       % N paramètres d'asymétrie
alpha=abs(param(N+2));      % Paramètre alpha du DCC
```

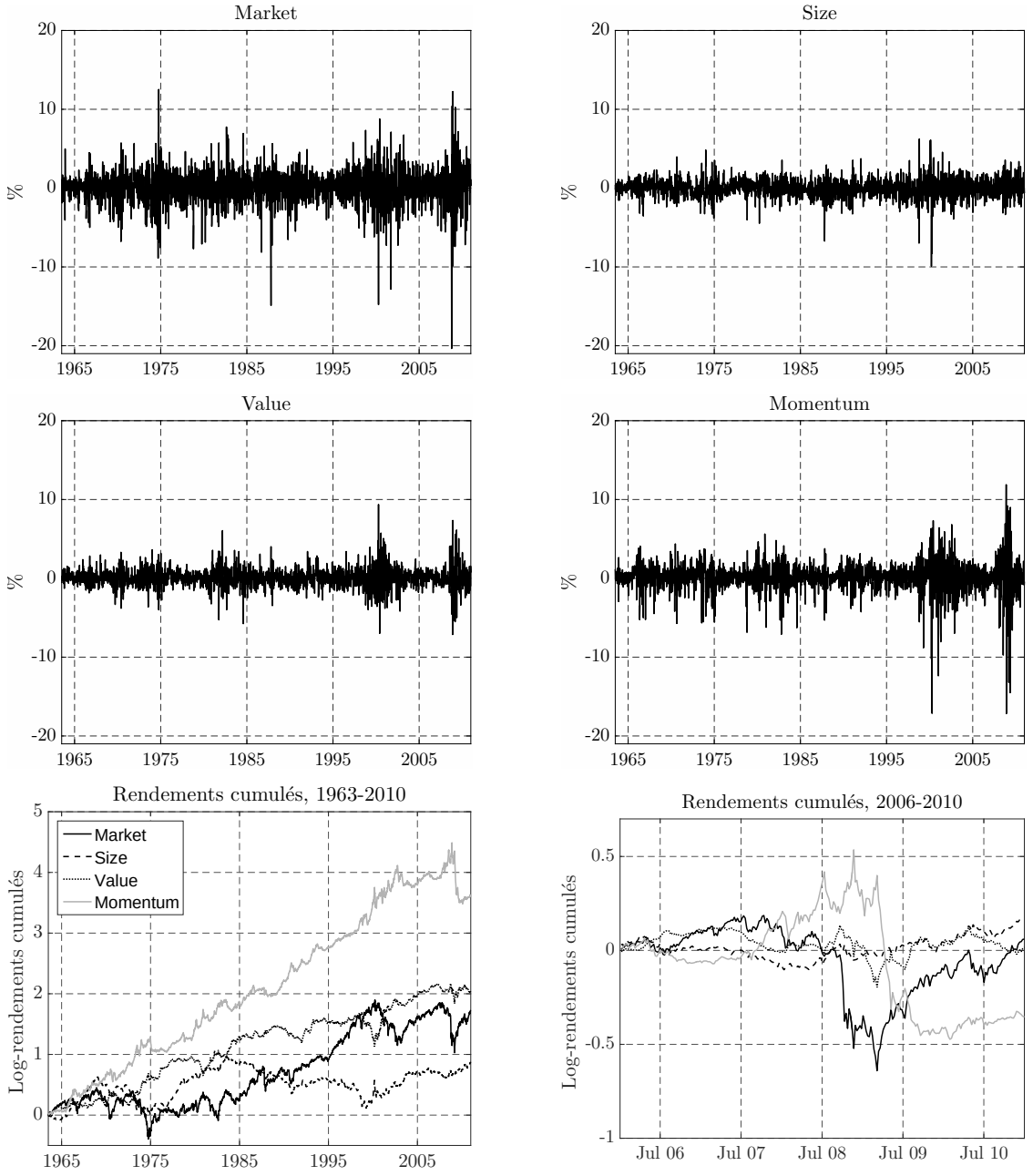
```
beta=abs(param(N+3));      % Paramètre beta du DCC

%Creation des matrices de corrélation

Psi=zeros(N,N,T);
Q=zeros(N,N,T);

%Calcul de l'espérance et de l'écart-type des fractiles

mu=nu*lambda./(nu-2);
std= sqrt( nu/(nu-2)+(2*nu^2.*lambda.^2)./((nu-2)^2*(nu-4)) )  ;

% Calcul des fractiles

%Nombre de simulations pour l'inversion des probabilités(u)

nsimul=100000;

rng('default');

W = 0.5 * nu ./ randg( 0.5 * nu, nsimul, 1 );
WZ = sqrt( W ) .* Z;

for i = 1:N

%Simulation de la distribution t asymétrique marginale i

X=lambda(i ) * W + WZ;
```

```
%Interpolation linéaire des fractiles

X=sort( X, 1 );
x( :, i )= interp1( Support, [ X( 1, : ); X( :, : ); X( nsimul, : ) ],
u( :, i ) );

%Standardisation des fractiles

z(:,i)=(x(:,i)-mu(i))./std(i);

end

%Correction d'Aielli

p=zeros(T,4);
p(1,:)=ones(1,4);

for t=2:T

p(t,:)= (1-alpha-beta)+beta.*p(t-1,:)+alpha.* z(t-1,:).^2;
z(t,:)= z(t,:).*sqrt(p(t,:));
end

%Ciblage des corrélations non conditionnelles des factiles

S=corrcoef(z);

%Matrice de corrélations non conditionnelles de la copule
```

```
C= (diag(std)*S*diag(std) - 2*(nu^2)*(lambda'*lambda)./((
(nu-2)^2)*(nu-4)) ) ./ (nu/(nu-2));

Q(:,:,1)=S;
Psi(:,:,1)=C;

%Calcul des matrices de corrélations conditionnelles du DCC et
des matrices de corrélations conditionnelles de la copule

for t=2:T

Q(:,:,t)=S.*(1-alpha-beta)+ alpha.*z(t-1,:)'*z(t-1,:)
+ beta.* Q(:,:,t-1);
Psi(:,:,t)= diag(diag(Q(:,:,t)))^-0.5 * Q(:,:,t)
* diag(diag(Q(:,:,t)))^-0.5;
Psi(:,:,t)=(diag(std)*Psi(:,:,t)*diag(std)- 2*(nu^2)*
(lambda'*lambda)./(((nu-2)^2)*(nu-4)) ) ./ (nu/(nu-2));
end

%Calcul de la log-vraisemblance

ll=zeros(T,1);

for t=1:T

arg1=(nu+N)*0.5;
arg2= sqrt((nu+x(t,1:N)*inv(Psi(:,:,t))*x(t,1:N)')*lambda
*inv(Psi(:,:,t))*lambda');
```

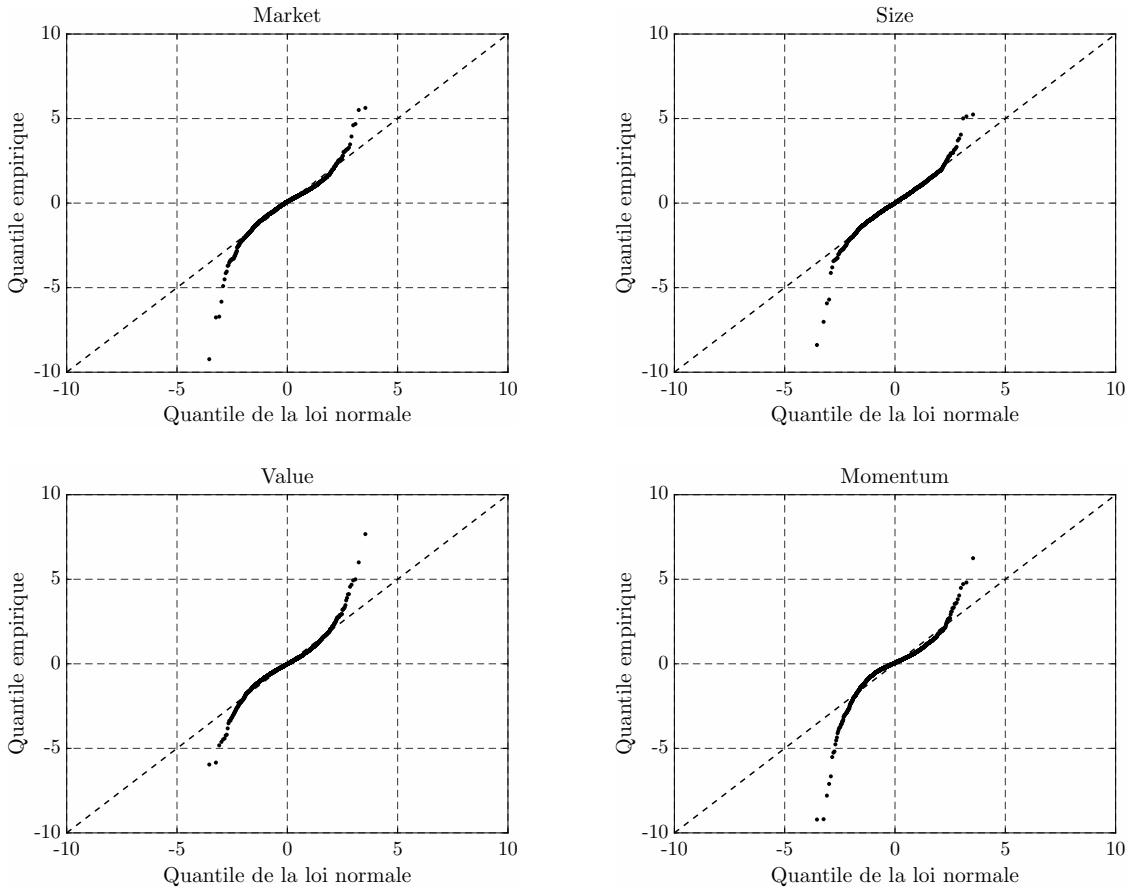
```
NUM=2^((nu-2)*(N-1)*0.5)*besselk(arg1,arg2)*exp(x(t,1:N)
*inv(Psi(:,:,t))*lambda');

DENOM= gamma(nu/2)^(1-N)* det(Psi(:,:,t))^0.5 * arg2^(-arg1)
* (1+(1/nu)*x(t,1:N)*inv(Psi(:,:,t))*x(t,1:N)')^(arg1);

arg3= sqrt(((nu + x(t,1:N).^2).*lambda.^2));
arg4=0.5*(nu+1);

PROD= prod( (arg3.^(-arg4).*(1+x(t,1:N).^2 ./nu).^(arg4) )
./ (besselk(arg4,arg3).* exp(x(t,1:N).* lambda)) );

ll(t,1)= log(NUM)-log(DENOM)+log(PROD);

end

fv=sum(ll);

end
```

FIGURE 1 – Séries chronologiques des rendements des facteurs



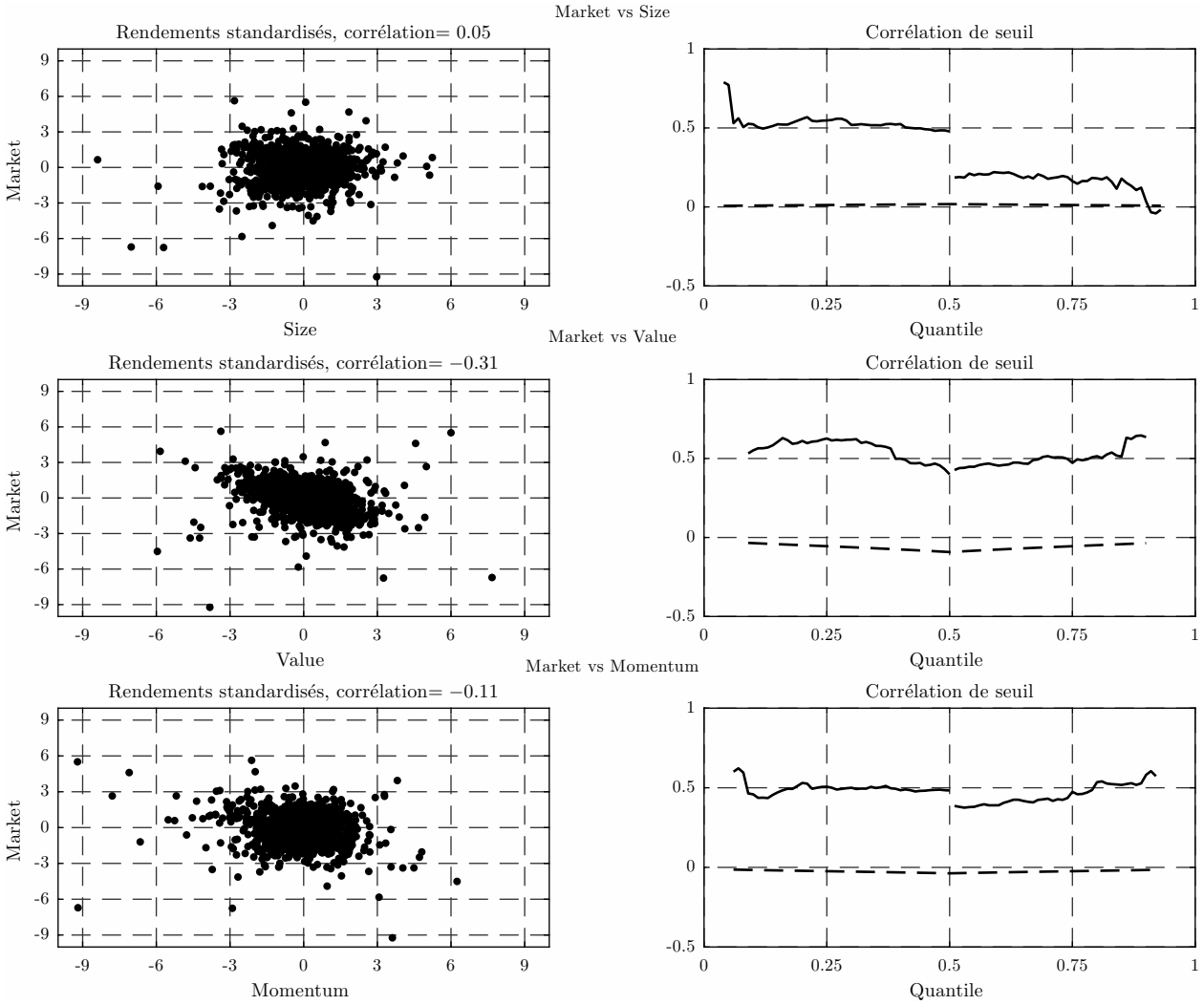
Notes : Les graphiques du haut et du milieu montrent les séries chronologiques des log-rendements hebdomadaires des quatre facteurs pour la période allant du 5 juillet 1963 au 31 décembre 2010. Les graphiques du bas montrent les log-rendements hebdomadaires cumulés pour la période allant du 5 juillet 1963 au 31 décembre 2010 (à gauche) et la période allant du 5 janvier 2006 au 31 décembre 2010 (à droite).

FIGURE 2 – Diagrammes Quantile-Quantile des rendements des facteurs, 1963-2010



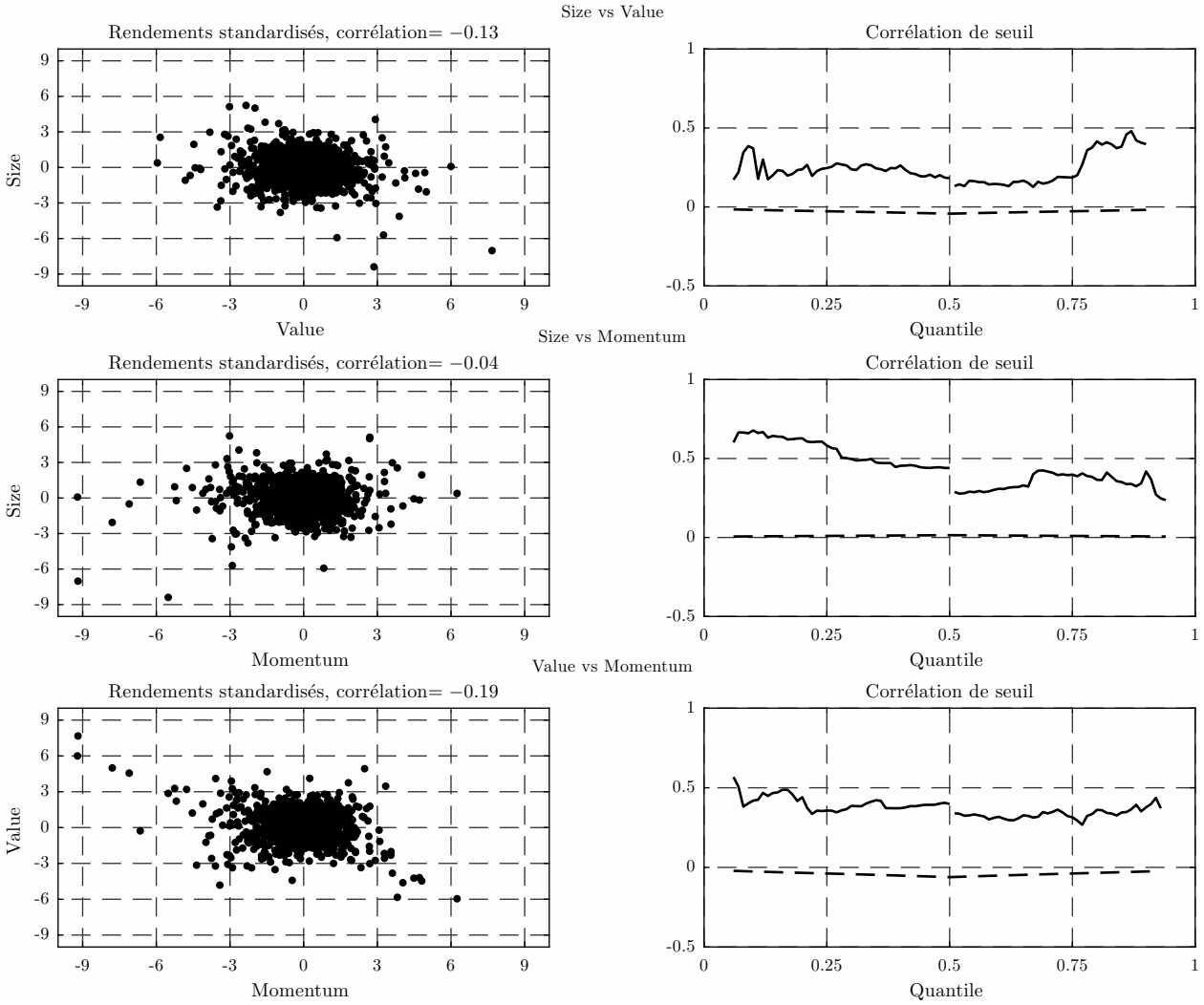
Notes : Chaque quantile empirique des rendements standardisés (axe vertical) est comparé à celui de la distribution normale centrée réduite(axe horizontal). Si les rendements sont normalement distribués, toutes les observations devraient se situer sur la ligne de 45 degrés.

FIGURE 3 – Diagrammes de dispersion et corrélations de seuil pour les paires de facteurs incluant le facteur *Market*



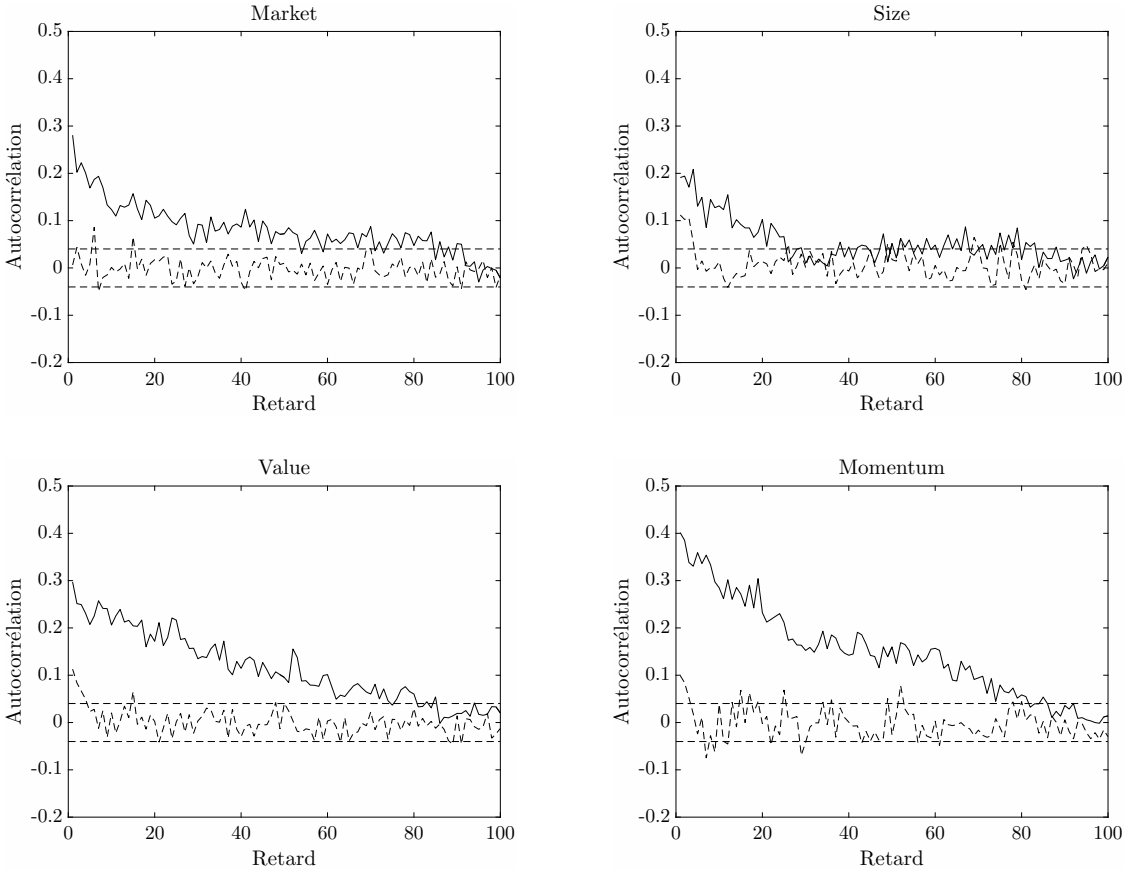
Notes : Les graphiques de gauche montrent les diagrammes de dispersion des rendements standardisés des paires de facteurs incluant le facteur Market. Les rendements sont des log-rendements hebdomadaires et couvrent la période allant du 5 juillet 1963 au 31 décembre 2010. Les graphiques de droite montrent les corrélations de seuil entre les paires facteurs. La ligne pleine indique la corrélation empirique entre les paires de facteurs lorsque ceux-ci sont en dessous ou au-dessus de la médiane. La ligne pointillée indique la corrélation impliquée par la distribution normale bivariée avec le même coefficient de corrélation entre les facteurs.

FIGURE 4 – Diagrammes de dispersion et corrélations de seuil pour les paires de facteurs excluant le facteur *Market*



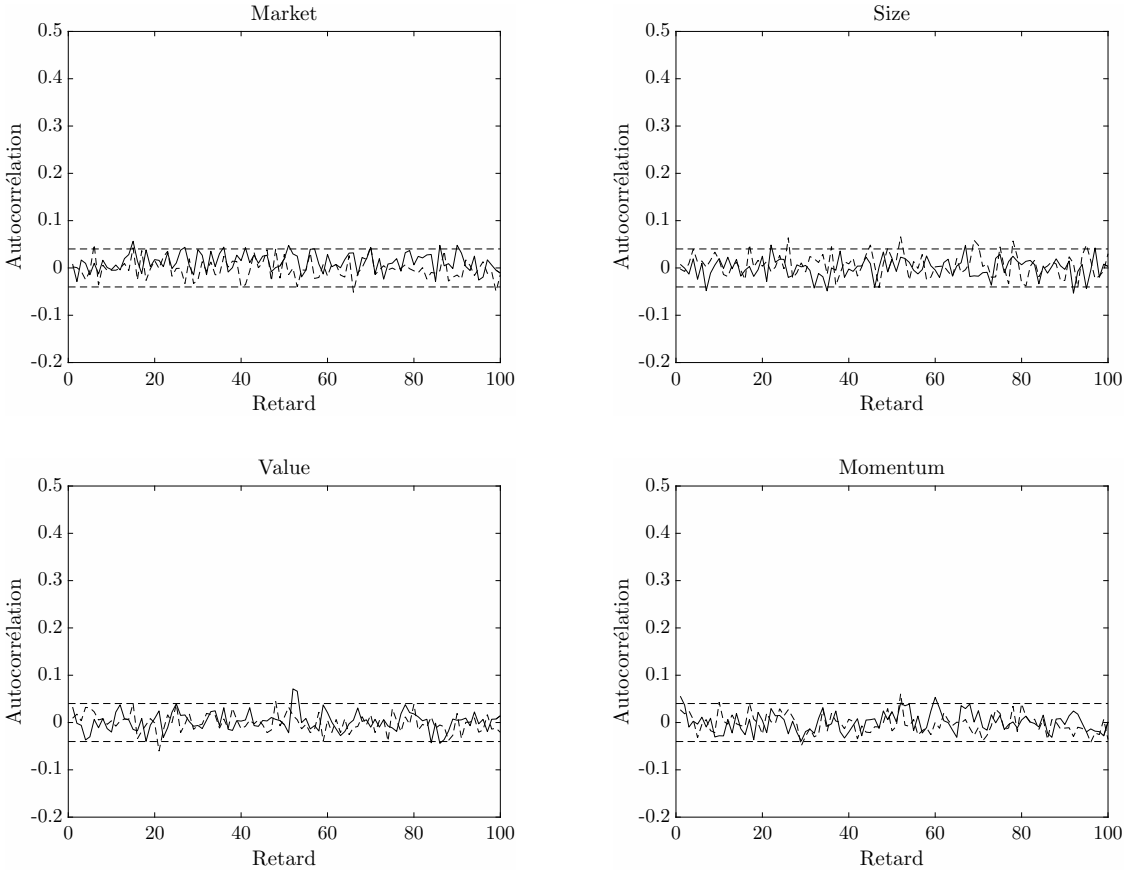
Notes : Les graphiques de gauche montrent les diagrammes de dispersion des rendements standardisés des paires de facteurs excluant le facteur Market . Les rendements sont des log-rendements hebdomadaires et couvrent la période allant du 5 juillet 1963 au 31 décembre 2010. Les graphiques de droite montrent les corrélations de seuil entre les paires facteurs. La ligne pleine indique la corrélation empirique entre les paires de facteurs lorsque ceux-ci sont en dessous ou au-dessus de la médiane. La ligne pointillée indique la corrélation impliquée par la distribution normale bivariée avec le même coefficient de corrélation entre les facteurs.

FIGURE 5 – Autocorrélations des rendements et des rendements absolus



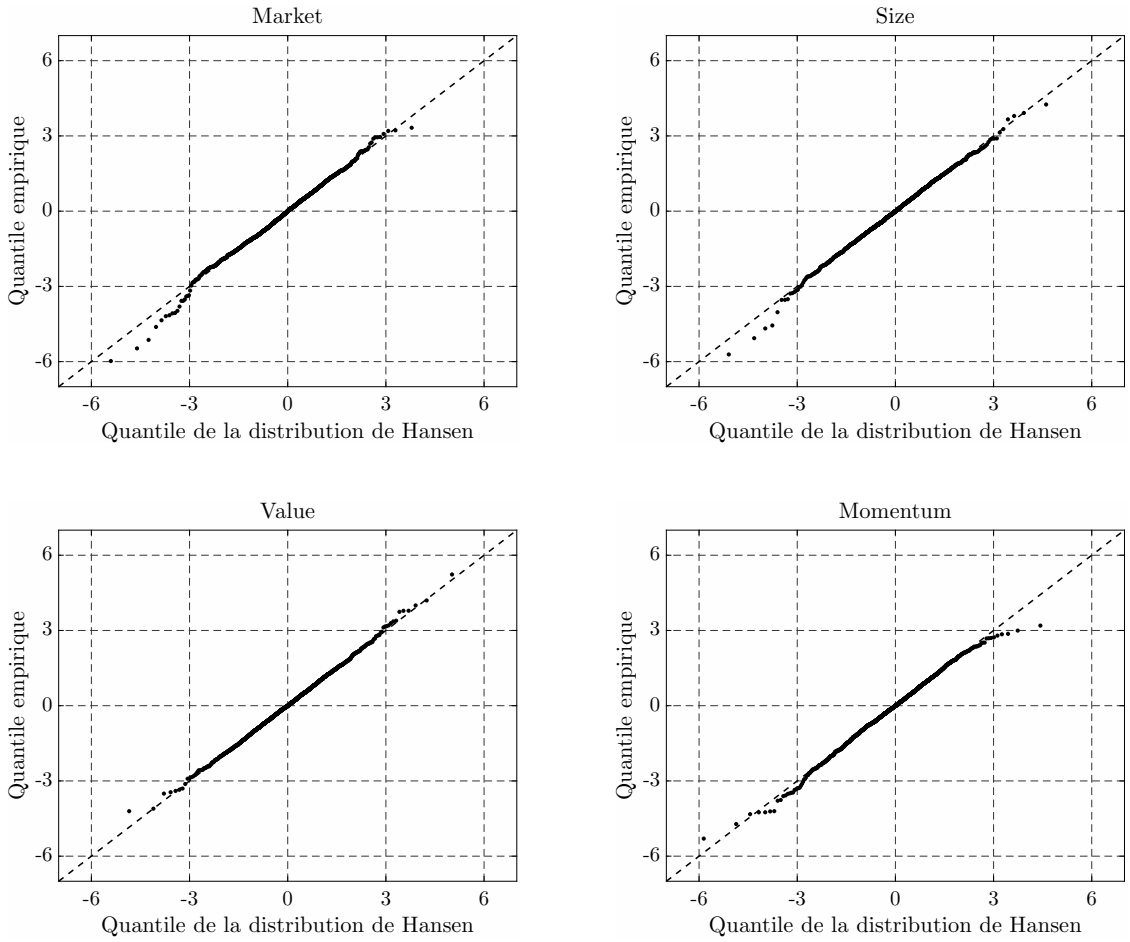
Notes : Autocorrélations des log-rendements hebdomadaires (ligne pointillée) et des log-rendements hebdomadaires absolus (ligne pleine) pour la période allant du 5 juillet 1963 au 31 décembre 2010. Les lignes pointillées horizontales indiquent les bornes pour un intervalle de confiance de 95% autour de zéro.

FIGURE 6 – Autocorrélations des résidus et des résidus absolus



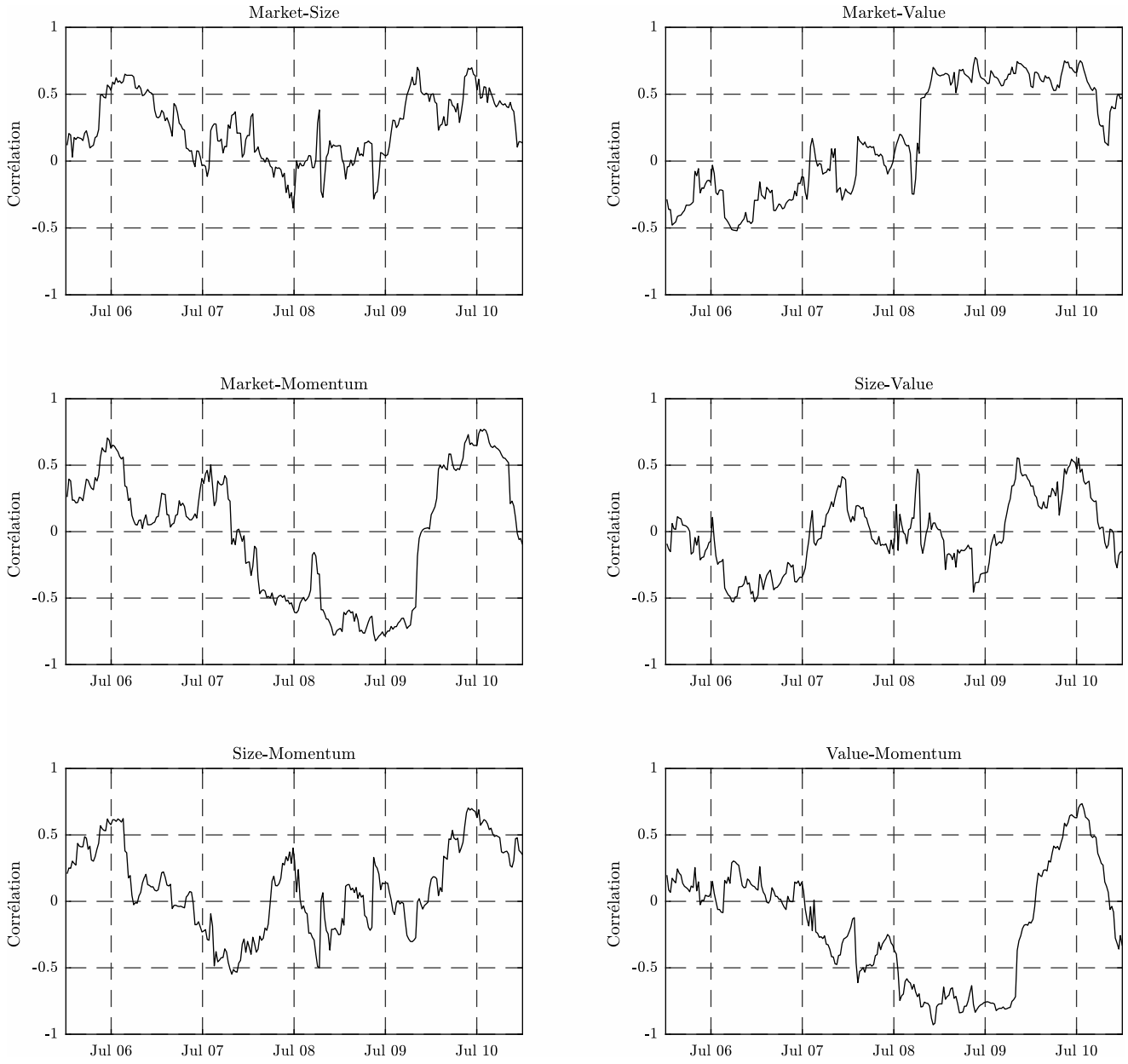
Notes : Autocorrélations des résidus(ligne pointillée) et des résidus absolus (ligne pleine) pour la période allant du 5 juillet 1963 au 31 décembre 2010. Les résidus sont obtenus par l'estimation du modèle AR(3)-NGARCH(1,1) avec une distribution normale. Les lignes pointillées horizontales indiquent les bornes pour un intervalle de confiance de 95% autour de zéro.

FIGURE 7 – Diagrammes Quantile-Quantile pour les résidus du modèle AR(3)-NGARCH(1,1) avec la distribution de Hansen



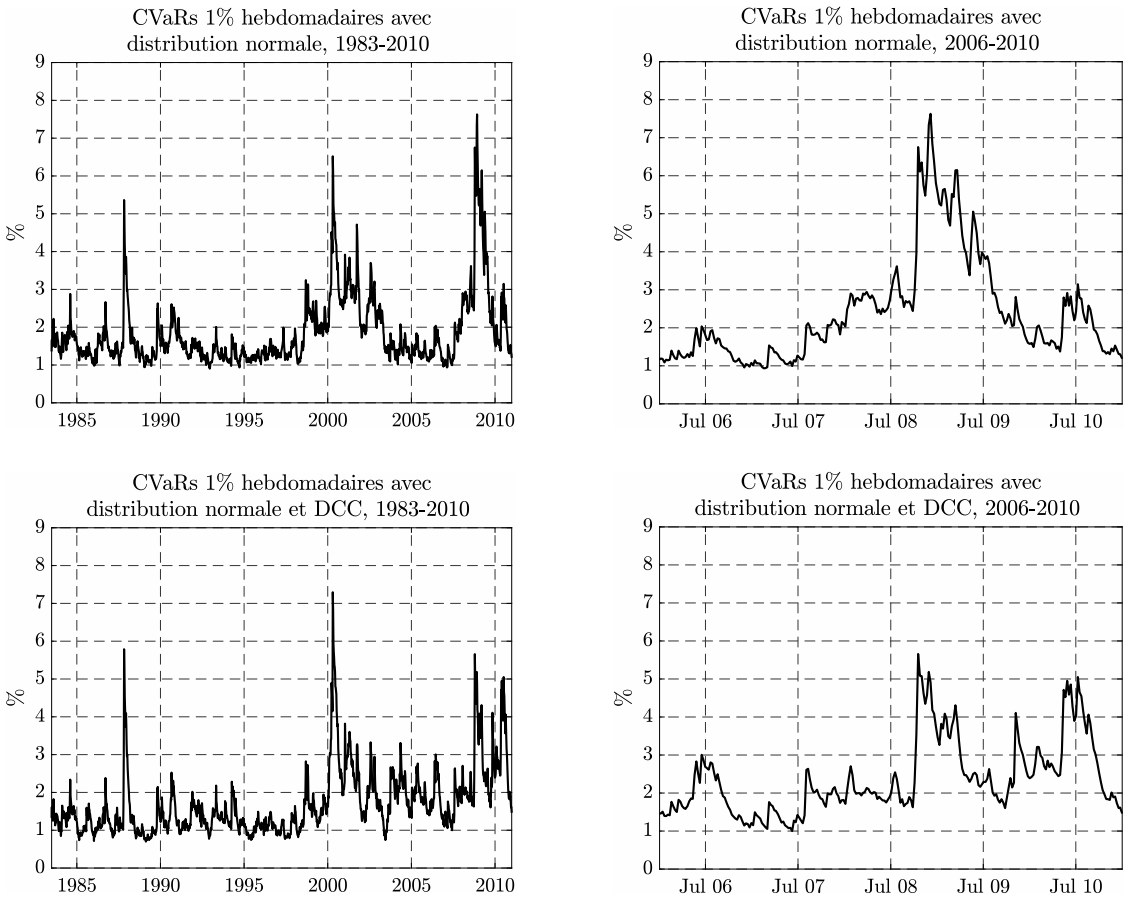
Notes : Chaque quantile empirique des résidus du modèle AR(3)-NGARCH(1,1) (axe vertical) est comparé à celui de la distribution t asymétrique de Hansen (axe horizontal). Si les résidus suivent la distribution t asymétrique de Hansen, toutes les observations devraient se situer sur la ligne de 45 degrés.

FIGURE 8 – Corrélations dynamiques de la copule t asymétrique, 2006-2010



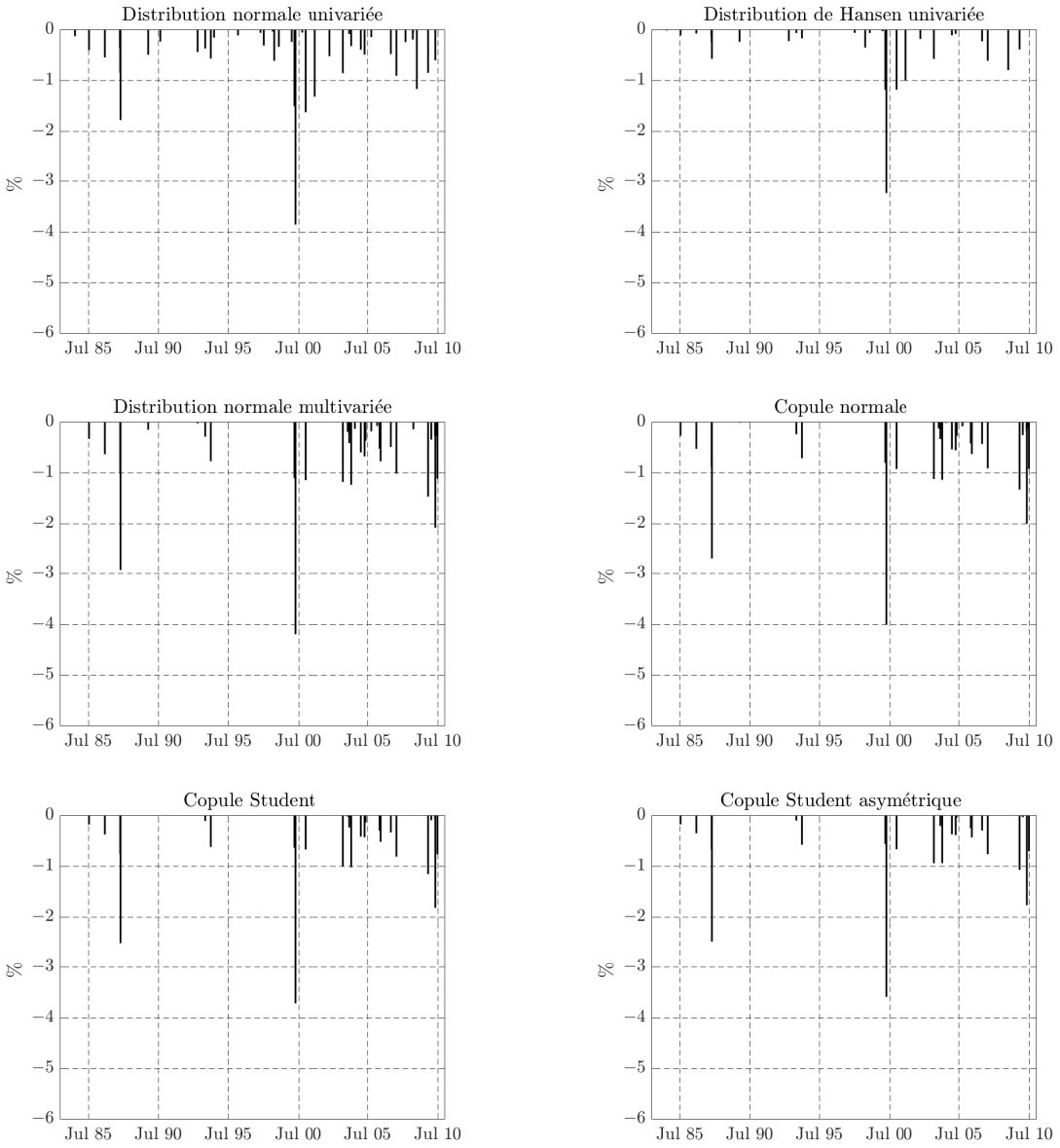
Notes : La figure présente les corrélations conditionnelles dynamiques de chaque paire de facteurs pour la période allant de janvier 2006 à décembre 2010. Les corrélations sont obtenues en estimant la copule t asymétrique dynamique sur les résidus des modèles AR(3)-NGARCH(1,1). Toutes les observations de 1963 à 2010 sont utilisées pour l'estimation.

FIGURE 9 – Prévisions de CVaRs hors échantillon pour un portefeuille équipondéré des quatre facteurs



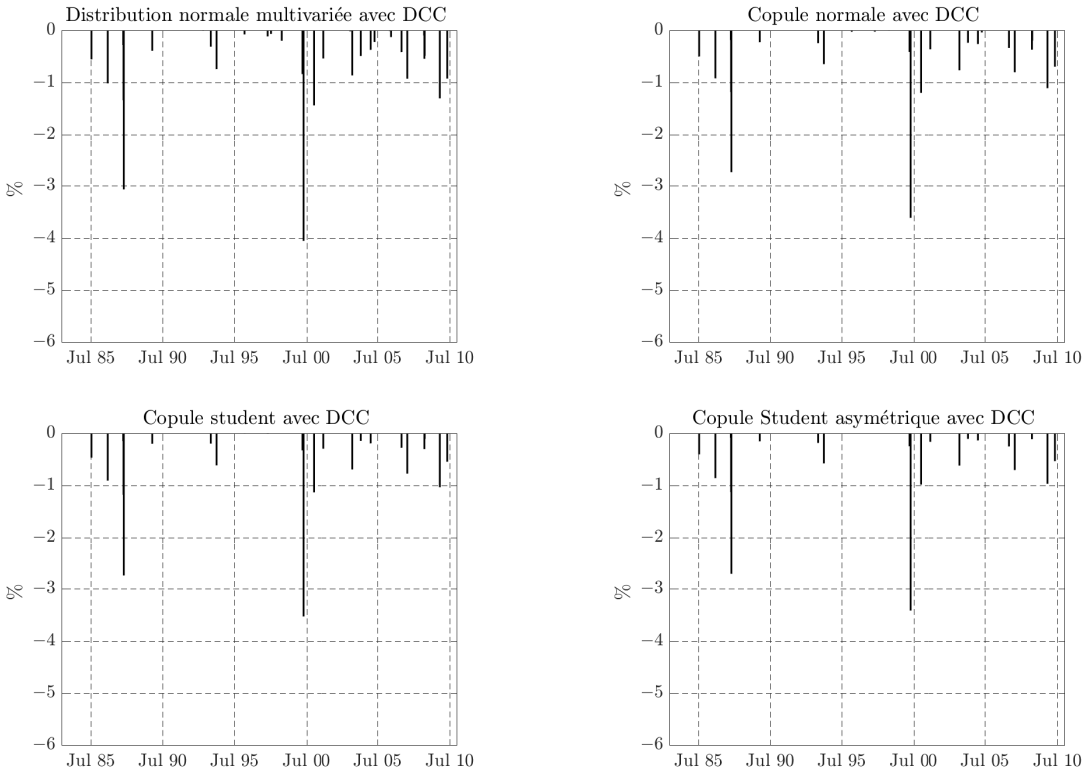
Notes : La figure présente des prévisions de CVaRs 1% hebdomadaires hors échantillon avec les deux modèles multivariés de référence. Les deux figures de gauche montrent les CVaRs d'une distribution normale multivariée avec corrélations constantes. Celles de gauches sont calculées avec distribution normale multivariée avec DCC. Les deux périodes considérées sont de juillet 1963 à décembre 2010 et de janvier 2006 à décembre 2010.

FIGURE 10 – Dépassements de VaRs 1% pour les modèles non dynamiques, 1983-2010



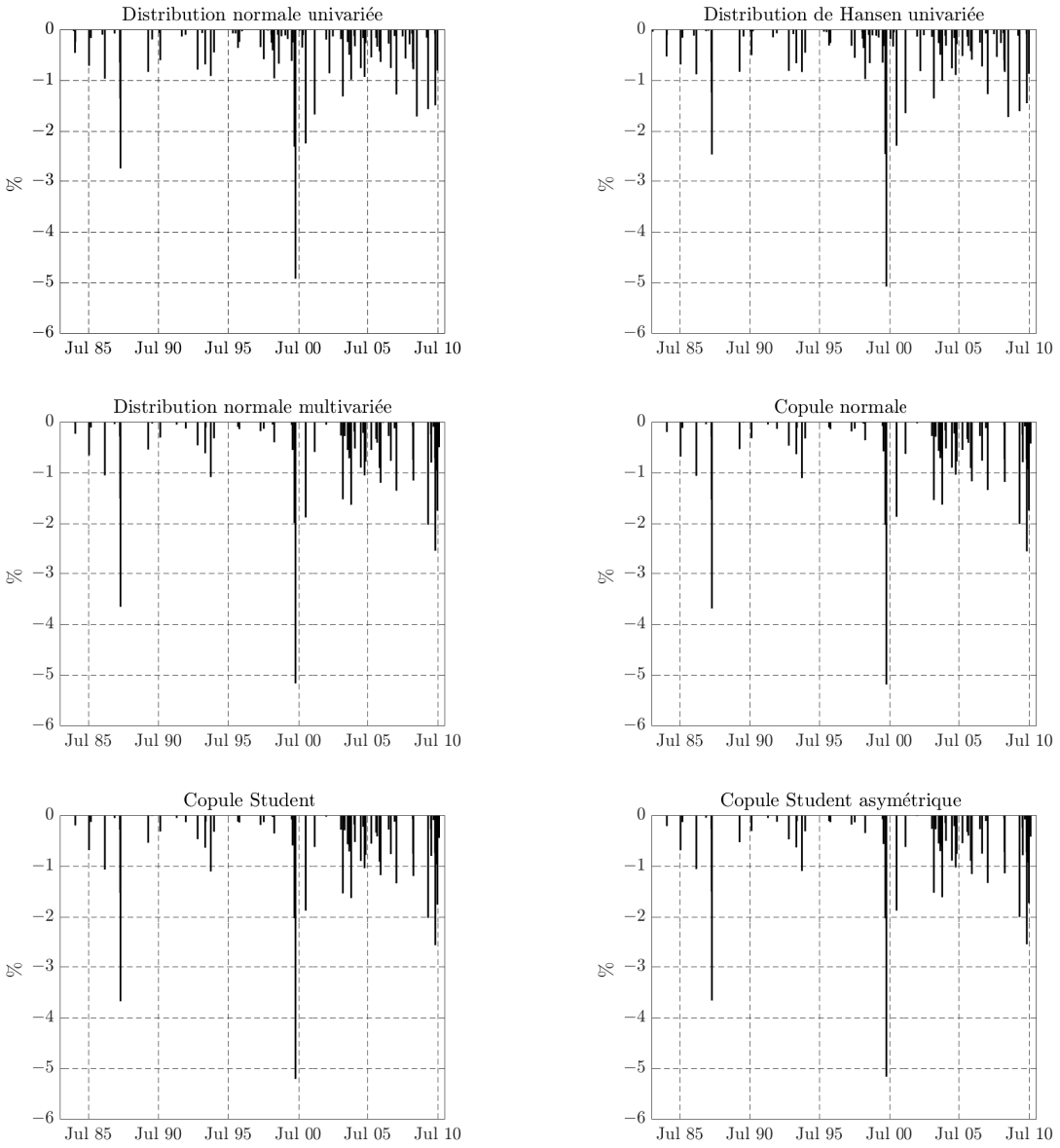
Notes : Les figures montrent la série chronologique de $\min\{r_{p,t+1} + VaR_{t+1}^{0,01}, 0\}$ pour les modèles univariés et les modèles multivariés sans corrélations dynamiques. La période hors échantillon va de la première semaine de juillet 1983 jusqu'à la fin de décembre 2010.

FIGURE 11 – Dépassements de VaRs 1% pour les modèles dynamiques, 1983-2010



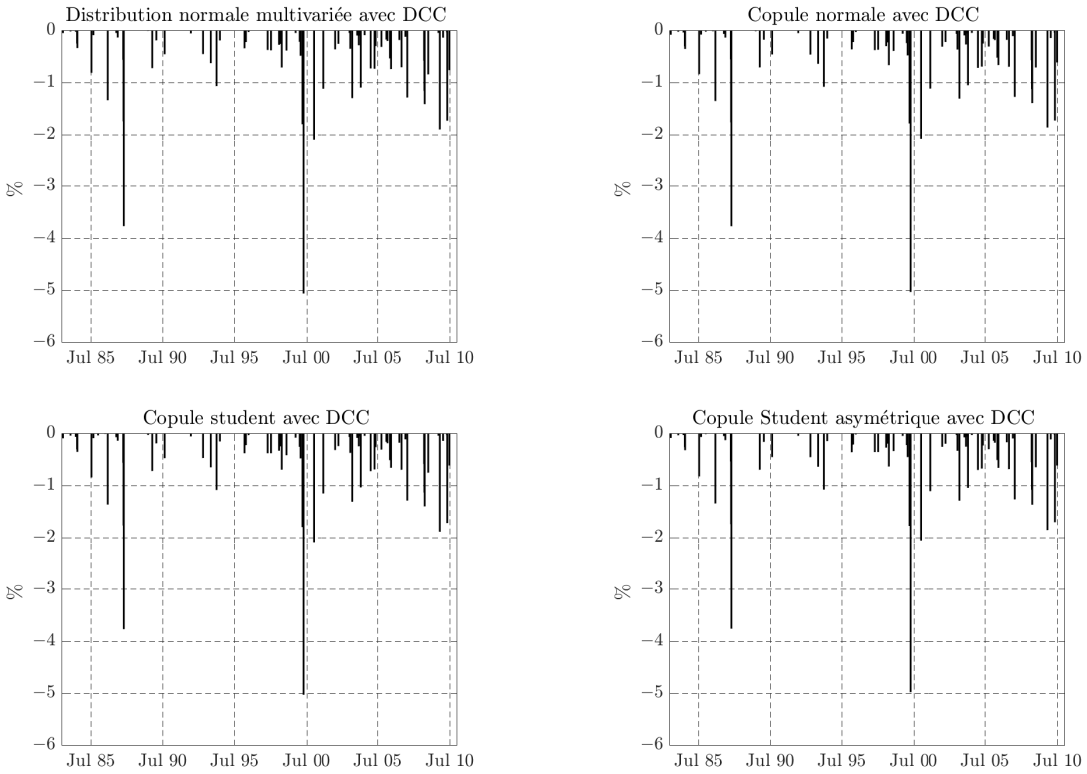
Notes : Les figures montrent la série chronologique de $\min\{r_{p,t+1} + VaR_{t+1}^{0,01}, 0\}$ pour les modèles multivariés avec DCC. La période hors échantillon va de la première semaine de juillet 1983 jusqu'à la fin de décembre 2010.

FIGURE 12 – Dépassements de VaRs 5% pour les modèles non dynamiques, 1983-2010



Notes : Les figures montrent la série chronologique de $\min\{r_{p,t+1} + VaR_{t+1}^{0,05}, 0\}$ pour les modèles univariés et les modèles multivariés sans corrélations dynamiques. La période hors échantillon va de la première semaine de juillet 1983 jusqu'à la fin de décembre 2010.

FIGURE 13 – Dépassements de VaRs 5% pour les modèles dynamiques, 1983-2010



Notes : Les figures montrent la série chronologique de $\min\{r_{p,t+1} + VaR_{t+1}^{0,05}, 0\}$ pour les modèles multivariés avec DCC. La période hors échantillon va de la première semaine de juillet 1983 jusqu'à la fin de décembre 2010.

TABLE 1 – Statistiques descriptives des rendements hebdomadaires des facteurs, 1963-2010

	Market	Size	Value	Momentum
Moyenne annuelle	3.63%	1.81%	4.30%	7.55%
Volatilité annuelle	15.95%	8.54%	8.73%	13.56%
Coefficient d'asymétrie	-0.75	-0.44	0.18	-1.44
Kurtosis	7.01	5.04	5.16	12.38
Autocorrélations				
Ordre 1	0.112	0.112	0.113	0.099
Ordre 2	0.103	0.103	0.084	0.084
Ordre 3	0.106	0.106	0.067	0.052
Corrélations croisées				
Market	-	0.048	-0.306	-0.110
Size	-	-	-0.128	0.039
Value	-	-	-	-0.190

Notes : La table présente les statistiques descriptives des rendements hebdomadaires des quatre facteurs pour la période allant du 5 juillet 1963 au 31 décembre 2010. La moyenne et la volatilité annuelle sont obtenues en multipliant la moyenne et la volatilité hebdomadaire par 52 et $\sqrt{52}$ respectivement. Le kurtosis présenté est normalisé.

TABLE 2 – Modèles AR-GARCH univariés des rendements de portefeuille, 1963-2010

Paramètres	Distribution normale	Distribution de Hansen
ϕ_0	7.29e-04 (1.11e-04)	7.37e-04 (1.09e-04)
ϕ_1	0.135 (0.021)	0.119 (0.021)
ϕ_2	0.069 (0.021)	0.064 (0.021)
ϕ_3	0.028 (0.020)	0.047 (0.020)
β	0.705 (0.015)	0.697 (0.028)
α	0.233 (0.012)	0.235 (0.021)
θ	0.067 (0.040)	0.045 (0.065)
ν	-	5.603 (0.546)
κ	-	-0.169 (0.546)
Log-likelihood	9213	9327

Notes : La table présente les coefficients du modèle AR(3)-NGARCH(1,1) estimés à l'aide des rendements d'un portefeuille équipondéré des quatre facteurs de 1963 à 2010. Les écart-types sont indiqués entre parenthèses. La première colonne utilise la distribution normale pour le terme d'erreur tandis que la deuxième colonne utilise la distribution t asymétrique de Hansen (1994).

TABLE 3 – Modèles AR-GARCH univariés des rendements des facteurs, 1963-2010

	Distribution Normale				Distribution de Hansen			
Paramètres	Market	Size	Value	Momentum	Market	Size	Value	Momentum
ϕ_0	6.34e-04	1.54e-04	5.43e-04	1.55e-03	7.30e-04	1.44e-04	4.80e-04	1.58e-03
	(3.34e-04)	(2.05e-04)	(1.83e-04)	(1.86e-04)	(2.91e-04)	(2.03e-04)	(1.83e-04)	(2.22e-04)
ϕ_1	0.030	0.086	0.137	0.107	0.011	0.101	0.147	0.101
	(0.020)	(0.021)	(0.019)	(0.019)	(0.021)	(0.021)	(0.020)	(0.019)
ϕ_2	0.053	0.108	0.036	0.007	0.046	0.124	0.042	0.013
	(0.021)	(0.021)	(0.022)	(0.020)	(0.021)	(0.021)	(0.021)	(0.019)
ϕ_3	0.022	0.073	0.069	-0.003	0.024	0.076	0.076	0.006
	(0.019)	(0.019)	(0.021)	(0.020)	(0.020)	(0.020)	(0.021)	(0.020)
β	0.752	0.849	0.862	0.832	0.793	0.849	0.877	0.831
	(0.018)	(0.017)	(0.011)	(0.011)	(0.020)	(0.022)	(0.015)	(0.016)
α	0.127	0.112	0.118	0.112	0.105	0.109	0.105	0.117
	(0.012)	(0.012)	(0.009)	(0.008)	(0.014)	(0.015)	(0.013)	(0.012)
θ	0.758	0.124	-0.058	-0.646	0.797	0.152	-0.074	-0.596
	(0.091)	(0.048)	(0.051)	(0.056)	(0.131)	(0.086)	(0.077)	(0.080)
ν	-	-	-	-	10.000	8.777	8.392	7.455
					(1.583)	(1.307)	(1.302)	(1.078)
κ	-	-	-	-	-0.221	-0.061	0.021	-0.161
					(0.027)	(0.029)	(0.029)	(0.029)
Diagnostic								
Log-vraisemblance	6290	7714	7860	7098	6357	7752	7891	7164
Persistance de variance	0.951	0.963	0.981	0.991	0.965	0.960	0.983	0.989
P-valeur L-B(20)	0.26	0.50	0.71	0.44	0.22	0.49	0.79	0.46
P-valeur abs L-B(20)	0.37	0.74	0.40	0.13	0.64	0.74	0.40	0.17
Asymétrie empirique	-0.56	-0.26	0.11	-0.56	-0.59	-0.26	0.11	-0.56
Asymétrie du modèle	0.00	0.00	0.00	0.00	-0.52	-0.16	0.06	-0.46
Kurtosis empirique	1.61	1.42	1.08	1.50	1.74	1.48	1.14	1.52
Kurtosis du modèle	0.00	0.00	0.00	0.00	1.27	1.28	1.37	1.97

TABLE 4 – Estimation des modèles de dépendance entre les facteurs, 1963-2010

	Corrélations constantes			Corrélations dynamiques		
Paramètres	Copule Normale	Copule Student	Copule Student Asymétrique	Copule Normale	Copule Student	Copule Student Asymétrique
ν_c		4.509 (0.228)	4.565 (0.172)		9.023 (0.736)	9.286 (0.002)
λ_{Market}			-0.008 (0.018)			-0.063 (0.047)
λ_{Size}			-0.059 (0.021)			-0.153 (0.050)
λ_{Value}			0.039 (0.018)			0.041 (0.044)
$\lambda_{Momentum}$			-0.150 (0.030)			-0.171 (0.048)
β_c				0.886 (0.003)	0.884 (0.005)	0.885 (0.005)
α_c				0.089 (0.004)	0.094 (0.006)	0.092 (0.006)
$\rho_{Market,Size}$	-0.014	0.000	0.001	0.013	0.039	0.034
$\rho_{Market,Value}$	-0.358	-0.352	-0.351	-0.378	-0.379	-0.374
$\rho_{Market,Momentum}$	0.107	0.115	0.119	0.082	0.089	0.088
$\rho_{Size,Value}$	-0.047	-0.047	-0.036	-0.086	-0.109	-0.104
$\rho_{Size,Momentum}$	0.013	0.022	-0.021	0.032	0.040	0.028
$\rho_{Value,Momentum}$	-0.077	-0.094	-0.072	-0.110	-0.139	-0.136
Log-likelihood	188.7	435.0	450.4	1051.0	1151.9	1161.2

TABLE 5 – *Backtest* des VaRs et CVaRs des modèles, 1983-2010

	$p = 1\%$					$p = 5\%$				
	LR_{uc}	LR_{ind}	LR_{cc}	Z_1	Z_2	LR_{uc}	LR_{ind}	LR_{cc}	Z_1	Z_2
Modèles univariés										
Distribution Normale	0.00	0.05	0.00	0.00	0.00	0.10	0.22	0.12	0.00	0.00
Distribution Hansen	0.01	0.04	0.00	0.93	0.02	0.10	0.21	0.13	0.38	0.05
Modèles multivariés										
Distribution Normale	0.00	0.01	0.00	0.00	0.00	0.82	0.04	0.08	0.00	0.01
Copule Normale	0.00	0.02	0.00	0.01	0.00	0.72	0.07	0.14	0.00	0.06
Copule Student	0.00	0.02	0.01	0.20	0.00	0.71	0.07	0.16	0.00	0.17
Copule t asymétrique	0.03	0.14	0.04	0.24	0.01	0.71	0.06	0.17	0.00	0.25
Distribution Normale-DCC	0.00	0.00	0.00	0.00	0.00	0.95	0.05	0.11	0.00	0.04
Copule Normale-DCC	0.01	0.00	0.00	0.07	0.00	0.74	0.06	0.14	0.00	0.11
Copule Student-DCC	0.04	0.00	0.00	0.18	0.01	0.67	0.08	0.15	0.02	0.14
Copule t asymétrique-DCC	0.15	0.01	0.01	0.22	0.05	0.96	0.05	0.10	0.09	0.31

Notes : Les cinq premières colonnes présentent le *backtest* des VaRs/CVaRs 1% tandis que les cinq dernières colonnes présentent le le *backtest* des VaRs/CVaRs 5%. Les colonnes LR_{uc}, LR_{ind} et LR_{cc} indiquent respectivement les p-valeurs du test de couverture non conditionnelle, d'indépendance et de couverture conditionnelle. Les colonnes Z_1 et Z_2 indiquent les p-valeurs des deux premiers tests d'Acerbi et Szekely (2014).

TABLE 6 – Comparaison de la moyenne des scores des modèles, 1983-2010

	Score moyen VaR		Score moyen VaR/CVaR	
	$p = 1\%$	$p = 5\%$	$p = 1\%$	$p = 5\%$
Modèles univariés				
Distribution Normale	3.13e-04	8.66e-04	6.57e-06	1.22e-05
Distribution Hansen	2.66e-04	8.66e-04	4.93e-06	1.17e-05
Modèles multivariés				
Distribution Normale	3.52e-04	9.38e-04	7.89e-06	1.36e-05
Copule Normale	3.39e-04	9.40e-04	7.34e-06	1.36e-05
Copule Student	3.27e-04	9.41e-04	6.83e-06	1.35e-05
Copule t asymétrique	3.22e-04	9.40e-04	6.68e-06	1.34e-05
Distribution Normale-DCC	3.06e-04	8.62e-04	6.91e-06	1.24e-05
Copule Normale-DCC	2.90e-04	8.63e-04	6.06e-06	1.22e-05
Copule Student-DCC	2.86e-04	8.62e-04	5.92e-06	1.21e-05
Copule t asymétrique-DCC	2.83e-04	8.62e-04	5.80e-06	1.21e-05

Notes : Le tableau présente la moyenne des scores hors échantillon de VaRs et VaRs/CVaRs pour chaque modèle à 1% et 5%. Pour les scores VaRs/CVaRs la constante W est fixée à 2.

TABLE 7 – Test CPA de Giacomini et White (2006) des VaRs 1% pour chaque paire de modèles, 1983-2010

	Distribution Hansen	Distribution Normale multivarié	Copule Normale	Copule Student	Copule t asymétrique	Distribution Normale multivarié-DCC	Copule Normale-DCC	Copule Student-DCC	Copule t asymétrique-DCC
Distribution Normale univariée	0.01↑ (1.00)	0.28 (0.01)	0.47 (0.00)	0.44 (0.01)	0.47 (0.01)	0.31 (0.91)	0.19 (0.99)	0.16 (1.00)	0.15 (1.00)
Distribution Hansen		0.01← (0.00)	0.02← (0.01)	0.04← (0.01)	0.05 (0.01)	0.28 (0.02)	0.52 (0.04)	0.59 (0.05)	0.64 (0.05)
Distribution Normale multivariée			0.04↑ (1.00)	0.04↑ (1.00)	0.03↑ (1.00)	0.02↑ (0.99)	0.00↑ (1.00)	0.00↑ (1.00)	0.00↑ (1.00)
Copule Normale				0.06 (1.00)	0.05↑ (1.00)	0.02↑ (0.99)	0.01↑ (0.99)	0.01↑ (1.00)	0.00↑ (1.00)
Copule Student					0.05↑ (1.00)	0.03↑ (0.98)	0.01↑ (0.99)	0.02↑ (0.99)	0.02↑ (1.00)
Copule t asymétrique						0.05↑ (0.97)	0.01↑ (0.99)	0.01↑ (0.99)	0.01↑ (1.00)
Distribution Normale multivariée-DCC							0.06 (1.00)	0.07 (1.00)	0.09 (1.00)
Copule Normale-DCC								0.21 (1.00)	0.26 (1.00)
Copule Student-DCC									0.34 (0.93)

Notes : Le tableau indique les p-valeurs du test CPA comparant le modèle ligne au modèle colonne à un seuil de 5%. La fonction de score est $S_{t+1} = (r_{t+1} + VaR_{t+1})\left(p - 1_{\{r_{t+1}+VaR_{t+1}<0\}}\right)$ et la fonction de test $h_t = (1, \Delta S_t)'$. Les nombres entre parenthèses indiquent la fraction des périodes hors échantillon où le modèle colonne performe mieux que le modèle ligne. Une flèche vers le haut(vers la gauche) indique que l'hypothèse nulle est rejetée et que le modèle colonne(ligne) performe mieux que l'autre plus de 50% du temps.

TABLE 8 – Test CPA de Giacomini et White (2006) des VaRs 5% pour chaque paire de modèles, 1983-2010

	Distribution Hansen	Distribution Normale multivarié	Copule Normale	Copule Student	Copule t asymétrique	Distribution Normale multivarié-DCC	Copule Normale-DCC	Copule Student-DCC	Copule t asymétrique-DCC
Distribution Normale univariée	0.81	0.02←	0.02←	0.02←	0.02←	0.39	0.42	0.39	0.43
	(0.43)	(0.00)	(0.00)	(0.00)	(0.00)	(0.66)	(0.60)	(0.62)	(0.61)
Distribution Hansen		0.02←	0.02←	0.02←	0.02←	0.46	0.49	0.46	0.51
		(0.00)	(0.01)	(0.01)	(0.01)	(0.62)	(0.57)	(0.60)	(0.57)
Distribution Normale multivariée			0.00←	0.00←	0.00←	0.00↑	0.00↑	0.00↑	0.00↑
			(0.02)	(0.01)	(0.04)	(0.98)	(0.98)	(0.98)	(0.98)
Copule Normale				0.06	0.42	0.00↑	0.00↑	0.00↑	0.00↑
				(0.08)	(0.52)	(0.98)	(0.98)	(0.98)	(0.98)
Copule Student					0.24	0.00↑	0.00↑	0.00↑	0.00↑
					(0.81)	(0.98)	(0.98)	(0.98)	(0.98)
Copule t asymétrique						0.00↑	0.00↑	0.00↑	0.00↑
						(0.98)	(0.98)	(0.98)	(0.98)
Distribution Normale multivariée-DCC							0.64	0.50	0.72
							(0.03)	(0.28)	(0.04)
Copule Normale-DCC								0.59	0.43
								(0.97)	(0.25)
Copule Student-DCC									0.59
									(0.05)

Notes : Le tableau indique les p-valeurs du test CPA comparant le modèle ligne au modèle colonne à un seuil de 5%. La fonction de score est $S_{t+1} = (r_{t+1} + VaR_{t+1})\left(p - 1_{\{r_{t+1}+VaR_{t+1}<0\}}\right)$ et la fonction de test $h_t = (1, \Delta S_t)'$. Les nombres entre parenthèses indiquent la fraction des périodes hors échantillon où le modèle colonne performe mieux que le modèle ligne. Une flèche vers le haut (vers la gauche) indique que l'hypothèse nulle est rejetée et que le modèle colonne (ligne) performe mieux que l'autre plus de 50% du temps.

TABLE 9 – Test CPA de Giacomini et White (2006) des VaRs/CVaRs 1% pour chaque paire de modèles, 1983-2010

	Distribution Hansen	Distribution Normale multivarié	Copule Normale	Copule Student	Copule t asymétrique	Distribution Normale multivarié-DCC	Copule Normale-DCC	Copule Student-DCC	Copule t asymétrique-DCC
Distribution Normale univariée	0.09 (1.00)	0.08 (0.00)	0.21 (0.00)	0.83 (0.01)	0.94 (0.02)	0.59 (0.03)	0.11 (0.99)	0.10 (1.00)	0.10 (1.00)
Distribution Hansen		0.04← (0.00)	0.04← (0.00)	0.04← (0.00)	0.04← (0.00)	0.22 (0.00)	0.40 (0.01)	0.50 (0.02)	0.58 (0.02)
Distribution Normale multivariée			0.12 (1.00)	0.14 (1.00)	0.15 (1.00)	0.05 (0.99)	0.01↑ (0.99)	0.01↑ (0.99)	0.01↑ (0.99)
Copule Normale				0.17 (1.00)	0.19 (1.00)	0.08 (0.98)	0.03↑ (1.00)	0.02↑ (1.00)	0.01↑ (0.99)
Copule Student					0.25 (1.00)	0.37 (0.58)	0.03↑ (0.99)	0.03↑ (0.99)	0.03↑ (1.00)
Copule t asymétrique						0.46 (0.23)	0.04↑ (0.98)	0.02↑ (0.99)	0.02↑ (1.00)
Distribution Normale multivariée-DCC							0.15 (1.00)	0.15 (1.00)	0.16 (1.00)
Copule Normale-DCC								0.29 (1.00)	0.26 (1.00)
Copule Student-DCC									0.27 (0.99)

Notes : Le tableau indique les p-valeurs du test CPA comparant le modèle ligne au modèle colonne à un seuil de 5%. La fonction de score est $S_{t+1} = \frac{p}{2}CVaR_{t+1}^2 + \frac{Wp}{2}VaR_{t+1}^2 - pCVaR_{t+1}VaR_{t+1} + \left[CVaR_{t+1}(VaR_{t+1} + r_{t+1}) + \frac{W}{2}(r_{t+1}^2 - VaR_{t+1}^2)\right] 1_{\{r_{t+1}+VaR_{t+1}<0\}}$ et la fonction de test $h_t = (1, \Delta S_t)'$. Les nombres entre parenthèses indiquent la fraction des périodes hors échantillon où le modèle colonne performe mieux que le modèle ligne. Une flèche vers le haut (vers la gauche) indique que l'hypothèse nulle est rejetée et que le modèle colonne (ligne) performe mieux que l'autre plus de 50% du temps.

TABLE 10 – Test CPA de Giacomini et White (2006) des VaRs/CVaRs 5% pour chaque paire de modèles, 1983-2010

	Distribution Hansen	Distribution Normale multivarié	Copule Normale	Copule Student	Copule t asymétrique	Distribution Normale multivarié-DCC	Copule Normale-DCC	Copule Student-DCC	Copule t asymétrique-DCC
Distribution Normale univariée	0.06	0.02←	0.03←	0.06	0.10	0.82	0.43	0.36	0.35
	(1.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.09)	(0.57)	(0.67)	(0.77)
Distribution Hansen		0.01←	0.01←	0.01←	0.01←	0.55	0.78	0.81	0.83
		(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.05)	(0.05)	(0.06)
Distribution Normale multivariée			0.03↑	0.25	0.26	0.02↑	0.01↑	0.01↑	0.01↑
			(0.84)	(0.97)	(0.98)	(0.99)	(1.00)	(1.00)	(1.00)
Copule Normale				0.29	0.30	0.02↑	0.01↑	0.01↑	0.01↑
				(1.00)	(1.00)	(0.98)	(0.99)	(0.99)	(1.00)
Copule Student					0.31	0.01↑	0.01↑	0.01↑	0.01↑
					(0.98)	(0.97)	(0.99)	(0.99)	(0.99)
Copule t asymétrique						0.01↑	0.01↑	0.01↑	0.01↑
						(0.97)	(0.98)	(0.98)	(0.99)
Distribution Normale multivariée-DCC							0.36	0.35	0.36
							(1.00)	(1.00)	(1.00)
Copule Normale-DCC								0.48	0.42
								(1.00)	(1.00)
Copule Student-DCC									0.45
									(0.94)

Notes : Le tableau indique les p-valeurs du test CPA comparant le modèle ligne au modèle colonne à un seuil de 5%. La fonction de score est $S_{t+1} = \frac{p}{2}CVaR_{t+1}^2 + \frac{Wp}{2}VaR_{t+1}^2 - pCVaR_{t+1}VaR_{t+1} + \left[CVaR_{t+1}(VaR_{t+1} + r_{t+1}) + \frac{W}{2}(r_{t+1}^2 - VaR_{t+1}^2)\right] 1_{\{r_{t+1}+VaR_{t+1}<0\}}$ et la fonction de test $h_t = (1, \Delta S_t)'$. Les nombres entre parenthèses indiquent la fraction des périodes hors échantillon où le modèle colonne performe mieux que le modèle ligne. Une flèche vers le haut(vers la gauche) indique que l'hypothèse nulle est rejetée et que le modèle colonne(ligne) performe mieux que l'autre plus de 50% du temps.