Mémoire longue, volatilité et gestion de portefeuille

Jérôme Coulon

▶ **To cite this version:**

Jérôme Coulon. Mémoire longue, volatilité et gestion de portefeuille. Economies et finances. Université Claude Bernard - Lyon I, 2009. Français. NNT : 2009LYO10068 . tel-00657711

HAL Id: tel-00657711
https://theses.hal.science/tel-00657711

Submitted on 9 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Claude Bernard Lyon 1

I.S.F.A. **École Doctorale de Sc. Économiques et de Gestion**

Mémoire longue, volatilité et gestion de portefeuille

THÈSE

Numéro d'ordre 068-2009

présentée et soutenue publiquement le 20 mai 2009

pour l'obtention du

Doctorat de l'université Claude Bernard – Lyon 1

par

Jérôme COULON

Composition du jury

Pr. Yannick MALEVERGNE	Université de Saint-Etienne	(co-Directeur de thèse)
Pr. Vêlayoudom MARIMOUTOU	Institut français de Pondicherry	(Rapporteur)
Pr. Patrick NAVATTE	Université de Rennes 1	
Pr. François QUITTARD-PINON	Université Lyon 1	(Directeur de thèse)
Pr. Jean-Charles ROCHET	Université Toulouse I	(Rapporteur)
Pr. Michael ROCKINGER	HEC Lausanne	

Laboratoire de Sciences Actuarielle et Financière — EA 2429

Remerciements

Sur les marchés financiers comme dans la vie, le hasard joue un rôle déterminant. Je dois me rendre à l'évidence que mon projet de thèse ne fait pas exception à la règle. En effet, c'est de la combinaison de deux évènements assez improbables qu'est née l'envie de poursuivre mes études en me lançant dans une thèse de doctorat. Premièrement, il y eut mon désir, après l'obtention de mon DESS de Statistiques, Informatique et Techniques Numériques, de m'orienter vers la finance en m'inscrivant au DEA de Sciences Actuarielle et Financière à l'ISFA (Université Claude Bernard - Lyon 1). Deuxièmement, ce furent les rencontres déterminantes des Professeurs François Quittard-Pinon et Yannick Malevergne qui devinrent par la suite respectivement mon directeur et mon co-directeur de thèse. Je leur suis reconnaissant de tout ce qu'ils m'ont apporté durant cette tranche de vie et les remercie tout particulièrement de leur disponibilité, de la qualité de leurs remarques, de leur gentillesse et surtout de leur patience !

Je tiens aussi à remercier monsieur Vêlayoudom Marimoutou, Professeur et Directeur de l'Institut Français de Pondicherry, et monsieur Jean-Charles Rochet, Professeur à l'université Toulouse I, pour avoir accepté d'être rapporteur ainsi que monsieur Patrick Navatte, Professeur à l'université de Rennes 1, et monsieur Michael Rockinger, Professeur à HEC Lausanne, pour leur participation au jury de thèse.

Mes remerciements vont également à Jean-Paul Laurent, pour m'avoir accueilli au sein du Laboratoire de Sciences Actuarielle et Financière de l'ISFA, donné accès à ses ressources et permis de participer à diverses conférences durant ma thèse. J'associe aussi toute l'équipe administrative qui a su, par sa gentillesse et sa compétence, me simplifier la vie à maintes reprises.

Le hasard a encore frappé lorsqu'en 2006 j'ai eu l'occasion de rejoindre pour un an la chaire "Entrepreneurial risk" du Professeur Didier Sornette à l'ETH Zürich, auquel j'exprime aujourd'hui ma profonde reconnaissance. Cette étape a été non seulement l'occasion de poursuivre mes travaux de thèse dans un environnement international et multi-culturel, mais également celle de découvrir un autre pays, une autre ville et une autre culture.

Ce séjour n'aurait pas été un aussi bon souvenir sans les stagiaires, doctorants et docteurs que j'ai eu l'occasion de rencontrer et avec lesquels j'ai pu discuter, participer à des présentations, et... jouer au poker. Eh oui, parfois une bonne soirée de détente est salutaire. Un grand merci donc à Felix, Georges, Gilles, Max, Moritz, Riley, Shengsui, Thomas, sans oublier Alison, Barbara, Kerstin et Uli. En particulier, à ceux qui ont bien voulu perdre leurs jetons !

J'ai également une pensée pour tous mes collègues thésards du laboratoire SAF, Areski, Matthieu, Florent, Romain, des deux premières années lyonnaises. Courage à tous ceux qui n'en ont pas encore fini avec cette période où se mélangent, entre autres, excitation, doute, découragement, espoir et soulagement (souvent dans cet ordre d'ailleurs).

Et que dire des amis, de plus ou moins longue date, rencontrés au hasard des bancs d'un amphi, d'un court de tennis ou encore d'un terrain de foot et qui ont su, à leur manière, m'encourager tout au long de cette quête du savoir ? Je remercie Flo, Greg, Nico et tous les

autres chaleureusement pour leur soutien, les moments de détentes partagés mais également pour leur humour si réconfortant du genre “tu comptes bien la finir un jour cette thèse ?”...

C’est aussi pour moi l’occasion d’exprimer ma profonde reconnaissance et mes plus sincères remerciements à ma famille, Benjamin et surtout mes parents pour m’avoir toujours soutenu et mis dans les meilleures conditions pour mener à bien mes études.

Je souhaiterais également exprimer ma gratitude à celles et ceux qui ont eu la gentillesse de consacrer du temps et de l’énergie à la relecture de ma thèse, bien que ce ne fût pas leur spécialité.

Et enfin, un grand merci à celle qui a partagé et embelli mon quotidien durant toutes ces années, et m’a constamment encouragé par sa présence, ses mots et ses petits plats... Pour toutes ces attentions passées et celles à venir, merci Fanny.

Résumé

Cette thèse porte sur l'étude de la mémoire longue de la volatilité des rendements d'actions. Dans une première partie, nous apportons une interprétation de la mémoire longue en termes de comportement d'agents grâce à un modèle de volatilité à mémoire longue dont les paramètres sont reliés aux comportements hétérogènes des agents pouvant être rationnels ou à rationalité limitée. Nous déterminons de manière théorique les conditions nécessaires à l'obtention de mémoire longue. Puis nous calibrons notre modèle à partir des séries de volatilité réalisée journalière d'actions américaines de moyennes et grandes capitalisations et observons le changement de comportement des agents entre la période précédant l'éclatement de la bulle internet et celle qui la suit. La deuxième partie est consacrée à la prise en compte de la mémoire longue en gestion de portefeuille. Nous commençons par proposer un modèle de choix de portefeuille à volatilité stochastique dans lequel la dynamique de la log-volatilité est caractérisée par un processus d'Ornstein-Uhlenbeck. Nous montrons que l'augmentation du niveau d'incertitude sur la volatilité future induit une révision du plan de consommation et d'investissement. Puis dans un deuxième modèle, nous introduisons la mémoire longue grâce au mouvement brownien fractionnaire. Cela a pour conséquence de transposer le système économique d'un cadre markovien à un cadre non-markovien. Nous fournissons donc une nouvelle méthode de résolution fondée sur la technique de Monte Carlo. Puis, nous montrons toute l'importance de modéliser correctement la volatilité et mettons en garde le gérant de portefeuille contre les erreurs de spécification de modèle.

Mots-clés: Mémoire longue, volatilité réalisée, modèle de volatilité, agents hétérogènes, rationalité limitée, mouvement brownien fractionnaire, modèle de choix de portefeuille.

Abstract

This PhD thesis is about the study of the long memory of the volatility of asset returns. In a first part, we bring an interpretation of long memory in terms of agents' behavior through a long memory volatility model whose parameters are linked with the bounded rational agents' heterogeneous behavior. We determine theoretically the necessary condition to get long memory. Then we calibrate our model from the daily realized volatility series of middle and large American capitalization stocks. Eventually, we observe the change in the agents' behavior between the period before the internet bubble burst and the one after. The second part is devoted to the consideration of long memory in portfolio management. We start by suggesting a stochastic volatility portfolio model in which the dynamics of the log-volatility is characterized by an Ornstein-Uhlenbeck process. We show that when the uncertainty of the future volatility level increases, it induces the revision of the consumption and investment plan. Then in a second model, we introduce a long memory component by the use of a fractional Brownian motion. As a consequence, it transposes the economic system from a Markovian framework to a non-Markovian one. So we provide a new resolution method based on Monte Carlo technique. Then we show the high importance to well model the volatility and warn the portfolio manager against the misspecification errors of the model.

Keywords: Long memory, realized volatility, volatility modelling, heterogeneous agents, bounded rationality, fractional brownian motion, portfolio modelling.

Table des matières

Liste des tableaux **xi**

Liste des figures **xiii**

Introduction générale

Partie I Modélisation de la volatilité

Introduction **13**

Chapitre 1 Revue de la littérature **17**

1.1 Les rendements d'actifs 19

1.1.1 Facteurs influençant les rendements 19

1.1.2 Faits stylisés des rendements d'actifs 21

1.2 La volatilité des rendements du prix d'une action 27

1.2.1 Facteurs influençant la volatilité 28

1.2.2 Estimateurs de la volatilité 29

1.2.3 Estimateur de la volatilité réalisée 30

1.2.4 Faits stylisés de la volatilité réalisée 33

1.3 La mémoire longue 40
1.3.1 Définitions 40
1.3.2 Modèles appliqués à la représentation de la volatilité 42
1.3.3 Origine de la mémoire longue 47
1.3.4 Signification 49
1.4 La prise en compte de l'hétérogénéité des agents 50
Conclusion 52

Chapitre 2 Le Modèle d'hétérogénéité **53**
2.1 Cadre général 53
2.2 Propriétés de la (log-)volatilité 56
2.3 Analyse du paramètre d'hétérogénéité du modèle 58
2.3.1 Absence d'hétérogénéité 58
2.3.2 Absence de référence à la volatilité passée 59
2.3.3 Indépendance entre les paramètres ψ et φ 60
2.3.4 Dépendance entre les paramètres ψ et φ 60
2.4 Exemples de comportements 63
2.4.1 Agents rationnels 65
2.4.2 Ecart à la rationnalité 65
2.4.3 Extension de l'écart à la rationalité 73
Conclusion 79

Annexes
2.A Démonstration de la proposition 1 84
2.B Démonstration de la proposition 2 86
2.C Calcul de la fonction d'autocorrélation par FFT 87

Chapitre 3 Résultats empiriques **89**
3.1 Inférence statistique 89
3.1.1 Choix de l'estimateur 91
3.1.2 Commentaires sur la mise en œuvre 93
3.2 Résultats empiriques 97
3.2.1 Validation des faits stylisés des rendements et de la volatilité . . . 98
3.2.2 Calibrage du modèle sur l'intervalle de temps tout entier 106
3.3 Étude de l'effet d'éclatement de la bulle 112
3.3.1 Étude d'une action de nouvelle technologie : Cisco Systems 113
3.3.2 Étude d'une action non technologique : Coca Cola 119
3.3.3 Observations générales 120

Conclusion 124

Annexes

3.A Coefficients du moment d'ordre k 127
3.B Méthodes semi-paramétriques d'estimation de d 127
3.B.1 Méthode de Geweke et Porter-Hudak 128
3.B.2 Méthode de Hurst 128

Partie II Gestion de portefeuille

Introduction **133**

Chapitre 4 Revue de littérature **137**
4.1 Les critères prix en compte dans la modélisation 138
4.1.1 L'approche 138
4.1.2 Le cadre 139
4.1.3 La prise en compte des imperfections de marché 139
4.1.4 La prise en compte des facteurs de risque 140
4.1.5 L'objectif de l'investisseur 141
4.1.6 La méthode de résolution 141
4.2 Un cas particulier du modèle de Merton à volatilité constante 142
4.3 Modèles en temps continu avec volatilité stochastique 143
4.3.1 Le modèle de Chacko et Viceira 144
4.3.2 Le modèle de Liu 145
4.3.3 Le modèle de Benth, Karlsen et Reikvam 146
Conclusion 147

Chapitre 5 Modèle avec volatilité stochastique **149**
5.1 Modèle à volatilité constante 150
5.1.1 Définition du modèle 150
5.1.2 Étude des critères de gestion avec consommation de l'investisseur 152
5.1.3 Étude du critère de gestion sans consommation de l'investisseur . 153

5.2 Modèle à volatilité stochastique 153
5.2.1 Définition du modèle 154
5.2.2 Étude des critères de gestion avec consommation de l'investisseur 155
5.2.3 Étude du critère de gestion sans consommation de l'investisseur . 156
5.3 Impact de la volatilité stochastique 157
5.3.1 Définition du problème étudié 157
5.3.2 Méthodologie 158
5.3.3 Choix des paramètres 161
5.3.4 Résultats 162
Conclusion 174

Annexes

5.A Rappels sur le mouvement brownien 176
5.B Démonstration de la section 5.2.2 177
5.C Démonstration de la section 5.2.3 180

Chapitre 6 Modèle avec volatilité stochastique à mémoire longue 183
6.1 Présentation du modèle avec volatilité stochastique à mémoire longue . . 184
6.1.1 Définition du modèle 185
6.1.2 Étude du critère de gestion sans consommation de l'investisseur . 187
6.2 Étude du processus de la log-volatilité à mémoire longue 189
6.2.1 Impact du paramètre de mémoire longue 189
6.2.2 Autocovariance de la log-volatilité 189
6.3 Impact de la mémoire longue 191
6.3.1 Définition du problème 191
6.3.2 Méthodologie 191
6.3.3 Paramétrage 194
6.3.4 Résultats 195
Conclusion 212

Annexes

6.A Mouvement brownien fractionnaire 215
6.B Démonstration de l'équation (6.6) 217
6.C Démonstration de la section 6.1.2 220
6.D Démonstration de l'équation (6.20) 222
6.E Démonstration de la proposition 5 224
6.E.1 Nouvelle formulation du poids optimal de l'actif risqué 224
6.E.2 Simulation des mouvements browniens 227

6.E.3 Formulation définitive 228
6.F Démonstration de la proposition 6 229
6.G Démonstration de la proposition 7 230

Conclusion générale

Annexes **241**

Annexe A Évaluation de la conduite du projet de thèse **241**
A.1 Bref résumé du sujet 241
A.2 Éléments de contexte 242
A.2.1 Choix du sujet 242
A.2.2 Choix de l'encadrement et du laboratoire d'accueil 242
A.2.3 Financement du projet 243
A.3 Évolution du projet 244
A.3.1 Élaboration du projet 244
A.3.2 Conduite du projet 245
A.3.3 Retombées scientifiques 245
A.4 Compétences acquises et enseignements personnels 246
A.5 Conclusion 247

Bibliographie **249**

Liste des tableaux

3.1 Descriptif des actions constituant notre base de données 98
3.2 Statistiques de la série des rendements journaliers 101
3.3 Statistiques de la log-volatilité réalisée obtenue par l'estimateur $\hat{\sigma}_{ABDL}$ . . . 105
3.4 Statistiques des rendements journaliers normalisés par la VR journalière . . . 107
3.5 Estimations du paramètre de mémoire longue, d, par notre modèle 112
3.6 Estimations du paramètre de mémoire longue, d, par plusieurs méthodes . . . 113
3.7 Estimation de d, α et w pour différentes phases du marché 125

5.1 Paramètres du modèle avec volatilité stochastique 162

6.1 Paramètres du modèle à mémoire longue 194

A.1 Récapitulatif des différents frais engagés pour la thèse 244

Liste des figures

1.1 Densité des rendements de CSCO à 3 différentes fréquences 23
1.2 Fonction d'autocorrélation des rendements journaliers de CSCO 26
1.3 Fonction d'autocorrélation des rendements 1 minute de CSCO 27
1.4 Volatilité réalisée journalière de WPO et KO 34
1.5 Fonction d'autocorrélation de la volatilité réalisée de WPO et KO 35
1.6 Densité de $\log(\hat{\sigma}_t)$ et de sa loi normale approchée pour MCHP et PG 37
1.7 Densité des rendements journaliers de VLO et GE 38
1.8 Densité des rendements journaliers normalisés de VLO et GE 39
1.9 Fonctions d'autocorrélation de la log-volatilité réalisée de 3 actions 44
1.10 Fonctions d'autocorrélation de processus ARMA(1,1) et ARFIMA(1,d,1) . . . 45
1.11 Fonctions d'autocorrélation de plusieurs processus ARFIMA(1,d,1) 46

2.1 ACF pour différents paramètres 62
2.2 Densité du coefficient d'hétérogénéité pour différents paramètres 63
2.3 Impact de β sur la fonction d'autocorrélation 64
2.4 Impact de $\bar{\varphi}'$ sur la fonction d'autocorrélation 68
2.5 Impact de p sur la fonction d'autocorrélation 70
2.6 Impact de p sur la densité du paramètre d'hétérogénéité 71
2.7 Impact de q sur la fonction d'autocorrélation 72
2.8 Impact de q sur la densité du paramètre d'hétérogénéité 73
2.9 Impact de α sur la fonction d'autocorrélation 74
2.10 Impact de w sur la fonction d'autocorrélation 76
2.11 Impact de w sur la densité du paramètre d'hétérogénéité 77
2.12 Impact de σ sur la fonction d'autocorrélation 78
2.13 Impact de σ sur la densité du paramètre d'hétérogénéité 79
2.14 Impact de m sur la fonction d'autocorrélation 80
2.15 Impact de m sur la densité du paramètre d'hétérogénéité 81

3.1 Efficacité des algorithmes de minimisation 96
3.2 Rendements journaliers de l'action TCB 100
3.3 Densité des rendements journaliers de l'action TCB 102
3.4 Densité de la log-volatilité réalisée de l'action KO 104
3.5 Densité de la log-volatilité réalisée de l'action FNF 106
3.6 Densité des rendements journaliers normalisés de l'action TCB 108
3.7 ACF à horizon 120 jours et densité de la log-VR avec IC 95% de FNF 109
3.8 ACF à horizon 120 jours et densité de la log-VR avec IC 95% de MSFT . . . 111
3.9 Cours de l'action Cisco Systems 114

3.10 ACF empiriques de la log-volatilité de CSCO sur plusieurs périodes 116
3.11 ACF et densité du coefficient d'hétérogénéité de CSCO, période pré-bulle . . 117
3.12 ACF et densité du coefficient d'hétérogénéité de CSCO, période post-bulle . . 118
3.13 Densité du paramètre d'hétérogénéité de CSCO des sous-périodes 119
3.14 Cours de l'action Coca-Cola 120
3.15 ACF empiriques de l'action KO pour plusieurs phases 121
3.16 ACF et densité du coefficient d'hétérogénéité de KO, période pré-bulle 122
3.17 ACF et densité du coefficient d'hétérogénéité de KO, période post-bulle . . . 123
3.18 Densités des paramètres d'hétérogénéité de Coca-Cola des sous-périodes . . . 124

5.1 Consommation optimale en fonction de t et de la volatilité σ_s 163
5.2 Poids de l'actif risqué en fonction de la volatilité σ_s 164
5.3 Variation de $c^*(t)$ en fonction de la variation de la volatilité, $\kappa = 2$ 166
5.4 Variation de $w^*(t)$ en fonction de la variation de la volatilité, $\kappa = 2$ 167
5.5 Variation de $c^*(t)$ en fonction de la variation de la volatilité, $\sigma = 1$ 169
5.6 Variation de $w^*(t)$ en fonction de la variation de la volatilité, $\sigma = 1$ 170
5.7 Variation de $w^*(t)$ en fonction de la variation de la volatilité, $\sigma = 1$ et $\rho = 0.25$ 172
5.8 Couverture intertemporelle pour différents κ, $\sigma = 1$ et $\rho = 0.25$ 173
5.9 Taux de croissance en fonction du temps pour différents μ 174

6.1 Trajectoires de mouvements browniens fractionnaires 186
6.2 Trajectoires de volatilité pour différents d 190
6.3 Convergence des modèles 196
6.4 Volatilité et poids optimal de l'action, gestion sur 1 an 197
6.5 Cours de l'action et richesse du portefeuille, gestion sur 1 an 199
6.6 Richesse corrigée du portefeuille, gestion sur 1 an 200
6.7 Richesse et richesse corrigée du portefeuille, gestion sur 5 ans 201
6.8 Richesse et richesse corrigée du portefeuille, gestion sur 10 ans 203
6.9 Poids de l'action, pour différentes aversions au risque, gestion sur 5 ans . . . 204
6.10 Richesse et richesse corrigée du portefeuille, $\gamma = 0.5$, gestion sur 5 ans 205
6.11 Richesse et richesse corrigée du portefeuille, $\gamma = -19$, gestion sur 5 ans 206
6.12 Volatilité de l'action pour différents κ, gestion sur 5 ans 208
6.13 Richesse et richesse corrigée du portefeuille, $\kappa = 12$, gestion sur 5 ans 209
6.14 Richesse et richesse corrigée du portefeuille, $\kappa = 50$, gestion sur 5 ans 210
6.15 Volatilité et poids optimal de l'action, forte volatilité, gestion sur 1 an 211
6.16 Richesse et richesse corrigée du portefeuille, forte volatilité, gestion sur 1 an . 213

Introduction générale

Les experts financiers ainsi que les dirigeants politiques s'accordent à dire depuis quelques mois que la crise financière actuelle est la plus grave depuis la Grande Dépression de 1929. En effet, son ampleur est telle qu'elle s'est propagée à l'économie réelle, provoquant par effet domino un ralentissement économique, puis une récession dans les pays du monde entier. Une des causes de cette crise majeure réside dans la mauvaise appréciation du risque des titres adossés aux crédits hypothécaires américains, notamment par les agences de notation, et de leur prolifération (Krugman 2007, Dodd et Mills 2008). Deux décennies auparavant, les options et le trading automatique ont été accusés d'être parmi les principales causes du krach boursier de 1987, notamment du fait de l'inadéquation du modèle de Black & Scholes pour répliquer les options du fait de la présence d'un risque résiduel négligé par le modèle (Brady 1988). Cela illustre de manière générale l'importance cruciale de correctement estimer le risque en finance.

Les deux exemples évoqués ci-dessus sont qualifiés de risques extrêmes étant donné l'ampleur de leur impact. Toutefois, les risques financiers ne se résument pas à des évènements de cette importance. En effet, l'activité financière est soumise à de nombreux risques plus ou moins grands et dont l'apparition est aléatoire. En résumé, le hasard est omniprésent sur les marchés financiers et joue un rôle déterminant. A ce propos, Mandelbrot reconnaît deux types de hasard : le hasard *benin* donnant lieu à des évènements s'écartant peu de la moyenne, ou au contraire le hasard *sauvage*, source d'évènements de très grande ampleur voire extrêmes (Malevergne et Sornette 2005). Par conséquent, gérer le risque en finance revient à éviter de se faire surprendre par le hasard, pour reprendre le thème du livre de Thaleb (2001). Pour y parvenir, il convient de se doter d'outils permettant de mieux appréhender les risques sur les marchés financiers.

Sur le marché des actions, une mesure élémentaire du risque est la volatilité des rendements. Or, contrairement à d'autres variables financières telles que le prix des actions, la volatilité n'est pas une variable directement observable. Il faut donc l'évaluer. Dans cette optique, de nombreux estimateurs et modèles dédiés à la représentation de la volatilité ont été développés. L'apport des séries temporelles a permis d'effectuer de réels progrès dans ce domaine, notamment autour du *concept ARCH* introduit par Engle (1982) qui a connu un très grand succès du fait de sa simplicité de mise en œuvre et de sa capacité à rendre compte de la majorité des caractéristiques des séries financières (variance conditionnelle non constante principalement).

Dans l'espoir d'obtenir une meilleure évaluation du niveau de risque réel des marchés financiers, la modélisation doit prendre en compte les caractéristiques les plus importantes de la volatilité (Campbell, Lo, et MacKinlay 1997) et une tout particulièrement que la plupart des modèles développés jusqu'à présent négligent : la mémoire longue. Ce concept, introduit par Hurst (1951), alors qu'il cherchait à modéliser les crues du Nil, caractérise une persistance à long terme des autocorrélations d'une série temporelle. On parle alors de décroissance hyperbolique. En finance, les séries de volatilité journalière présentent de la mémoire longue. Cela signifie que le niveau de volatilité journalière datant de six mois a un impact sur la volatilité actuelle, et que celle-ci aura encore un impact sur la valeur de volatilité dans six, neuf mois voire un an. Cela se révèle particulièrement intéressant dans l'optique de la gestion à long terme.

Afin de prendre en compte la mémoire longue dans un modèle de volatilité, il convient de s'interroger sur l'origine de ce phénomène. Trois hypothèses, pouvant agir conjointement ou non, ont été avancées. La plus ancienne, connue sous le nom de théorie de l'agrégation, est due à Granger (1980). Selon lui, l'agrégation de micro-variables recélant de la mémoire courte conduit à la formation d'une variable présentant de la mémoire longue. Quant à la plus récente hypothèse, traitée entre autres par Liu (2000) et Granger et Hyung (2004), elle fait état de changements structuraux occasionnels dans les séries de volatilité, conséquence de chocs financiers de grande ampleur. Ces deux hypothèses apportent une interprétation de ce qui se passe sur le marché. Au contraire, la troisième, récemment étudiée par Deo, Hurvich, et Li (2006) attribue la mémoire longue à la mauvaise spécification des modèles.

Depuis le début des années 1980 et les travaux de Granger (1980) et Granger et Joyeux (1980), la mémoire longue a été à l'origine d'une abondante littérature (voir entre autres Hosking (1981), Geweke et Porter-Hudak (1983), Sowell (1992), Engle et Smith (1999) ou encore Diebold et Inoue (2001)). Toutefois, force est de constater que très peu d'articles débattent de ce que traduit concrètement sur le marché cette caractéristique. Di Matteo, Aste, et Dacorogna (2005) sont parmi les premiers à établir une relation entre la mémoire longue des indices boursiers et l'état de développement du marché : plus les marchés sont matures et liquides, plus la mémoire longue est faible. Or, Dacorogna, Müller, Olsen, et Pictet (2001) supposent que la présence d'un grand nombre d'agents hétérogènes (le critère de différenciation est la fréquence à laquelle ils opèrent) caractérise un marché mature. En conséquence de quoi, la mémoire longue est reliée à l'hétérogénéité des acteurs du marché : elle est d'autant plus forte que l'hétérogénéité des comportements est faible.

A nos yeux, la prise en compte de la mémoire longue est un point clé pour une meilleure appréciation du risque action et constitue un défi comportant deux volets. Premièrement, il est indispensable de mieux interpréter ce phénomène sur les marchés financiers, et cela passe par le développement d'un modèle de volatilité capable de relier la mémoire longue au comportement des investisseurs. Deuxièmement, il convient de s'interroger sur la manière d'intégrer cette caractéristique des séries de volatilité de manière opérationnelle, notamment en gestion de portefeuille et les conséquences que cela implique aussi bien dans le cadre théorique que pratique. En effet, il est primordial, dans ce contexte, de savoir si la théorique classique du choix de portefeuille peut toujours s'appliquer, ou au contraire, s'il faut envisager d'autres moyens de mise en œuvre.

Dans le domaine de la modélisation de la volatilité, la mémoire longue a déjà été introduite, notamment dans les processus fractionnairement intégrés ARFIMA de Granger et Joyeux (1980) et FIGARCH de Baillie, Bollerslev, et Mikkelsen (1996), extensions des modèles ARMA de Box et Jenkins (1970) et GARCH de Bollerslev (1986). Dans ces modèles, la mémoire longue est contrôlée explicitement au moyen d'un paramètre. Bien que de nombreux articles concluent qu'un modèle à mémoire longue est plus efficace pour caractériser une série temporelle telle que la volatilité (voir Deo, Hurvich, et Li (2006) par exemple), le modèle ARFIMA souffre de plusieurs critiques émises, entre autres, par Man (2003) et Ellis et Wilson (2004). En effet, une grande variance dans l'estimation des paramètres nuit à son efficacité. Hélas, toutes ces représentations de la volatilité sont purement économétriques dans le sens où l'estimation de leurs paramètres n'a pas la faculté d'apporter une interprétation de ce qui se passe sur le marché.

Pour relier les paramètres du modèle que nous proposons au comportement des investisseurs, il reste à discuter de la manière de modéliser leurs comportements. Deux possibilités s'offrent à nous. La première a été employée pendant des décennies et jusqu'au début des années 1990. Elle préconisait l'utilisation d'un agent représentatif pour traduire l'agrégation de l'ensemble des comportements des agents opérant sur le marché. La deuxième, apparue à partir des années 1990 suite aux vives critiques à l'encontre de cette approche (Kirman 1992), consiste à prendre en compte l'hétérogénéité des agents. En effet, il s'est avéré que la complexité des relations entre les agents ne pouvait pas être traduite au travers d'un agent unique. C'est pourquoi les modèles comportementaux actuels (voir par exemple Mankiw (2000), Calvet et Common (2003) et Ghoulmie, Cont, et Nadal (2005)) se doivent de prendre en compte l'hétérogénéité des investisseurs.

Les modèles comportementaux d'agents considèrent les marchés financiers comme des systèmes complexes non figés, composés d'un grand nombre d'agents faisant preuve de rationalité limité. Ces agents interagissent d'une part au travers de stratégies d'investissement en adaptant constamment leur comportement à l'arrivée d'information et à la performance de leur stratégie (concept de *adaptively rational equilibrium* introduit par Brock et Hommes (1997)), et d'autre part au travers d'interactions sociales. Selon Hommes (2008), dans un système d'agents interactifs, le marché agrège l'arrivée aléatoire d'information transformée et amplifiée par les agents pour donner lieu à une variable traduisant des faits stylisés importants tels que des prix et des rendements imprévisibles, des bulles spéculatives et des krachs soudains, les *clusters* de volatilité (Alfarano, Wagner, et Lux 2005) ou encore de la mémoire longue.

Après avoir considéré la mémoire longue d'un point de vue théorique et conceptuel, nous nous plaçons dans un contexte axé sur la gestion de portefeuille pour le côté pratique. Dans ce domaine, la modélisation des rendements des actifs financiers joue un rôle central dans le processus de sélection, d'abord qualitatif, puis quantitatif, des constituants du portefeuille. En effet, la modélisation permet d'appréhender de manière rigoureuse deux éléments essentiels que sont le rendement espéré et le niveau de risque. Cet impératif survient en fait de façon évidente dans la théorie du choix d'allocation optimale, que se soit en temps discret ou en temps continu.

Pour modéliser le plus fidèlement possible la volatilité et ainsi obtenir le portefeuille le plus proche du portefeuille optimal, deux caractéristiques doivent impérativement être introduites : la dynamique de la volatilité d'une part et la mémoire longue d'autre part. Or, en gestion de portefeuille, la prise en compte de la mémoire longue entraîne un changement de cadre du problème étudié : de markovien, il devient non-markovien. La conséquence du point de vue pratique est que la stratégie de gestion doit prendre en compte tout le risque passé, contrairement à une approche classique. Du point de vue théorique, les méthodes classiques de résolution des problèmes de choix de portefeuille telles que l'approche par programmation dynamique et l'utilisation de l'équation d'Hamilton-Jacobi-Bellman deviennent inapplicables. Il convient donc de trouver une autre approche.

Après avoir exposé nos motivations, venons en maintenant au plan de la thèse. Dans une première partie, notre objectif est de proposer un modèle à mémoire longue de la volatilité réalisée d'actions afin de fournir une interprétation de ce phénomène. Par conséquent, notre

modèle ne doit pas être uniquement économétrique. En effet, l'estimation de ses paramètres doit pouvoir apporter une explication de ce qui se passe sur le marché. Pour y parvenir, nous avons pour ambition de relier les paramètres de notre modèle au comportement des investisseurs.

Dans le premier chapitre, nous effectuons une revue de littérature concernant les rendements d'action, leur volatilité, la mémoire longue ainsi que la prise en compte de l'hétérogénéité des agents dans la modélisation économique. Cette revue de littérature est l'occasion de passer en revue les faits stylisés des rendements d'action et de leur volatilité mais surtout de s'interroger sur plusieurs points importants intervenant dans le développement d'un modèle comme le nôtre. Premièrement, il s'agit de connaître l'origine de la mémoire longue et sa signification. Deuxièmement, il convient de se demander de quelle manière les paramètres de notre modèle peuvent être reliés au comportement des investisseurs.

Le chapitre 2 est consacré à la présentation de notre modèle. Nous partons du constat que les investisseurs se distinguent les uns des autres par de nombreuses caractéristiques telles que leurs moyens financiers, bien évidemment, mais aussi les informations et les techniques dont ils disposent. Compte tenu du grand nombre d'agents intervenant sur les marchés, cela se traduit par une hétérogénéité de comportements malgré la présence de phénomènes qui confèrent une uniformité de décision comme par exemple le mimétisme (Orléan 1995, Kirman 1993). Par conséquent, notre modélisation doit en tenir compte. Par ailleurs, pour être de bonne qualité, un modèle de volatilité doit pouvoir prendre en compte un certain nombre de faits stylisés et en particulier la mémoire longue présente dans les séries temporelles de la volatilité (Campbell, Lo, et MacKinlay 1997). Conformément à l'hypothèse de Granger (1980), nous supposons que l'origine de la mémoire longue résulte de l'agrégation de micro-variables qui suivent des processus à mémoire courte. En effet, nous supposons que chaque investisseur anticipe une valeur future de la volatilité réalisée, et que l'agrégation de l'ensemble de ces anticipations conduit à la formation de la volatilité observée. Nous supposons également que l'anticipation de la volatilité réalisée de chaque agent va dépendre de quatre facteurs : leur anticipation précédente, la réalisation de la volatilité précédente, des informations publiques communes à tous les investisseurs, et des informations privées. Ce modèle, simple en apparence, est capable de traduire un nombre intéressant de comportements en fonction de l'importance accordée à chacun de ces facteurs. Des investisseurs basant leur anticipation sur l'information disponible auront un comportement rationnel, tandis qu'il feront preuve de rationalité limitée si lors de la formation de leur anticipation, ils accordent davantage d'importance aux évènements passés (anticipation passée, réalisation passée) qu'à l'information disponible.

Le deuxième chapitre s'achève par une étude théorique de notre modèle montrant les différentes formes que peut prendre la rationalité limitée des investisseurs. Puis, nous analysons et discutons de l'impact de chaque paramètre du modèle sur la mémoire longue et la mémoire courte. En particulier, nous mettons en lumière un paramètre, que l'on qualifie de *paramètre d'hétérogénéité* dont la densité fournit de précieux renseignements sur le comportement des investisseurs. Grâce à ce paramètre, nous caractérisons les comportements des investisseurs qui génèrent de la mémoire longue dans les séries de volatilité. Plus précisément, nous relions l'hétérogénéité des comportements des investisseurs au paramètre de mémoire longue.

Après avoir présenté les résultats théoriques de notre modèle, nous explorons l'aspect

empirique dans le chapitre 3. La diversité des comportements des investisseurs se traduit par un paramètre d'hétérogénéité défini par une distribution de probabilité caractérisée par sa densité. Par conséquent, nous commençons par traiter l'inférence statistique du modèle qui consiste à définir et justifier la densité du paramètre d'hétérogénéité. Ensuite, nous réalisons une estimation non paramétrique des deux éléments essentiels du modèle : la densité de la distribution d'hétérogénéité et le paramètre de mémoire longue.

Nous entreprenons une étude empirique à partir des prix à la minute sur une période de dix ans entre le 1er janvier 1994 et le 31 décembre 2003, de 24 actions américaines cotées sur le NYSE ou le Nasdaq, de différentes tailles en terme de capitalisation et représentant différents secteurs d'activités. Pour chaque action, nous calculons les séries des rendements et estimons les séries de volatilité au moyen de l'estimateur de la volatilité réalisée de Barndorff-Nielsen et Shephard (2002b). Nous commençons par vérifier certains faits stylisés importants des rendements et de la volatilité évoqués dans le premier chapitre. Puis, nous estimons les paramètres de notre modèle pour chacune des séries de volatilité réalisée de notre base de données. En particulier, nous obtenons une estimation du paramètre de mémoire longue pour chacune de ces séries que l'on compare à celle obtenue grâce à d'autres estimateurs connus : celui de Hurst (1951) et celui de Geweke et Porter-Hudak (1983). En terme de comportement des acteurs du marché, deux situations distinctes, indépendantes à la fois de l'ordre de grandeur de la capitalisation et du secteur d'activité de l'entreprise apparaissent. Nous clôturons le troisième et dernier chapitre de la première partie, en étudiant le comportement des agents en fonction de la phase du marché. Cette étude est menée en considérant la bulle spéculative des nouvelles technologies de la fin des années 1990. Nous sélectionnons quelques actions d'entreprises appartenant au domaine des nouvelles technologies et d'autres d'entreprises exerçant leur activité dans des secteurs différents. Nous séparons nos séries en deux sous-séries, la première précédant l'éclatement de la bulle et la seconde le suivant. Nous estimons ensuite les paramètres de notre modèle sur chacune de ces sous-séries. Les résultats obtenus montrent clairement que le comportement des investisseurs est affecté par la situation du marché.

Dans la deuxième partie de la thèse, nous avons pour but d'intégrer la mémoire longue de la volatilité dans le développement d'un modèle de choix de portefeuille en temps continu avec volatilité stochastique. En effet, nous avons constaté que la plupart des modèles de choix de portefeuille, y compris les plus récents, ignorent le phénomène de mémoire longue. Or, en la prenant en compte et donc en traduisant encore plus fidèlement le risque des actions, nous pouvons raisonnablement espérer aboutir à une meilleure stratégie d'allocation et donc à une meilleure performance du portefeuille.

Dans le chapitre 4 nous effectuons une revue de littérature des modèles de choix de portefeuille. Etant donné l'abondante recherche dans ce domaine résultant du grand nombre de dimensions à prendre en compte dans la modélisation, nous ne pouvons pas réaliser une revue de littérature exhaustive. Toutefois, nous discutons brièvement de ces différentes dimensions. Puis, nous nous interrogeons, conformément à la problématique retenue dans cette thèse, sur la manière d'introduire la volatilité stochastique en gestion de portefeuille.

Dans le chapitre 5, notre objectif est de proposer un modèle de choix de portefeuille, en temps continu, avec une volatilité stochastique et d'étudier l'impact du caractère stochastique

de la volatilité sur l'allocation optimale. Le portefeuille est constitué d'un actif sans risque et d'un actif risqué dont le cours suit une équation de diffusion avec une dérive constante et une volatilité stochastique, modélisée au moyen de l'exponentiel d'un processus d'Ornstein-Uhlenbeck. En outre, la corrélation entre le rendement de l'actif risqué et sa volatilité est contrôlée de manière classique (voir Chacko et Viceira (2005), Liu (2007) ou Heston (1993) dans un autre domaine).

Nous dérivons les résultats analytiques de notre modèle. Nous traitons trois critères de gestion couramment rencontrés dans la littérature et la fonction d'utilité caractérisant l'aversion absolue au risque hyperbolique, appelée HARA. Pour résoudre notre problème, nous faisons appel à la méthode classique par programmation dynamique et l'utilisation de l'équation d'Hamilton-Jaccobi-Bellman initiée par Merton (1971) et reprise entre autres par Benth, Karlsen, et Reikvam (2003), Chacko et Viceira (2005) et Liu (2007). Nous fournissons les expressions du poids optimal de l'actif risqué et de la consommation optimale le cas échéant. Dans notre cas, l'équation aux dérivées partielles obtenue est trop compliquée pour être résolue analytiquement. Par conséquent, nous développons une méthode de résolution numérique fondée sur les différences finies.

Dans l'étude de l'impact du caractère stochastique de la volatilité sur l'allocation optimale, nous prenons en compte deux critères. Premièrement, le fait que la volatilité prend dans le temps des valeurs plus grandes ou plus petites que sa valeur moyenne, que nous appelons *stochasticité* de la volatilité. Deuxièmement, il faut tenir compte de l'ampleur moyen de la stochasticité que nous qualifions de *degré de stochasticité* qui agit tel un degré d'incertitude sur le niveau de volatilité. Nous mettons en lumière les paramètres le contrôlant. Par ailleurs, nous montrons quel est impact du degré de stochasticité sur le plan de consommation et d'investissement et l'influence du temps restant avant l'échéance. En particulier, nous observons dans quelle mesure l'augmentation de l'incertitude induit un changement de comportement de l'investisseur.

Le sixième et dernier chapitre a pour objet de présenter notre modèle de choix de portefeuille, en temps continu, avec une volatilité stochastique à mémoire longue et d'étudier l'impact de la mémoire longue de la volatilité sur l'allocation optimale. Ce modèle est développé dans la logique du modèle précédent : le processus d'Ornstein-Uhlenbeck est remplacé par un processus d'Ornstein-Uhlenbeck fractionnaire faisant appel au mouvement brownien fractionnaire, introduit par Mandelbrot et Wallis (1968), qui confère une corrélation sérielle aux incréments du processus. La particularité d'une telle modélisation est double puisque d'une part la mémoire longue de la volatilité est contrôlée explicitement par un paramètre, et d'autre part le fait d'utiliser un mouvement brownien fractionnaire transpose le système économique dans un cadre non markovien. En raison de ce deuxième point, la situation se complique singulièrement. En effet, l'approche classique par programmation dynamique et l'utilisation de l'équation d'Hamilton-Jacobi-Bellman n'est plus adaptée. Nous recourons donc à une autre méthode et traitons uniquement le problème de maximisation de l'héritage qui est en réalité un pur cas de choix de portefeuille. Pour résoudre ce problème, nous utilisons la théorie du calcul stochastique appliquée au mouvement brownien fractionnaire. Nous dérivons la stratégie de gestion optimale et nous formulons un algorithme de résolution récursif fondé sur la technique de Monte Carlo.

Le chapitre 6 s'achève sur l'étude de l'impact de la mémoire longue sur l'allocation optimale. Nous commençons par nous intéresser au rôle joué par la variable contrôlant la mémoire longue sur les trajectoires de la log-volatilité. Cela nous permet de fixer les paramètres de notre modèle et de débuter l'étude. Nous mesurons l'impact de la mémoire longue sur la performance du portefeuille dans différentes situations de gestion, par exemple pour différents horizons de gestion, ou encore pour un investisseur plus ou moins risquophobe. Enfin, nous discutons de l'importance de la bonne spécification du modèle de choix de portefeuille pour obtenir une allocation optimale de la richesse du portefeuille.

Cette thèse se termine par une annexe dans laquelle nous menons une réflexion personnelle sur le travail de thèse qui aborde les différentes dimensions d'une telle entreprise (annexe A), notamment les éléments de contexte, l'évolution du projet ainsi que les compétences acquises et les enseignements personnels.

Première partie

Modélisation de la volatilité

Introduction

La lente décroissance hyperbolique de la fonction d'auto-corrélation caractérisant de nombreuses séries temporelles économiques et financières a été le sujet de nombreuses recherches depuis plus de deux décennies. Plusieurs explications et modèles ont été fournis pour tenter d'expliquer ce phénomène, aussi connu sous le nom de phénomène de mémoire longue. Parmi les explications les plus pertinentes proposées jusqu'à présent, trois principaux mécanismes, pouvant agir séparément ou conjointement retiennent particulièrement l'attention : d'abord (i) l'approche par agrégation suggérée par Granger et Joyeux (1980) qui ont montré que les séries temporelles résultant de l'agrégation de micro-variables à la mémoire courte peut conduire à de la mémoire longue, puis (ii) la présence de rares changements de régime qui permet de reproduire la non-stationnarité à long terme de l'activité économique et financière (Hamilton 1989, Diebold et Inoue 2001, Gouriéroux et Jasiak 2001, Granger et Hyung 2004, Gadea, Sabaté, et Serrano 2004, entre autres), et enfin (iii) la présence de non-linéarités dans les systèmes économiques et financiers (voir Davidson et Sibbersten (2005) pour une revue de la littérature).

Notre but dans cette partie est de fournir un modèle reliant la mémoire longue de la volatilité réalisée des rendements d'actifs, qui est une caractéristique importante des séries temporelles financières (Taylor 1986, Ding, Granger, et Engle 1993, Dacoragna, Müller, Nagler, Olsen, et Pictet 1993, pour les travaux pionniers), au comportement hétérogène des agents économiques. Nous fondant sur le fait que, sur le marché, les participants forment des anticipations hétérogènes concernant le niveau futur de la volatilité et qu'ils agissent comme des agents à rationalité limitée (Simon 1957, Simon 1991), nous proposons une explication du phénomène de mémoire longue qui repose sur l'agrégation par le marché des opinions hétérogènes des investisseurs, révélées au travers du processus de fixation des prix du marché.

L'hétérogénéité des investisseurs intervenant sur un marché est à présent un fait bien reconnu – en particulier par les partisans de la finance comportementale (Barberis et Thaler (2003) pour une revue de la littérature) – qui peut prendre plusieurs formes. L'hétérogénéité peut tout d'abord être considérée comme résultant de la diversité de la nature, de la taille, et de la stratégie des agents économiques : des investisseurs individuels, des traders ou des investisseurs institutionnels, pour ne citer qu'eux, n'ont pas les mêmes ressources financières (dépendant de leur taille), les mêmes buts (des profits à court terme ou long terme, ou des fréquences d'allocation par exemple) et les mêmes qualités et aptitudes. Toutes ces différences les font se concentrer sur différents critères et, par conséquent, anticiper différemment la valeur future des entreprises.

A cet égard, Di Matteo, Aste, et Dacorogna (2005) ont récemment suggéré l'existence d'une relation entre le degré d'hétérogénéité parmi les intervenants du marché et le degré de développement des marchés financiers. En effet, en rattachant le degré de maturité d'un marché financier à la vitesse de la décroissance hyperbolique de la fonction d'auto-corrélation de la volatilité des actifs échangés sur le marché étudié, Di Matteo *et al.* montrent que plus le marché est mature et efficient, plus grand est le nombre de classes d'agents différentes et plus petit est l'effet de la mémoire longue, en accord avec l'approche de Granger. De plus, d'autres phénomènes tels que la psychologie de masse et la contagion doivent être pris en compte. En effet, il a été clairement montré que les rumeurs (Banerjee 1993), le mimétisme (Orléan 1995) ou effet moutonnier (Banerjee 1992, Froot, Scharfstein, et Stein 1992, Kirman 1993, Cont et Bouchaud 2000), les modes (Shiller 1989) affectent le comportement des agents.

Afin de justifier ces différentes sources d'hétérogénéité, mais de garder encore une représentation parcimonieuse, nous proposons un modèle qui établit un lien entre le comportement des investisseurs et un petit nombre de paramètres d'hétérogénéité qui permettent de justifier la tendance individuelle de ces investisseurs à gérer leur portefeuille sur la base des informations qu'ils reçoivent ou, au contraire, sur la base d'une approche auto-référentielle qui les conduit à s'intéresser principalement à leurs anticipations passées et leurs observations passées de la volatilité du marché. Cette approche nous permet de retenir qu'en plus de l'hétérogénéité des tailles et des stratégies, une autre source importante d'hétérogénéité provient du fait qu'en réalité les agents anticipent les nombreux facteurs qui influencent les gains futurs de la firme et leur volatilité. Ces facteurs sont capturés, dans notre modèle, à l'aide d'un flux d'informations publiques et privées qui peut être considéré comme englobant plusieurs variables macro-économiques affectant la volatilité des actions telles que les cycles économiques (Schwert 1989), le prix du pétrole dont la volatilité est importante dans l'explication de la volatilité des rendements d'actifs technologiques (Sadorsky 2003), ou l'inflation et les taux d'intérêts qui ont de grands impacts sur la volatilité des marchés d'actions (Kearney et Daly 1998).

Ensuite, nous établissons le lien entre les propriétés distributionnelles des paramètres d'hétérogénéité et les fonctions d'auto-covariance/auto-corrélation de la volatilité réalisée. Cela nous permet d'examiner le genre de comportements économiques pouvant produire de la mémoire longue dans les séries temporelles de la volatilité réalisée. En particulier, cela nous permet de mettre en perspectives les affirmations de Di Matteo, Aste, et Dacorogna (2005) selon lesquelles plus les classes d'agents et les stratégies différentes sont nombreuses, plus petit est l'effet de la mémoire longue. Enfin, le calibrage de notre modèle sur la dernière décennie à partir d'une base de données d'actions américaines de moyenne et grande capitalisation, nous permet de reconstruire la distribution des paramètres d'hétérogénéité et ainsi d'avoir accès au comportement global des investisseurs. En fonction de la phase du marché considérée, différents comportements sont observés. En particulier, nous constatons (i) une forte tendance à faire référence à leurs propres anticipations passées avant le krach de la bulle internet, et (ii) une augmentation de la proportion d'investisseurs effectuant leurs prédictions sur la base des informations leur parvenant après le krach.

Cette première partie est organisée de la manière suivante. Dans le chapitre 1, nous effectuons une revue de la littérature rappelant les faits stylisés élémentaires concernant les rendements d'actifs puis la volatilité réalisée, une mesure de la volatilité présentée par An-

dersen, Bollerslev, Diebold, et Ebens (2001) et Barndorff-Nielsen et Shephard (2002a), parmi d'autres. Nous y présentons en détail le phénomène de mémoire longue, de l'origine de sa présence dans les séries temporelles économiques et financières aux modèles mathématiques qui lui sont consacrés. Enfin, nous achèverons ce chapitre en discutant de la prise en compte de l'hétérogénéité des comportements d'agents dans la modélisation des séries temporelles économiques et financières, par opposition à la notion d'agent représentatif longtemps employée par le passé. Ensuite, dans le chapitre 2 nous présentons notre modèle de volatilité d'actions dans lequel les investisseurs à rationalité limitée ont des opinions hétérogènes. Nous étudions à la fois l'impact de la rationalité limitée et celle des opinions hétérogènes des agents dans un marché supposé agréger les anticipations individuelles. Enfin, dans le troisième chapitre nous discutons des questions de calibrage du modèle et dérivons la loi asymptotique de l'estimateur des paramètres du modèle. Nous présentons et discutons également les résultats empiriques obtenus par le modèle à partir de 24 actions américaines. Cette section et ce chapitre s'achèvent en tentant de répondre à la question suivante : "Les agents adoptent-ils un comportement différent en fonction de la phase du marché ?".

Chapitre 1

Revue de la littérature

Depuis le début du XXème siècle et la thèse de Bachelier (1900), et probablement depuis que l'on a commencé à reporter le cours des actions sur un graphique, les professionnels de la finance tout comme les chercheurs se sont demandés dans quelle mesure l'historique des prix d'un actif financier pouvait être utilisé pour prédire le prix futur de cet actif ? Autrement dit, existe-t-il une corrélation entre les prix successifs d'une action ? Concernant cette question, deux points de vue radicalement différents sont apparus et se sont opposés : certains étaient persuadés de l'absence de corrélation dans la série des rendements d'action et comparaient le processus de prix à une marche aléatoire tels Osborne (1959) et Samuelson (1965), alors que d'autres assuraient que le processus de prix connaissait des phases ou des cycles et que le passé avait alors un impact sur le présent (Bechu, Bertrand, et Nebenzahl 2008).

Récemment, Cont (2001) a passé en revue les faits stylisés des rendements d'actions [1]. Il rappelle que l'autocorrélation dans la série des rendements instantanés [2] n'excède pas 2 lags (autrement dit la durée de 2 cotations). Au contraire, il constate que l'autocorrélation de la série de la valeur absolue des rendements décroît beaucoup plus lentement dans le temps, de façon hyperbolique : elle persiste à plus de 100 lags. Ce résultat reste valable à l'échelle journalière (Campbell, Lo, et MacKinlay 1997). Cette persistance à long terme des autocorrélations est appelée mémoire longue. Or, le fait que l'on trouve de la mémoire longue dans la série de la valeur absolue des rendements indique qu'il y a de la mémoire longue dans les séries de la volatilité d'actifs.

En finance, la volatilité est utilisée comme une mesure du risque des actifs dans de nombreux domaines tels que, entre autres, la gestion du risque ou les méthodes d'allocation optimale d'actifs en gestion de portefeuille. Par exemple, en gestion de portefeuille, si l'estimation du risque des actions constituant le portefeuille est éloignée de leur vraie valeur, alors l'allocation n'est pas optimale et les conséquences peuvent être graves. L'estimation de la

1. Le rendement du prix d'une action sur une période donnée est défini comme étant la différence entre le logarithme du prix de l'action à la fin de la période et le logarithme de son prix au début de la période. Si $S(t)$ est le cours de l'action à l'instant t, alors le rendement de cette action sur la période Δt vaut $r(t, \Delta t) = \ln S(t + \Delta t) - \ln S(t)$.

2. Un rendement instantané correspond au rendement entre deux cotations successives.

volatilité se révèle donc cruciale. C'est pourquoi, de nombreux estimateurs et modèles ont été développés ces dernières décennies. Il existe une grande disparité en terme d'efficacité entre les premiers estimateurs n'utilisant que deux données par jour, le minimum et le maximum du cours de l'action comme l'estimateur de Parkinson (1980), et les plus récents utilisant les données cotation par cotation comme l'estimateur de Andersen, Bollerslev, Diebold, et Ebens (2001). On retrouve également cette disparité dans les modèles de volatilité. En effet, les modèles capables de prendre en compte le phénomène de mémoire longue tels que le modèle ARFIMA de Granger et Joyeux (1980) ou FIGARCH de Baillie, Bollerslev, et Mikkelsen (1996) s'avèrent plus efficaces que les autres tels que le modèle ARMA de Box et Jenkins (1970) et le modèle GARCH de Bollerslev (1986).

Trois principales explications, pouvant agir individuellement ou conjointement ont été avancées pour justifier la présence de mémoire longue dans ce type de séries temporelles financières et économiques. La moins intéressante étant donné qu'elle n'apporte aucune interprétation de ce qui se passe sur le marché, étudiée notamment par Chen et Deo (2006), concerne la mauvaise spécification des modèles. La plus récente, traitée par Liu (2000), de Peretti (2003) ou encore Granger et Hyung (2004), fait état, dans les séries de rendements, de la présence de ruptures structurelles occasionnelles résultant de chocs financiers de grande ampleur. C'est le cas par exemple de l'éclatement de la bulle internet au début des années 2000, les deux chocs pétroliers des années 70, etc... Enfin, l'explication la plus ancienne, suggérée par Granger (1980), nous dit que l'agrégation de micro-variables présentant de la mémoire courte conduit à la formation d'une variable présentant de la mémoire longue. Cette théorie de l'agrégation a été reprise récemment, dans un domaine un peu différent, par Calvet et Common (2003) dans un modèle de la demande des consommateurs qui agrège leurs goûts hétérogènes.

La théorie de l'agrégation prend tout son sens lorsque l'on considère les travaux récents sur l'hétérogénéité des agents intervenant sur le marché. En effet, bien qu'il y ait des phénomènes de rumeurs (Banerjee 1993), de mimétisme (Orléan 1995), d'effets moutonniers (Banerjee 1992, Froot, Scharfstein, et Stein 1992, Kirman 1993, Cont et Bouchaud 2000), ou encore des effets de modes (Shiller 1989), l'hétérogénéité des investisseurs demeure, dans une certaine mesure, en raison de leur diversité. En effet, ces derniers se différencient tout d'abord par leur taille. Cela implique des moyens d'actions différents : du particulier à l'investisseur institutionnel en passant par les traders et gestionnaires de portefeuille, tous n'ont pas les mêmes ressources financières (quelques dizaines ou centaines d'euros pour des particuliers aux moyens limités, quelques dizaines de millions d'euros pour les gestionnaires de portefeuille, quelques milliards d'euros pour les institutions [3]). En outre, les investisseurs n'ont pas les mêmes connaissances des marchés, les mêmes stratégies (profit à court terme voire très court terme dans le cadre de petits profits immédiats ou bien à long terme dans le cadre d'une assurance vie), accès aux mêmes informations au même moment, etc...

Ce premier chapitre consacré à la littérature va nous permettre de mieux appréhender le contexte historique et les enjeux majeurs de la modélisation de la volatilité des rendements d'actifs. Il est organisé de la manière suivante. Nous commençons par consacrer une partie aux rendements d'actifs puisque d'une part comprendre le comportement des rendements permet dans une certaine mesure de comprendre celui de la volatilité et d'autre part cela nous permet

3. Le lecteur est invité à se référer à Gabaix, Gopikrishnan, Plerou, et Stanley (2006).

d'introduire l'intérêt de modéliser la volatilité. Puis, nous nous intéressons à la volatilité des rendements. Dans un troisième temps, nous abordons la problématique du phénomène de mémoire longue. Enfin, la dernière partie de cette revue de la littérature est axée sur l'intérêt grandissant manifesté à l'égard de la théorie de l'interaction d'agents hétérogènes sur le marché et sa prise en compte dorénavant quasi-systématique dans les modèles d'agents comme par exemple les modèles de Mankiw (2000), Calvet et Common (2003), Ghoulmie, Cont, et Nadal (2005).

1.1 Les rendements d'actifs

Les investisseurs ont pour objectif de générer des profits. Ceux qui opèrent sur le marché des actions se sont donc intéressés en premier lieu au cours des actions ou à leurs rendements. Les professionnels et les chercheurs ont analysé les séries des prix ou celles des rendements selon deux axes : l'un concerne les facteurs influençant le cours des actions et donc les rendements, et l'autre les caractéristiques des séries des rendements. En particulier, l'objectif de ce deuxième axe était de mettre en lumières des faits permettant de prévoir les rendements.

1.1.1 Facteurs influençant les rendements

Les principaux facteurs qui, à notre connaissance, ont fait l'objet d'études sont les informations diverses (telles que les recommandations, l'annonce des résultats trimestriels ou annuels, ou encore l'annonce du changement de dividendes), les manipulations et les facteurs psychologiques et comportementaux.

Les informations peuvent être classées en deux catégorie différentes : l'une regroupant les informations émises par les compagnies comme les annonces concernant le versement de dividendes, la signature de contrats, les prévisions ou encore les résultats, et l'autre concernant les informations extérieures aux compagnies, telles que les recommandations de banques ou sociétés de courtage ou les annonces gouvernementales par exemple.

D'après Fama, Fisher, Jensen, et Roll (1969), l'annonce du changement du montant des dividendes est une information très importante puisque Lintner (1956) et Michaelsen (1961) ont démontré qu'une fois les dividendes augmentés, les firmes étaient très réticentes à les baisser dans l'avenir. Par conséquent, suite à l'annonce de l'augmentation des dividendes, les directeurs envoient le signal selon lequel ils sont sûrs de pouvoir maintenir le niveau des dividendes les années suivantes. Cela implique qu'ils ont confiance en leur entreprise et leur capacité à dégager de gros bénéfices dans les mois, voire les années futures. C'est l'hypothèse du signal. Cela a été confirmé plus récemment par Baker, Farrelly, et Edelman (1985), Baker et Powell (1999) Baker, Powell, et Veit (2002). Un peu plus tôt, les travaux de Bhattacharya (1979), John et Williams (1985), Miller et Rock (1985) ont permis de conclure que les managers choisissaient le niveau de dividendes et l'augmentation de dividendes comme le moyen de communiquer une information privée aux investisseurs. De manière similaire, d'après DeAngelo, DeAngelo, et Skinner (1992) et Kumar et Lee (2001), l'annonce d'une

diminution des dividendes est aussi une information indiquant que les dirigeants redoutent une baisse des revenus. Toutefois, l'hypothèse du signal a été mise à rude épreuve par Benartzi, Michaely, et Thaler (1997) qui ont trouvé que lorsque les dividendes augmentent, les revenus n'augmentent pas par la suite, et lorsque les dividendes baissent, les revenus augmentent dans le futur. Selon eux le versement de dividendes est un signal qui informe sur le passé et le présent et non pas sur le futur. Malgré cela, ils reconnaissent que les compagnies ayant augmenté leurs dividendes ont une plus faible probabilité de voir leur revenus diminuer l'année d'après que les compagnies n'ayant pas changé le montant de leurs dividendes. C'est pourquoi ils concluent que "le modèle de dividendes de Lintner demeure la meilleure description du processus de fixation de dividende disponible".

Quant aux recommandations, toutes les études consacrées à leur impact à court terme sur le prix des actions s'accordent sur le fait que les recommandations ont une influence au moment de leur publication. Par exemple, Davies et Canes (1978) et Liu (1990) estiment que les recommandation d'achat (respectivement de vente) parues dans la colonne "Heard on the Street" du Wall Street Journal sont suivie d'un rendement positif (resp. négatif) significatif. Trahan et Bolster (1995) aboutissent la même conclusion avec un autre journal. Par ailleurs, Bjerring, Lakonishok, et Vermaelen (1983) affirment q'un investisseur suivant à la lettre les recommandations obtiendrait un meilleur rendement malgré les frais de transaction. En revanche, des désaccords apparaissent lorsqu'il s'agit de juger l'effet à long terme : pour Bjerring, Lakonishok, et Vermaelen (1983) l'impact se prolonge à long terme alors que Davies et Canes (1978), Pari (1987) ou encore Trahan et Bolster (1995) concluent le contraire. En particulier, ces derniers ont montré qu'au bout d'un mois, la performance des actions recommandées chutent fortement, puis disparaissent au-delà.

La manipulation est un autre facteur influençant le cours des actions. Elle intéresse particulièrement les agents puisque c'est un moyen de faire des profits de manière quasi-certaine. Il en existe principalement trois formes : (i) la manipulation d'informations, prohibée, qui consiste à diffuser de fausse informations dans le but modifier brutalement et significativement le cours une action et d'en tirer profit, par exemple au moyen d'une vente à découvert si l'on a fait chuter le prix. (ii) la manipulation basée sur une action, elle aussi prohibée, qui concerne les actions volontaires changeant la valeur réelle des actions. Un exemple célèbre concerne les dirigeants de l'American Steel and Wire Company qui, en 1901, après avoir vendu à découvert des actions de leur compagnie, ont fermé leur aciéries. Le prix de l'action a perdu 33%, chutant de 60$ à 40$. Ils ont alors racheté leurs actions, puis ont fait ré-ouvrir les aciéries pour faire remonter le cours. (iii) la manipulation basée sur les échanges [4], légale, dont le principe est qu'il est possible de manipuler le cours des actions en intervenant sur le marché. En achetant une grande quantité d'actions, on peut faire croire que l'on dispose d'une information privée, et ainsi influencer d'autres investisseurs. Cela se répercutera sur le prix de l'action.

Enfin, le lien entre la psychologie et le comportement [5] est un fait établi depuis longtemps. Ainsi, tous les facteurs influençant l'état psychologique des individus sont susceptibles d'influencer le prix des actions, par l'intermédiaire des agents du marché. Par exemple, des

4. Le lecteur est invité à se référer à Allen et Gale (1992), Kumar et Seppi (1992).

5. Pour en savoir plus sur la finance comportementale, le lecteur est invité à se référer, aux nombreux travaux de Kahneman, Tversky et Thaler entre autres.

études psychologiques ont montré que lorsqu'il fait beau, les gens sont de meilleure humeur que quand il fait gris ou qu'il pleut. Or l'humeur, et les émotions en général, influencent la prise de décision humaine. Dans ce sens, Rind (1996) a montré une relation entre le soleil et le fait de donner un pourboire, tandis que l'absence de soleil a été relié à la dépression par Eagles (1994) ou encore au taux de suicide par Tietjen et Kripke (1994). Cela a donné l'idée à Hirshleifer et Shumway (2003) d'étudier la relation entre le temps qu'il fait sur les places financières et les rendements des actions. Concernant les agents de marché, le soleil entraînerait une bonne humeur qui les rendrait optimistes et les ferait acheter plus facilement des actions. Plus précisément, la bonne humeur est associée à une diminution du sens critique et une augmentation de l'optimisme lors du traitement d'une information, alors que la mauvaise humeur engendre un scepticisme et une analyse plus attentive et plus approfondie. Leur étude, portant sur 26 marchés d'actions de 1982 à 1997, conclut que les journées ensoleillées sont positivement et fortement corrélées au rendement des actions, tandis que les jours de pluie et de neige ne le sont pas. Par exemple, à New York, le rendement du marché des jours parfaitement ensoleillés est de 24.8%, tandis qu'il n'est que de 8.7% lorsque le temps est nuageux. Toutefois, à cause des frais de transactions, il n'est pas possible aux agents de faire des profits avec une stratégie basée sur la météo.

1.1.2 Faits stylisés des rendements d'actifs

Les caractéristiques des rendements des actions admises par tous ont été résumées récemment par Cont (2001). Il en existe principalement deux particulièrement importantes. La première concerne la nature de la densité de la série des rendements qui, à cause de queues épaisses et d'un pic au coeur de la distribution, s'écarte d'une loi normale. La seconde, et peut être la plus importante, assure que la fonction d'autocorrélation de la série des rendements décroît si rapidement qu'au bout de quelques minutes l'autocorrélation disparaît, ce qui est synonyme d'absence de mémoire.

Dans cette section, dans un premier temps, nous illustrons l'écart à la normalité de la densité des rendements d'actifs au moyen de graphiques, puis dans un deuxième temps nous discutons et appuyons, toujours au moyen de graphiques, la théorie de la marche aléatoire.

Écart à la normalité de la densité des rendements

Andersen, Bollerslev, Diebold, et Labys (2001) rappellent que la densité des prix d'une action, présentant un pic au coeur de la distribution et faisant apparaître des queues épaisses, s'écarte de la densité d'une loi normale. C'est ce que nous avons choisi d'illustrer sur la figure 1.1 : nous y avons reporté trois graphiques montrant la distribution de la série des rendements à 1 minute, à 30 minutes et journaliers de l'action Cisco Systems sur une période de 10 ans à l'horizon fin 2003, obtenus grâce à un estimateur par noyau [6]. Pour chaque graphique, nous avons superposé la densité de la loi normale qui se rapproche le plus de chaque distribution des rendements.

6. Le lecteur est invité à se référer à Pagan et Ullah (1999)

Le troisième graphique (en partant du haut) de cette figure concerne la distribution des rendements journaliers. Un zoom a été réalisé sur les queues de la distribution pour mieux se rendre compte du phénomène de queues épaisses. Nous constatons que, bien que la distribution de la série des rendements journaliers semble symétrique par rapport à un axe vertical centré en zéro, elle s'écarte très nettement de la densité d'une loi normale par deux critères : premièrement les queues de la distribution des rendements sont bien plus épaisses [7] que celle de la loi normale et deuxièmement, autour de zéro, la distribution des rendements est beaucoup plus pointue que celle d'une loi normale, on dit qu'elle est leptokurtique [8].

En outre, Cont (2001) ajoute que plus l'intervalle de temps entre deux prix est faible, plus l'écart à une densité de loi normale est flagrant. Ce fait avait déjà été observé dans les années soixante par Mandelbrot (1963) lorsqu'il a révélé l'incapacité de la distribution gaussienne à modéliser la distribution des rendements du prix du coton de 1890 à 1937 et Fama (1965) lorsqu'il a étudié la normalité des rendements de 30 actions du Dow Jones sur une période d'environ 5 ans à horizon septembre 1962, alors qu'à cette époque l'accès aux données hautes-fréquences n'était pas d'actualité.

En observant la figure 1.1, on constate que la distribution de la série des rendements à 1 minute, s'écarte radicalement d'une distribution de loi normale : elle est très pointue autour de zéro et présente des bosses. Ces bosses ont déjà été observées, entre autres, par Gouriéroux et Jasiak (1998) lorsqu'ils ont tracé la distribution des rendements intraday de l'action Alcatel (ils disposaient des relevés de prix toutes les 52 secondes de juillet à août 1995) et Le Courtois (2003) lorsqu'il s'est intéressé à la modélisation des prix d'actions au moyen de processus de Lévy. Ce phénomène peut s'expliquer par l'effet de discrétisation du prix des actions : par exemple sur Euronext une action est cotée au 1/100 d'euro et sur le Nasdaq ou le NYSE au 1/100 de dollar.

Lorsque que l'on s'intéresse à une fréquence plus faible (toutes les 30 minutes), la distribution des rendements se rapproche un peu plus d'une loi normale. Toutefois, les queues très épaisses et le pic au coeur de la distribution ne plaident pas en la faveur d'une distribution de loi normale.

Enfin, quand on réduit encore la fréquence (rendements journaliers), l'épaisseur des queues et l'amplitude du pic diminuent. En conclusion, on observe effectivement que la fréquence des observations influence l'écart à la normalité : plus la fréquence est faible, plus la distribution des rendements tend vers une loi normale, ce qui est conforme au théorème de la limite centrale (Campbell, Lo, et MacKinlay 1997).

Ce résultat est en relation directe avec les travaux de Ané et Geman (2000) qui affirment qu'estimer les rendements sur des intervalles de temps fixes, qu'ils appellent "temps calendaire", ne convient pas. Il faut des intervalles de temps calendaire de l'ordre d'un mois pour que la distribution des rendements soit normale. Ils prétendent qu'au contraire il faut utiliser des intervalles de temps stochastiques, nommés "temps des affaires" ou "temps opérationnel",

7. Le lecteur est invité à se référer à Engle (1982) qui avance l'argument de l'hétéroscédasticité pour expliquer ce trait.

8. Pour Resnick et Samorodnitsky (1997) ce phénomène est la conséquence des caractéristiques des séries temporelles des rendements d'action et en particulier de la corrélation à long terme.

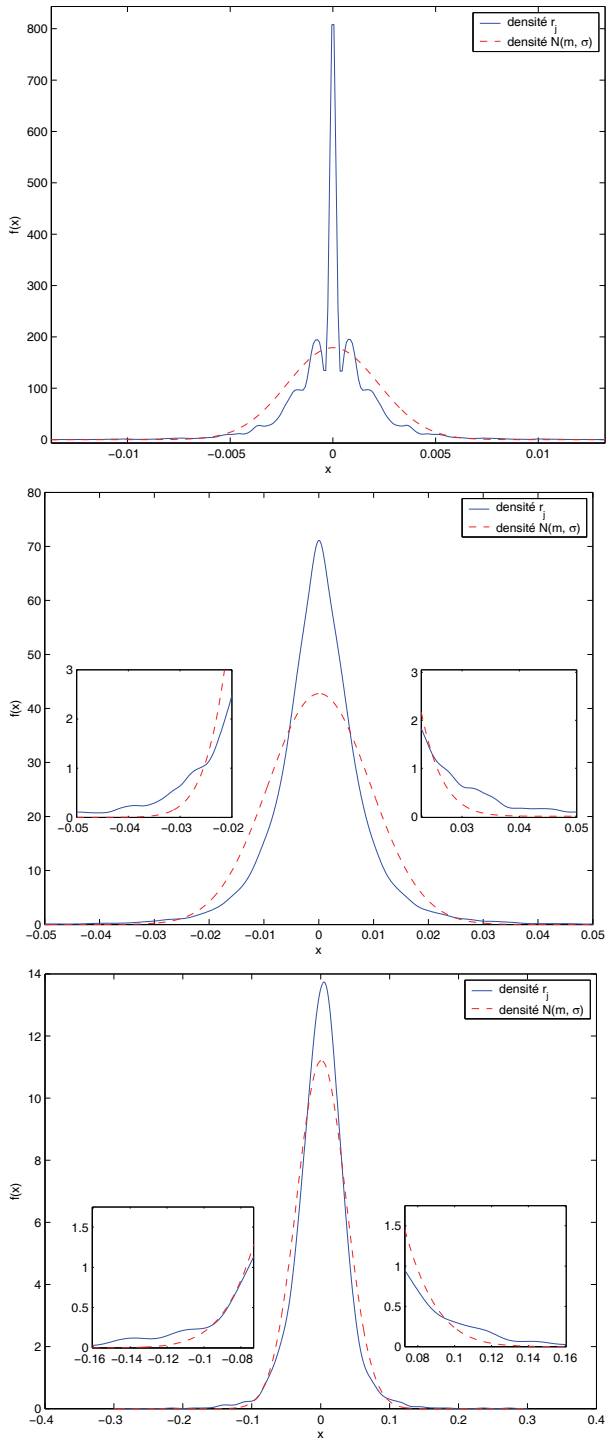

FIGURE 1.1 – Densité des rendements de l'action Cisco Systems sur une période de 10 ans à horizon fin 2003 et densité de la loi normale la plus proche pour trois échelles de temps : en haut les rendements à 1 minute, au milieu les rendements à 30 minutes et en bas les rendements journaliers.

dépendant du nombre d'échanges effectués durant ceux-ci. Ces travaux approfondissent ceux de Clark (1973) qui avait relié l'écart à la normalité des rendements d'actifs à l'existence d'une variation du volume durant la période d'échanges en introduisant en finance les processus subordonnés. Ané et Geman (2000) ont montré, en s'appuyant sur des données hautes fréquences, que c'est le nombre d'échanges [9] et non le volume qui définit le mieux de temps des affaires.

Dans la section 3.2.1 nous justifierons ces faits stylisés en appliquant des tests statistiques à notre jeu de données que l'on aura préalablement détaillé.

Absence de mémoire dans la série des rendements

S'interroger sur l'absence de mémoire dans la série des rendements revient à se demander si le processus de prix des actions suit une marche aléatoire, ou pas. C'est ce qu'ont fait de nombreux chercheurs et professionnels depuis le début des années 60. Cela a donné lieu à deux théories totalement différentes.

Pour les professionnels du marché, dont l'objectif est d'anticiper les mouvements futurs des prix des actions, il ne fait aucun doute que la réponse à cette question est non. Il existe principalement deux méthodes, pour prédire les prix futurs : la théorie chartiste (ou technique) et la théorie de l'analyse de la valeur fondamentale [10] (ou intrinsèque). Par la suite, les adeptes de la première méthode seront qualifiés de chartistes, et ceux de la deuxième méthode, d'analystes. Le point de vue des chartistes est que le passé se répète. Ainsi, leur but est d'observer l'historique des prix d'une action pour y isoler des phases et de reconnaître des phases similaires dans le présent. De cette façon, ils pourront prédire le comportement futur de la série des prix de l'action considérée. Les analystes, quant à eux, estiment que le prix d'une action reflète, en moyenne, la valeur fondamentale de la compagnie qui dépend des revenus potentiels et espèrent tirer profit des divergences passagères. Ces derniers peuvent résulter de facteurs fondamentaux tels que l'économie, la santé du secteur d'activité, la qualité de la gestion, les ressources, les goûts des consommateurs, etc... En résumé, si l'on considère que le prix réel d'une action tend vers la valeur fondamentale de la compagnie, alors prédire la valeur fondamentale revient à prédire le prix futur. Selon Fama (1965), le marché compte très peu de chartistes et beaucoup d'analystes.

A l'opposé des professionnels du marché, un grand nombre de chercheurs, depuis le début des années 60, à l'instar de Bachelier (1900), plaident en faveur de la marche aléatoire : il n'y a pas plus de corrélation (ou mémoire) dans la série des prix que dans une série de nombres aléatoires. Fama (1965) argumente en faveur de l'absence de mémoire dans les séries des prix des actifs financiers. A chacune des critiques envers la théorie de la marche aléatoire, à

9. Jones, Kaul, et Lipson (1994) l'avaient aussi montré.

10. C'est l'hypothèse de Fama dans les années 60 et de Samuelson (1965) : L'évolution des prix est déterminée par les fondamentaux que sont les ressources, les technologies, les préférences des consommateurs, les structures des marchés. Cette philosophie a été remise en cause depuis notamment par Orléan (1999) et Orléan, Bourghelle, Brandouy, et Gillet (2005) : ils ne pensent pas qu'il existe ex ante une "valeur intrinsèque" des entreprises, aussi appelée "valeur objective" ou encore "valeur fondamentale" et remettent en cause de manière radicale l'idée d'efficience informationnelle.

savoir le principe de la valeur fondamentale, l'arrivée d'information et la présence de leaders d'opinion, il oppose son raisonnement pour faire accepter l'idée que l'hypothèse de la marche aléatoire, bien qu'elle ne soit probablement pas une description exacte du comportement des prix du marché, est la meilleure. Ainsi, il n'y a pas de mémoire dans les séries des prix des actions.

Depuis 1965, malgré l'argumentation en faveur de l'indépendance [11] et les démonstrations empiriques (Mandelbrot 1963, Fama 1965, entre autres), de nombreux chercheurs ont continué à s'intéresser à ce sujet. Par exemple, Chow, Denning, Ferris, et Moronha (1995), Cheung et Lai (1995), tout comme Lo (1991) ne trouvent pas de preuves en faveur de la présence de mémoire dans les séries de rendements.

A la fin des années 80, des chercheurs se sont intéressés à l'existence de retour à la moyenne dans les prix d'actions. Si le prix d'une action suit un processus de retour à la moyenne, il existe alors une tendance qui incite le niveau de prix à revenir à un niveau antérieur. Les investisseurs seraient alors capables de prédire les rendements futurs en se basant sur les rendements passés. Les premiers à avoir apporté des preuves de la présence de retour à la moyenne dans des séries de prix d'actions américaines sont Fama et French (1988) et Poterba et Summers (1988). Au même moment, d'autres chercheurs comme Lo et MacKinlay (1988) ont abouti à la conclusion contraire. Plus nuancés, Kim, Nelson, et Startz (1991) prétendent que le retour à la moyenne n'est présent que dans les données d'avant-guerre. Quant à Richardson et Stock (1989) et Richardson (1993) après avoir corrigé des biais dans les séries utilisées par Fama et French (1988) et Poterba et Summers (1988) concluent que leurs résultats pourraient être inversés. Plus récemment, Chaudhuri et Wu (2003) se sont attelés à déterminer ce qui, entre une marche aléatoire et un processus de retour à la moyenne, caractérise le mieux l'indice des prix de 17 marchés émergents. Dans leur étude, ils prennent en compte les changements structurels qui peuvent se produire sur de tels marchés : c'est le cas pour 14 des 17 marchés avec un intervalle de confiance de 95%. Ces changements de structure sont souvent la conséquence de la libéralisation des marchés comme pour Taiwan au début de l'année 1991, ou les Philipines à la fin de l'année 1989. Lorsque ils sont ignorés, leur test statistique conclu en faveur de la marche aléatoire. Au contraire, quand ils sont pris en compte, 10 des 17 séries sont mieux caractérisées par un processus de retour à la moyenne que par une marche aléatoire.

A présent, afin de confirmer de manière visuelle l'hypothèse de l'absence de mémoire dans les séries des rendements, nous traçons la fonction d'autocorrélation des rendements journaliers de l'action Cisco Systems pour laquelle nous disposons des prix à 1 minute sur une période de 10 ans à horizon fin 2003 afin d'observer sa décroissance (cf. figure 1.2). Celle-ci nous renseignera sur la présence ou l'absence de mémoire. A titre de comparaison, nous avons superposé la fonction d'autocorrélation de la valeur absolue des rendements.

Nous observons que l'autocorrélation des rendements chute brutalement dès le premier lag : l'autocorrélation devient nulle au au bout d'un jour et est maintenue à ce niveau sur toute la période considérée, ce qui traduit l'absence de mémoire. Quant à la fonction d'autocorrélation de la valeur absolue de rendements journaliers, son comportement n'est pas le

11. L'indépendance en statistique signifie que la distribution de probabilité du prix pendant la période t est indépendant de la séquence des prix pendant les périodes précédentes

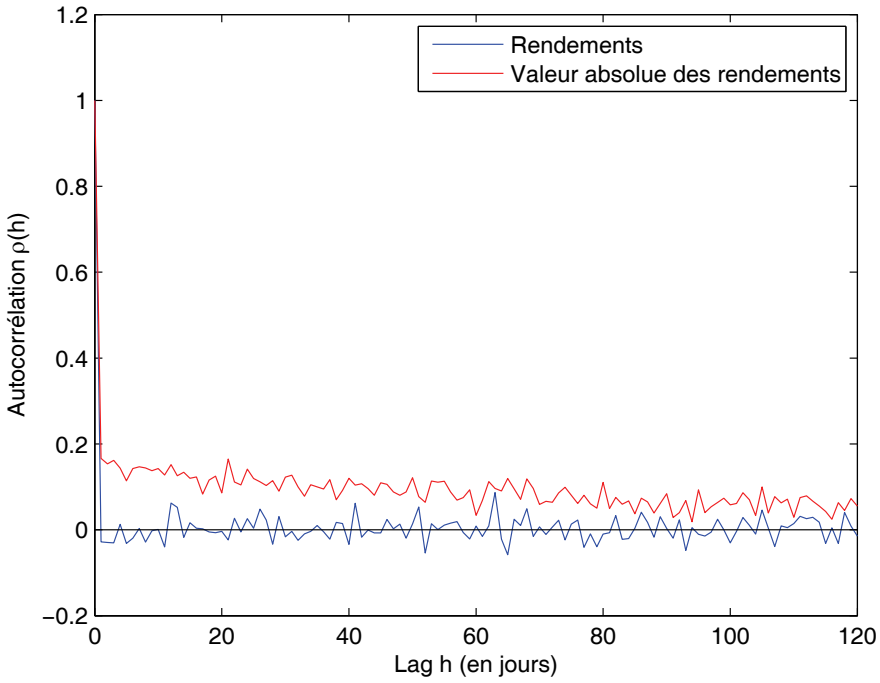

FIGURE 1.2 – Fonction d'autocorrélation des rendements journaliers et de la valeur absolue des rendements journaliers de l'action Cisco Systems sur une période de 120 jours.

même. Certes, la chute au premier lag est brutale, l'autocorrélation passant de 1 à environ 0.18, mais par la suite la décroissance devient très lente : après 120 jours, elle est encore de 0.08. Cela montre clairement la présence de mémoire longue.

La question de dépendance se pose également à des échelles de temps très courtes (pour le lien avec avec la microstructure des marché). Nous avons donc tracé le même graphique à partir cette fois-ci les rendements à la minute au lieu des rendements journaliers (voir figure 1.3. Nous constatons que l'autocorrélation des rendements chute brutalement dès le premier lag, à la différence qu'au lag 1 elle devient négative (-0.23). Puis, elle tend vers zéro au lag 2. Cette anti-corrélation au premier lag provient, selon Roll (1984) de l'écart, engendré par les frais de transaction, entre le prix d'achat et le prix de vente d'une action. Par la suite, la valeur de l'autocorrélation des rendements varie très faiblement autour de zéro, ce qui traduit l'absence de mémoire. En ce qui concerne l'autocorrélation de la série de la valeur absolue des rendements 1 minutes, ici aussi sa décroissance est très lente après la chute brute du premier lag (elle vaut 0.12 au bout de 120 minutes).

Revenons un instant au graphique 1.2. Il illustre le fait qu'une fonction non-linéaire des rendements affiche une persistance positive et significative. Or, puisque la volatilité en est une, la présence de mémoire longue dans les séries de la valeur absolue des rendements plaide en faveur de la présence de mémoire longue dans les séries de la volatilité d'actions. C'est ce à quoi nous allons maintenant nous intéresser.

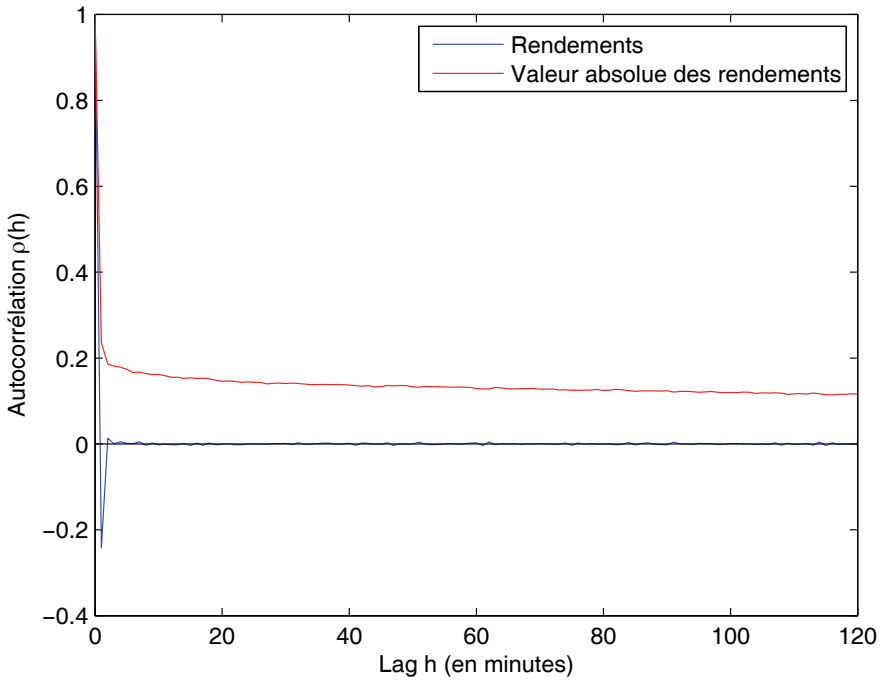

FIGURE 1.3 – Fonction d'autocorrélation des rendements 1 minute et de la valeur absolue des rendements 1 minutes de l'action Cisco Systems sur une période de 120 minutes.

1.2 La volatilité des rendements du prix d'une action

Contrairement au prix des actions, la volatilité n'est pas directement observable sur le marché. C'est une mesure du risque des actifs financiers. L'estimer présente alors un intérêt majeur d'autant plus qu'elle est utilisée dans de nombreux domaines en finance, parmi lesquels, la gestion du risque, la gestion de portefeuille, l'évaluation et la couverture de produits dérivés. Prenons le cas, par exemple, de la gestion de portefeuille et de la méthode d'allocation optimale des ressources de Markovitz (aussi appelée méthode moyenne-variance) où le gestionnaire doit faire un compromis entre le risque des titres considérés, représenté par leur volatilité (soit la racine carrée de la variance), et leur rendement espéré, décrit par leur moyenne. En effet, il peut décider d'une part de maximiser les gains en se fixant un niveau de risque et d'autre part de minimiser le risque en fonction d'un objectif de gain.

Il est donc primordial de mesurer avec la meilleure précision possible la volatilité des rendements, c'est à dire en tenant compte de toutes les informations disponibles, à savoir les rendements des actions. C'est la raison pour laquelle ce sujet a passionné de nombreux chercheurs et a donné naissance à de nombreuses définitions et estimateurs plus ou moins pratiques et plus ou moins efficaces.

La volatilité historique, comme son nom l'indique est basée sur les rendements passés. Elle est obtenue grâce à des modèles tel que ARMA ou GARCH par exemple. Une définition

récente, la volatilité réalisée, basée sur des données haute-fréquence et que nous étudierons en détail a été fournie par Barndorff-Nielsen et Shephard (2002b). Une approche alternative, basée sur le marché des options a été fournie par Latané et Rendleman (1976) : ils ont inversé la formule d'évaluation d'options de Black et Scholes (1973) pour inférer une volatilité implicite.

Dans cette section, nous détaillerons les facteurs qui influencent la valeur de la volatilité. Ensuite, nous passerons en revue les principaux estimateurs et modèles dédiés à la représentation de la volatilité. Enfin, nous achèverons cette section en nous intéressant à un estimateur en particulier : celui de la volatilité réalisée.

1.2.1 Facteurs influençant la volatilité

Lorsque les chercheurs se sont intéressés à la dynamique de la volatilité, et en particulier aux facteurs influençant cette dynamique, ils se sont tournés vers l'environnement économique et ses différentes variables.

Les chocs

Certains évènements économiques sont susceptibles d'engendrer des pics de volatilité comme, par exemple, la Grande Dépression de 1929-1939 aux États-Unis que Schwert (1989) a étudiée. Il a remarqué que, pendant toute cette période, la volatilité des actions était en moyenne 2 à 3 fois supérieure à sa valeur moyenne en période normale. Cet exemple reste toutefois un cas particulier, car pour les autres crises (certes de moins grande envergure) qu'a connues l'Histoire, l'augmentation de la volatilité ne persistait pas durant toute la crise, mais au contraire s'atténuait assez rapidement après avoir connu des pics en début période. En général, quelques jours suffisent. Schwert (1989) n'apporte pas d'explication à ces observations. Peut être est-ce le délai nécessaire pour que l'ensemble des acteurs du marché prenne conscience de la crise et l'assimile. Quoiqu'il en soit, une crise financière agit sur la volatilité des rendements des actions, mais ce n'est pas le seul facteur. D'autres évènements, politiques ou encore naturels sont également associés à des pics de volatilité (White 1996). Ce type de choc porte le qualificatif "exogène" par opposition à "endogène". Les chocs endogènes, c'est-à-dire ceux qui n'ont apparemment aucune cause extérieure, peuvent aussi être responsables de changements de volatilité. Sornette, Malevergne, et Muzi (2003) ont montré qu'en cas de chocs endogènes, le retour de la volatilité vers sa moyenne est plus lent qu'en cas de chocs endogènes.

L'effet de levier

Schwert (1989) analyse les relations entre la volatilité des rendements d'actions et la volatilité de l'inflation, la croissance monétaire, le niveau de l'activité économique ou encore le levier financier. Pour réaliser son étude, Schwert utilise les rendements journaliers du por-

tefeuille composite du Dow Jones de février 1885 à décembre 1927 puis celui de Standard & Poor's de janvier 1928 à décembre 1987 pour calculer la volatilité mensuelle estimée par la somme des rendements journaliers au carré (après avoir soustrait le rendement journalier moyen dans le mois).

Il résulte de cette étude que la volatilité des rendements est reliée à la santé générale de l'économie. Schwert (1989) s'appuie sur le levier financier pour interpréter ce résultat. Selon lui, le prix d'une action est un indicateur important qui, avant et durant une phase de récession, va chuter (relativement au prix des obligations). Ainsi, le levier augmente pendant les récessions, ce qui augmente le risque de la firme et contribue à l'accroissement de la volatilité. Cette interprétation est à mettre en relation directe avec Black (1976) et Christie (1982) qui ont observé une relation entre volatilité et levier financier : lorsque le levier augmente, alors la volatilité augmente. Par ailleurs, il existe une corrélation négative entre les rendements passés et la volatilité présente.

La volatilité passée

La volatilité passée peut être une cause de volatilité future. En effet, des études ont montré que les graphes de volatilité présentaient à certains endroits des concentrations de pics de volatilité ou au contraire des volatilités faibles. En d'autres termes, une forte volatilité a tendance à être suivie d'une forte volatilité, et inversement, une volatilité faible a tendance à être suivie d'une volatilité faible (Cont 2001). Bref, la volatilité présente dépend de la volatilité passée. On retrouve ici le principe du modèle GARCH.

1.2.2 Estimateurs de la volatilité

De nombreux estimateurs de la volatilité journalière ont été développés par le passé. Ces estimateurs ont d'abord été très simples en prenant en compte deux valeurs : le prix d'ouverture et le prix de clôture de l'actif. Bollen et Inder (2002) font un rappel des principaux estimateurs de la volatilité journalière. Les rendements journaliers s'écrivent $r_t = \ln p_t - \ln p_{t-1}$, où p_t est le prix de clôture du jour t. r_t est supposé suivre une loi normale avec une volatilité non constante σ_t :

$$r_t \sim N(0, \sigma_t^2). \tag{1.1}$$

Les estimateurs vont chercher à estimer le paramètre de la volatilité journalière σ_t.

L'estimateur le plus simple est basé sur l'observation que l'espérance de la valeur absolue des rendements est donnée par $\mathrm{E}[|r_t|] = \sigma_t\sqrt{2/\pi}$. L'estimateur simple de la volatilité est alors défini par :

$$\sigma_{S,t} = \frac{|r_t|}{\sqrt{2/\pi}}. \tag{1.2}$$

Cet estimateur ne dépend que de la valeur du rendement journalier, soit deux valeurs. C'est évidemment trop peu pour espérer que cet estimateur soit efficient. Parkinson (1980) propose un autre estimateur de la volatilité journalière. Il prend en compte le prix le plus haut H_t et

le prix le plus bas L_t de la journée t :

$$\sigma^2_{PARK,t} = \frac{(\ln H_t - \ln L_t)^2}{4 \ln 2}. \tag{1.3}$$

A première vue, cet estimateur ne prend en compte que deux valeurs. Ce n'est pas tout à fait exact car pour déterminer les prix maximum et minimum quotidiens, il faut l'ensemble des prix de la journée. Cet estimateur est donc plus évolué qu'il n'y paraît. Garman et Klass (1980) prétendent que leur estimateur est plus efficient que l'estimateur $\sigma_{PARK,t}$. Il est défini par :

$$\sigma^2_{GK,t} = 0.5 \left(\ln(H_t/L_t)\right)^2 - 0.39 \left(\ln(p_t/p_{t-1})\right)^2. \tag{1.4}$$

Quoi qu'il en soit, son inconvénient, comme celui tous de les estimateurs précédemment cités, réside dans le fait que l'on ne peut pas espérer incorporer une grande quantité d'informations et donc traduire correctement ce qui se passe sur le marché si l'on n'utilise pas explicitement toutes les observations des rendements de la journée. A titre d'exemple, si l'on considère une fenêtre de cotation de 9h30 à 16h, cela fait 390 rendements minute par minute. En conclusion, si l'on veut une mesure du risque qui soit plus fidèle à la réalité et donc plus juste, il faut un estimateur prenant en compte explicitement un plus grand nombre de données.

Récemment, Andersen, Bollerslev, Diebold, et Ebens (2001) ont proposé une mesure de la volatilité pour palier les différentes insuffisances des estimateurs précédents. Leur estimateur prend en compte un grand nombre de données puisqu'il utilise les prix minute par minute pour estimer la volatilité journalière. Cet estimateur, en plus de l'avantage de prendre en compte une grande quantité d'information sur les rendements, possède la propriété d'être asymptotique gaussien.

1.2.3 Estimateur de la volatilité réalisée

Andersen, Bollerslev, Diebold, et Ebens (2001), en se fondant sur la théorie de la variation quadratique dans un environnement standard d'absence d'opportunité d'arbitrage et sans friction, ont proposé un cadre général pour l'utilisation de données haute-fréquence pour la mesure, la modélisation et la prédiction de la volatilité journalière des rendements d'actions. Plus précisément, comme l'ont également rappelé Barndorff-Nielsen et Shephard (2002a, 2002b), lorsque le processus de prix des actifs sous-jacents est une équation de diffusion, la variance réalisée (le carré de la volatilité réalisée) fourni un estimateur convergent de la variation quadratique des rendements, dans la limite de grands échantillons :

Soit $p(t)$ le processus du logarithme des prix d'un actif vérifiant l'équation de diffusion

$$dp(t) = \mu(t)dt + \sigma(t)dW(t), \tag{1.5}$$

où $W(t)$ est un mouvement brownien standard, $\mu(t)$ est de variation finie et $\sigma(t)$ est strictement positif et de carré intégrable [12]. Ainsi, $\mu(t)$ représente le rendement espéré instantané et $\sigma(t)$ la volatilité des rendements.

12. C'est à dire $\mathrm{E}\left[\int_0^t \sigma^2(s)ds\right] < \infty$.

Le rendement cumulé sur l'intervalle de temps $[t-k,t]$, où $0 < k \leq t$, s'écrit

$$r(t,k) = p(t) - p(t-k) = \int_{t-k}^{t} \mu(s)ds + \int_{t-k}^{t} \sigma(s)dW(s), \tag{1.6}$$

et sa variation quadratique $VQ(t,k)$ est donnée par

$$VQ(t,k) = \int_{t-k}^{t} \sigma^2(s)ds. \tag{1.7}$$

En absence d'effets de microstructure et d'erreur de mesure, la variation quadratique des rendements peut être arbitrairement approchée de manière satisfaisante par le processus du carré des rendements cumulés. Soit une partition $\left\{t-k+\frac{j}{n}, j=1,\cdots,n\cdot k\right\}$ de l'intervalle $[t-k,t]$. Alors, la variance réalisée est définie par

$$VR(t,k;n) := \sum_{j=1}^{n\cdot k} r\left(t-k+\frac{j}{n},\frac{1}{n}\right)^2. \tag{1.8}$$

La théorie sur les semi-martingales assure que la mesure de la variance réalisée converge en probabilité vers la variation quadratique des rendements lorsque que n tend vers l'infini :

$$VR(t,k;n) \underset{n\to\infty}{\longrightarrow} VQ(t,k). \tag{1.9}$$

A présent, simplifions ces notations et adaptons-les au cadre de l'estimation de la volatilité réalisée journalière. Désignons par $P_{i,t}$ la i-ème observation du prix[13] d'un actif pendant le jour de cotation t et par $r_{i,t} = \ln(P_{i,t}) - \ln(P_{i-1,t}), i = 1,\cdots,n_t$ le rendement de l'actif considéré sur la période $i-1$ à i. Ainsi, la variance réalisée, au jour t, est définie par

$$\hat{\sigma}^2_{ABDL,t} = \sum_{i=1}^{n_t} r^2_{i,t}, \tag{1.10}$$

avec n_t le nombre d'observations pendant ce jour.

Comme l'ont montré Barndorff-Nielsen et Shephard (2002b), les propriétés asymptotiques de cet estimateur sont telles que

$$\frac{\hat{\sigma}^2_{ABDL,t} - \int_{t-1}^{t} \sigma^2(s)\ ds}{\sqrt{\frac{2}{3}\sum_{i=1}^{n_t} r^4_{i,t}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \tag{1.11}$$

$$\frac{\ln \hat{\sigma}^2_{ABDL,t} - \ln \int_{t-1}^{t} \sigma^2(s)\ ds}{\sqrt{\frac{2}{3}\frac{\sum_{i=1}^{n_t} r^4_{i,t}}{\left[\sum_{i=1}^{n_t} r^2_{i,t}\right]^2}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{1.12}$$

13. Le premier prix de la journée de cotation t est $P_{0,t}$ et le dernier prix est $P_{n_t,t}$

Si l'estimateur (1.10) ne requiert pas l'homoscédasticité de la série temporelle des rendements d'actif pendant le jour t, on suppose toutefois que les rendements ne sont pas corrélés. Par conséquent, afin de rendre compte des effets de microstructure, qui engendrent des corrélations (Roll 1984, par exemple), il est souvent nécessaire soit de filtrer les séries brutes des rendements intraday pour traiter ces corrélations [14], soit de se concentrer sur une échelle de temps suffisamment grande – 5 minutes au lieu de 1 minute par exemple – pour atténuer les effets de microstructure et les corrélations induites. L'inconvénient immédiat de cette dernière approche est de diminuer le nombre n_t d'observations disponibles, ce qui peut avoir pour conséquence de réduire la précision des estimations.

Appliqué à l'estimation de la volatilité d'actions, cet estimateur peut poser un problème. En effet, dans l'estimation de la volatilité, il faut également tenir compte de la volatilité overnight, c'est-à-dire de la volatilité de l'action entre le prix d'ouverture du marché et le prix de clôture de la veille, qui n'est pas négligeable comparé à la volatilité journalière. En effet, les variations des rendements overnight sont beaucoup plus importantes que les variations des rendements à 1 minute. Afin de rendre rendre compte de ce phénomène, Martens (2002) (voir aussi Koopman, Jungbacker, et Hol (2005)) propose d'appliquer un facteur d'échelle

$$\hat{\sigma}_{M,t} = \frac{\hat{\sigma}_{oc}^2 + \hat{\sigma}_{co}^2}{\hat{\sigma}_{oc}^2} \sum_{i=1}^{n_t} r_{i,t}^2, \tag{1.13}$$

où les variances empiriques *open-to-close* $\hat{\sigma}_{oc}^2$ et *close-to-open* $\hat{\sigma}_{co}^2$ sont données par

$$\hat{\sigma}_{oc}^2 = \frac{1}{n_t} \sum_{k=1}^{N} (\log P_{n_k,k} - \log P_{0,k})^2, \tag{1.14}$$

$$\hat{\sigma}_{co}^2 = \frac{1}{n_t} \sum_{k=1}^{N} (\log P_{0,k} - \log P_{n_{k-1},k-1})^2, \tag{1.15}$$

où $\log P_{0,k} - \log P_{n_{k-1},k-1}$ représente de rendement *overnight* entre la journée $k-1$ et la journée k.

Il existe d'autres méthodes pour estimer la variance d'ouverture à fermeture, comme par exemple Hansen et Lunde (2002) qui calculent la moyenne des volatilités réalisées de l'équation (1.10). Il existe aussi d'autres facteurs d'échelle comme celui d'Areal et Taylor (2002) qui assignent différents poids au carré des rendements intraday [15].

Pour illustrer nos propos, nous présentons sur la figure 1.4, la volatilité réalisée de deux séries temporelles. Sur la fenêtre de gauche, nous pouvons observer la volatilité des rendements journaliers d'une société de moyenne capitalisation, Washington Post, sur une période de 10 ans allant du 01/01/1994 au 31/12/2003, tandis que, sur la fenêtre de droite, est tracée la volatilité réalisée d'une société de grande capitalisation, Coca-Cola, sur la même période. Étant donné que nos études sur la dépendance entre les rendements intraday des prix des actifs de notre base de données n'ont pas révélé la présence d'une corrélation significative au-delà de l'échelle de temps d'une minute – pour tous les actifs, la fonction d'autocorrélation des

14. Voir l'étude comparative de Bollen et Inder (2002) par exemple.

15. Le lecteur est invité à se référer à Hansen et Lunde (2005) pour une discussion plus détaillée sur le choix du facteur d'échelle.

rendements à la minute (et au-delà) tombe à zéro au premier lag – les effets microstructuraux du marché sont négligeables, ce qui nous permet d'estimer directement la volatilité réalisée avec les rendements bruts à la minute.

Sur la figure 1.5, nous avons tracé la fonction d'autocorrélation des volatilités réalisées précédentes (avec toujours à gauche Washington Post et à droite Coca-Cola) du lag 0 au lag 250 (soit du jour 0 au 250ème jour), ce qui correspond à une période d'environ 1 an. La décroissance lente de la fonction d'autocorrélation est la caractéristique de la mémoire longue. Toutefois, les deux graphiques semblent différents. En ce qui concerne le Washington Post, l'autocorrélation chute rapidement à 0.4 et ensuite décroît très lentement, alors que pour Coca-Cola, la décroissance est plus régulière et l'autocorrélation n'est plus significativement différente de zéro au-delà de 180 jours. Cette caractéristique générale pourrait signifier que la fonction d'autocorrélation de la volatilité des rendements des grandes capitalisations américaines présentent moins de mémoire longue que celle des moyennes capitalisations. Cette remarque, qui sera confirmée par la suite, est dans la lignée de l'observation rapportée par Di Matteo, Aste, et Dacorogna (2005) et selon laquelle plus un titre est mature, plus la décroissance de la fonction d'autocorrélation de sa volatilité est rapide.

1.2.4 Faits stylisés de la volatilité réalisée

De nombreux faits stylisés concernant la volatilité des rendements des actifs financiers ont été rapportés dans plusieurs études ces dernières années (Ding, Granger, et Engle 1993, Lo 1991, parmi d'autres). En particulier, tout le monde est maintenant d'accord sur le fait que (i) les rendements montrent, à n'importe quelle échelle de temps, un grand degré de variabilité qui est mis en lumière par la présence de pics irréguliers de volatilité et (ii) la volatilité affiche une autocorrélation positive sur une longue période, ce qui quantifie le fait que des évènements à forte (resp. faible) volatilité ont tendance à se regrouper dans le temps comme cela a déjà été reporté par Mandelbrot (1963) et Fama (1965) qui ont été les premiers à mettre en évidence que des grands changements dans le prix des actifs sont souvent suivis par d'autres grands changements, et les petits changements sont souvent suivis par d'autres petits changements. Ce comportement a aussi été mis en lumière par plusieurs autres études, telles que Baillie, Bollerslev, et Mikkelsen (1996), Chou (1988) ou Schwert (1989) par exemple. Ce phénomène est en relation directe avec la non normalité des rendements : ces agrégats de grandes valeurs de la volatilité générera des queues épaisses dans la distribution des rendements.

Dans cette section, nous décrivons les caractéristiques de la volatilité réalisée et les relions à la littérature.

Normalité de la log-volatilité réalisée

Nous nous tournons à présent vers les propriétés de la distribution de la volatilité réalisée. Comme il l'a été suggéré auparavant dans de nombreuses études, (Andersen, Bollerslev, Diebold, et Ebens 2001, Andersen, Bollerslev, Diebold, et Labys 2001, Barndorff-Nielsen et

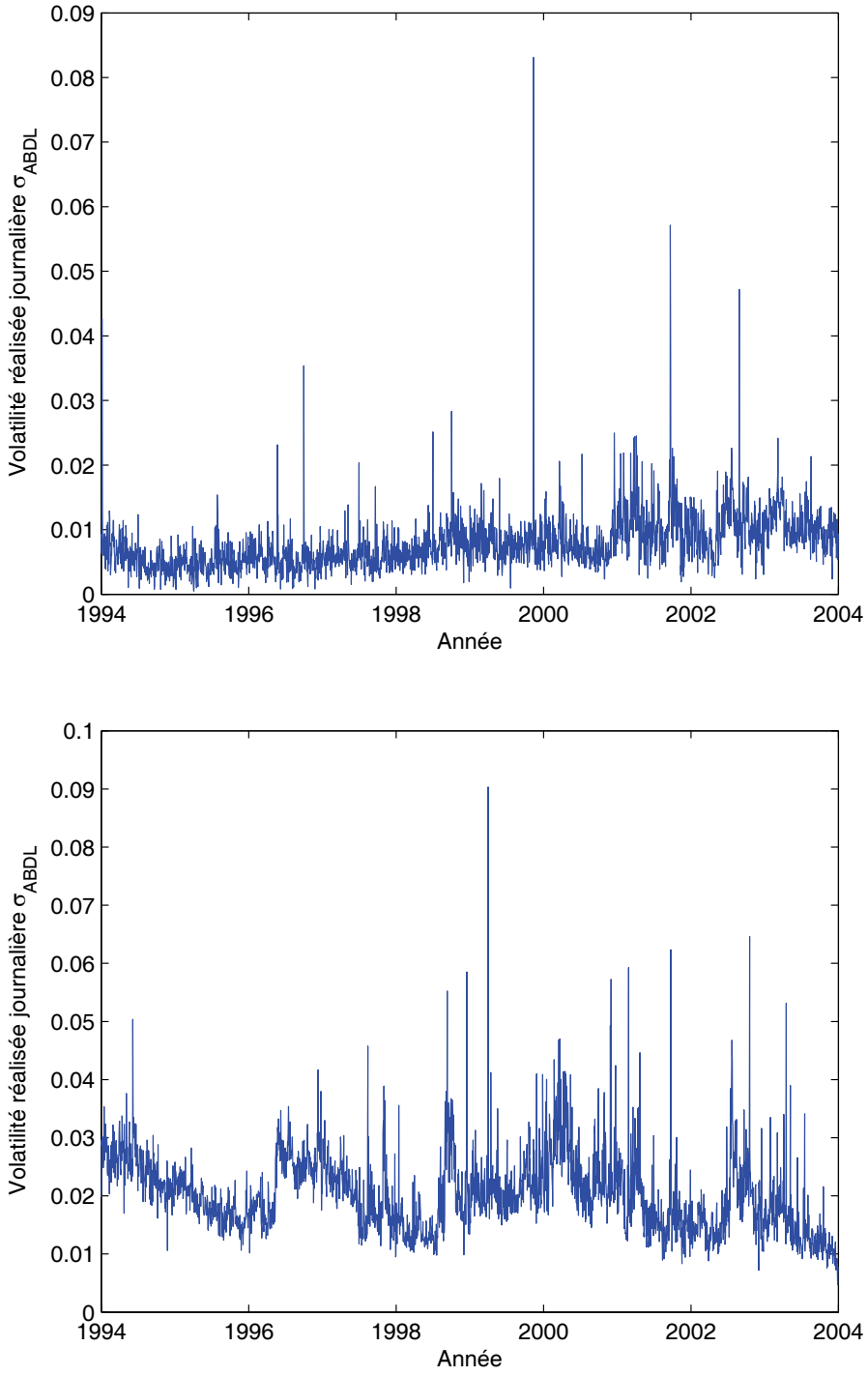

FIGURE 1.4 – Volatilité réalisée journalière de deux actions sur une période de 10 ans (du 01/01/1994 au 31/12/2003). En haut une moyenne capitalisation : Washington Post, et en bas une grande capitalisation : Coca Cola.

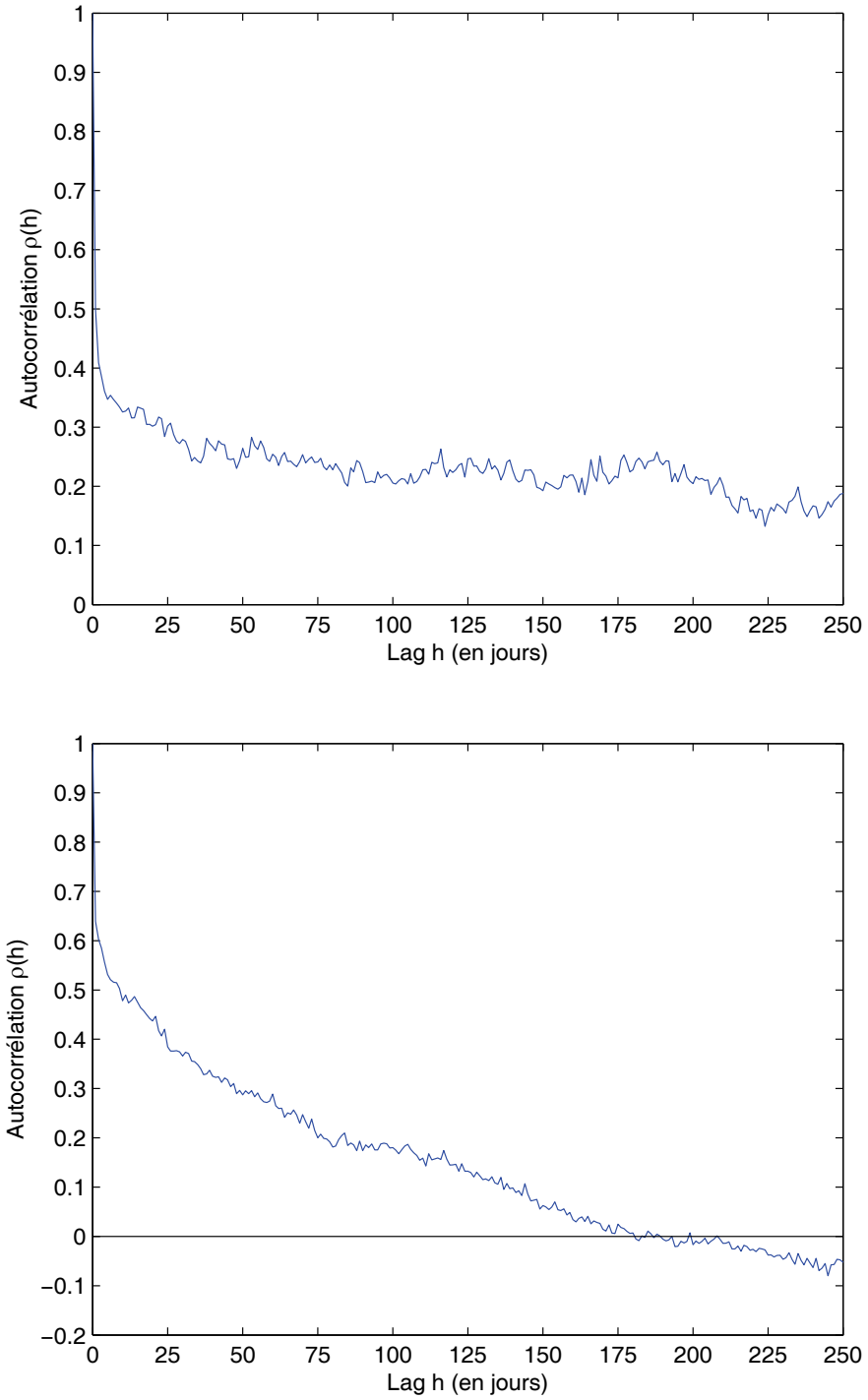

FIGURE 1.5 – Fonction d'autocorrélation de la volatilité réalisée de deux actions du 01/01/1994 au 31/12/2003. En haut une moyenne capitalisation : Washington Post, et en bas une grande capitalisation : Coca Cola.

Shephard 2002a, entre autres), la distribution de la loi log-normale [16] fournit une description adéquate de la volatilité, au moins au coeur de la distribution. Ces observations plaident clairement pour une modélisation de la volatilité réalisée au moyen de modèles log-normaux, qui renvoient à Clark (1973) (voir aussi Taylor (1986)).

En partant de ces observations, illustrons la pertinence d'un modèle log-normal de la volatilité réalisée des rendements des prix d'actions. Dans ce but, notons $\{\omega_t\}_{t\geq 1}$ le logarithme de la série de la volatilité réalisée. La figure 1.6 montre l'estimation par noyau (Pagan et Ullah 1999) de la densité de la log-volatilité d'une société de moyenne capitalisation, Microchip Technology, dans la fenêtre du haut, et d'une société de grande capitalisation, Procter & Gamble, dans celle du bas sur la même période de 10 ans (du 01/01/1994 au 31/12/2003). La densité de la log-volatilité des rendements des prix de Microchip Technology (courbe en trait plein) semblent très proche de la densité d'une loi normale (courbe en pointillés). De manière générale, c'est le cas pour toutes les moyennes capitalisations de notre échantillon. Au contraire, les densités des grandes capitalisations ont des pics pointus et des queues épaisses à tel point qu'elles s'éloignent de manière significative de la densité d'une loi normale. Ces impressions visuelles seront confirmées au moyen de tests statistiques dans la section 3.2.1.

Normalité des rendements journaliers normalisés par la volatilité réalisée

La densité des rendements des actions présentent un pic plus pointu qu'une densité de loi normale et des queues épaisses. Par conséquent, les rendements ne sont pas gaussiens. C'est ce qu'illustre la figure 1.7, où les densités des rendements de deux actions sont tracées avec les densités des lois normales les plus proches.

En revanche, les rendements journaliers normalisés par la volatilité réalisée journalière sont gaussiens. Pour les mêmes actions que celles considérées pour les rendements journaliers, nous avons tracé, sur la figure 1.8, les densités des rendements journaliers normalisés par la volatilité réalisée journalière ainsi que les densités des lois normales les plus proches. La similitude entre les densités des rendements et les densités de loi normale ne fait aucun doute.

Ce résultat n'est pas surprenant lorsque l'on considère les travaux de Clark (1973) et Ané et Geman (2000). Clark (1973) avait relié la normalité des rendements au volume des échanges, avant qu'Ané et Geman (2000) ne démontrent que c'était le nombre des échanges qui intervenait et constituait le temps des affaires (par opposition au temps calendaire). En divisant les rendements journaliers par la volatilité réalisée journalière, celle-ci est considérée comme étant l'horloge des affaires.

16. A noter que Barndorff-Nielsen et Shephard (2002a) montrent aussi que la loi gaussienne inverse fournit une approximation précise de la distribution de la log-volatilité. En fait, les distributions de la loi log-normale et de la gaussienne inverse sont proches à tel point que l'on ne peut pas les distinguer sur tout l'axe d'intérêt.

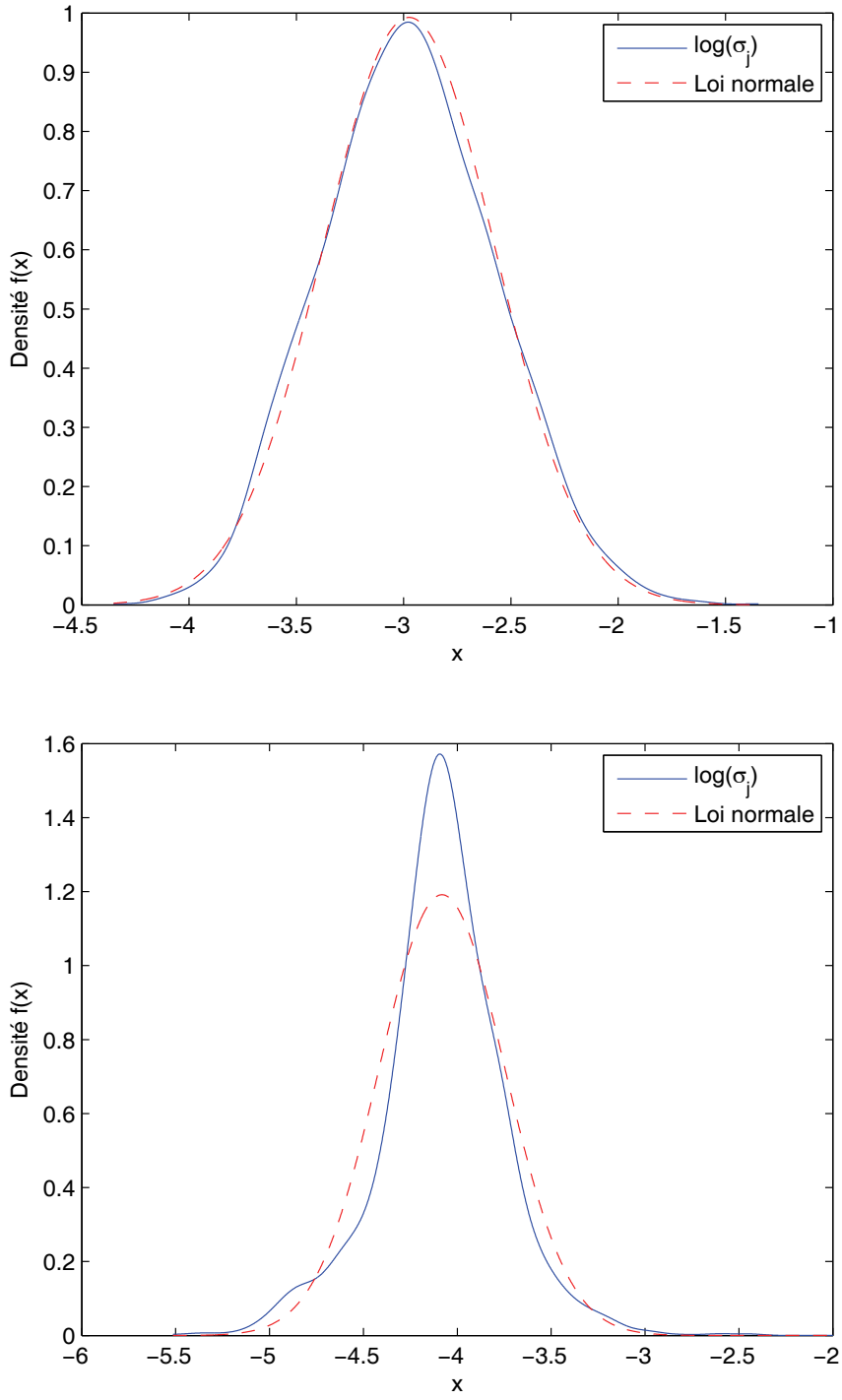

FIGURE 1.6 – Densité de $\log(\hat{\sigma}_t)$ et de sa loi normale approchée pour différents actifs. En haut une moyenne capitalisation : Microchip Technology, et en bas une grande capitalisation : Procter & Gamble.

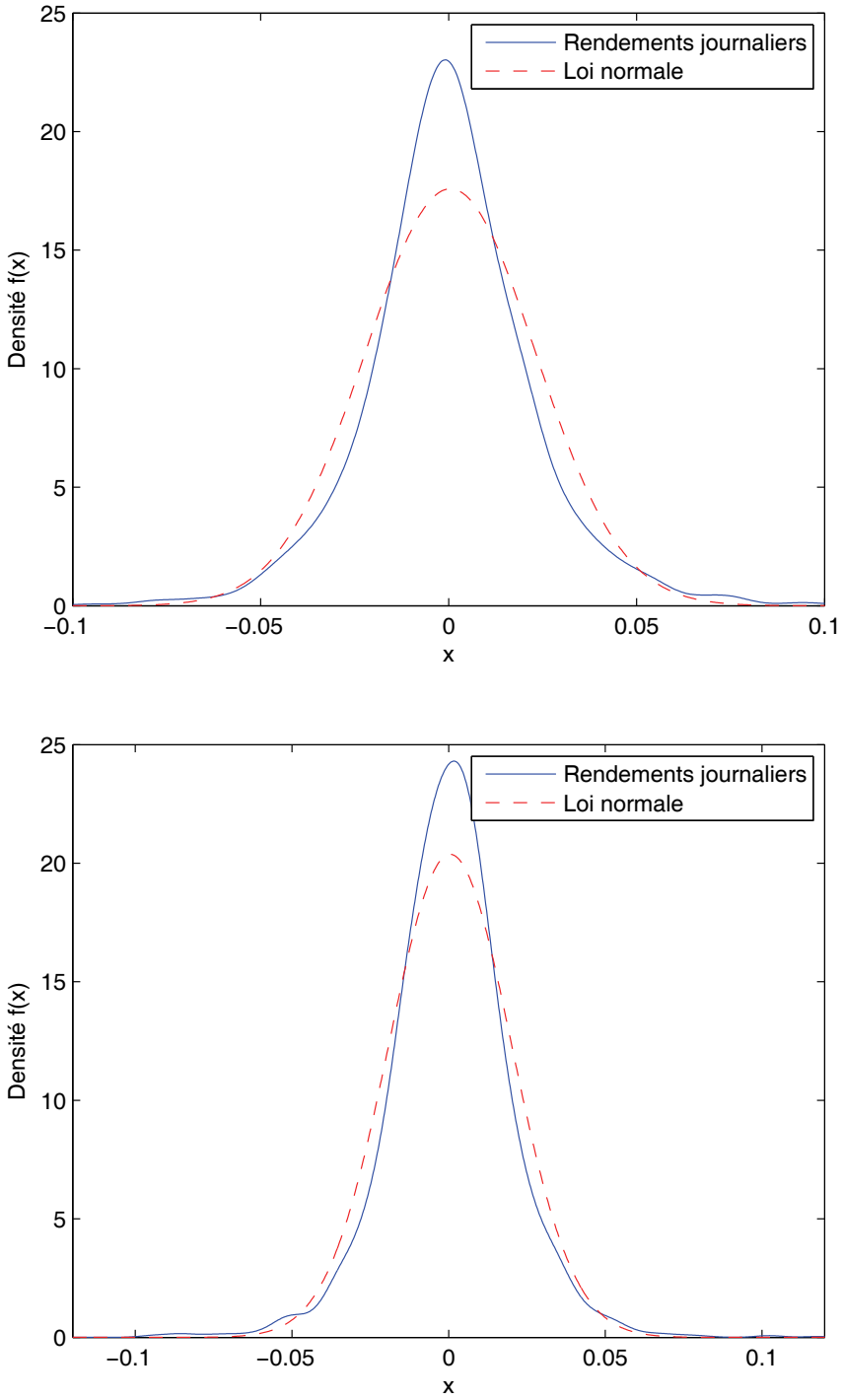

FIGURE 1.7 – Densité des rendements journaliers et de sa loi normale approchée pour deux actions : en haut une moyenne capitalisation, Valero, et en bas une grande capitalisation, Genéral Electrics.

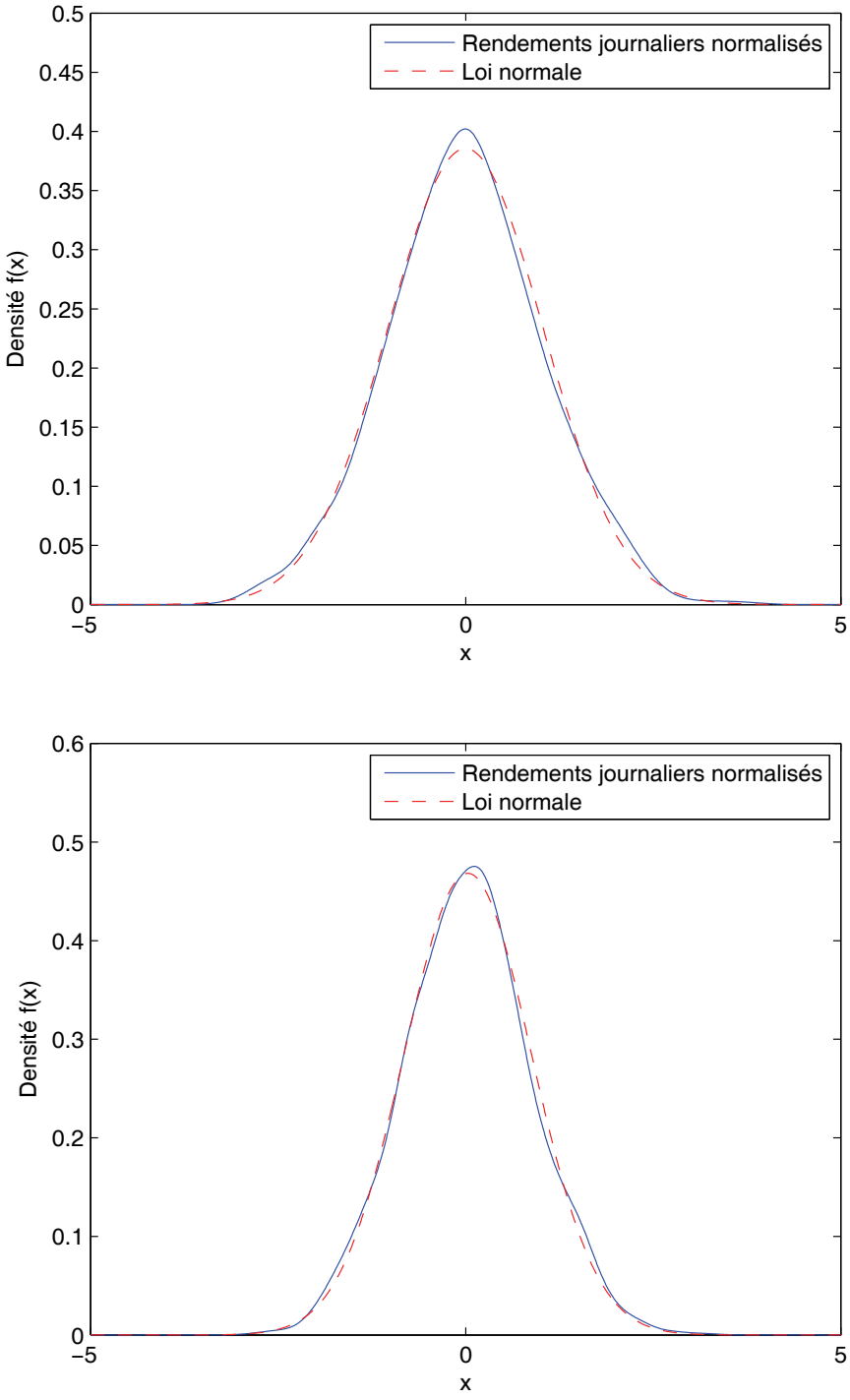

FIGURE 1.8 – Densité des rendements journaliers normalisés par la volatilité réalisée journalière et densité de sa loi normale approchée pour deux actions : en haut une moyenne capitalisation, Valero, et en bas une grande capitalisation, Genéral Electrics.

1.3 La mémoire longue

Nous avons observé dans la section précédente grâce à la figure 1.6 que la décroissance de la fonction d'autocorrélation des rendements des prix des actions était lente, on parle alors de décroissance hyperbolique. Ce trait est caractéristique de la présence de mémoire longue dans la série temporelle de la volatilité, par opposition à une décroissance rapide de type exponentiel, synonyme de mémoire courte. En d'autres termes, cela signifie qu'en présence de mémoire longue, la volatilité des rendements d'un actif le jour j peut avoir un impact sur la volatilité 6 mois après (voir figure 1.5), alors qu'en présence de mémoire courte, l'impact disparaît totalement le plus souvent après au plus une dizaine de jours.

La prise en compte du phénomène de mémoire longue est un aspect capital de la modélisation de la volatilité. En effet, elle va contribuer à une meilleure modélisation, et par conséquent une meilleure prédiction de la volatilité, à long terme. En outre, plus la modélisation sera précise sur le long terme et plus l'allocation des ressources du portefeuille sera performante. Le gestionnaire pourra avoir une idée plus précise du risque à moyen ou long terme et ainsi gérer son portefeuille à plus longue échéance. Il fera alors l'économie des frais de transaction qui grèvent la performance de son portefeuille.

C'est pourquoi l'analyse des séries temporelles financières en général, et de la volatilité des rendements en particulier, a suscité l'intérêt de nombreux chercheurs et est à l'origine d'une littérature très riche à ce sujet. Dans ce domaine, les pionniers Robert F. Engle et Clive W.J. Granger, se sont vus décerner le prix Nobel d'économie le 8 octobre 2003 par l'Académie Royale des Sciences de Suède pour leurs travaux respectivement sur "les méthodes d'analyse des séries temporelles économiques avec volatilité saisonnière" et sur "les méthodes d'analyse des séries temporelles économiques avec tendance commune". Plus précisément, Engle a été récompensé pour ses travaux relatifs aux modèles ARCH (AutoRegressive Conditional Heteroscedasticity) et Granger pour ses travaux relatifs à analyse des séries non-stationnaires au travers du concept de cointégration. Granger a été le premier à utiliser l'analyse spectrale et les modèles non-linéaires dans les séries temporelles économiques et financières.

Dans cette section, nous commençons par aborder brièvement les deux définitions mathématiques de la mémoire longue. Puis, nous présentons l'évolution des modèles appliqués à la représentation de la volatilité. Ensuite, nous cherchons à connaître la signification de la présence de mémoire longue dans les séries de la volatilité. En effet, puisque dans notre modèle nous cherchons à relier ses paramètres à la réalité sous-jacente du marché, nous nous intéressons aux travaux précédemment parus traitant de ce sujet. Enfin, nous détaillons les hypothèses permettant de rendre compte de l'origine de la mémoire longue dans certaines séries temporelles financières et économiques. C'est la première étape à réaliser pour modéliser la volatilité des rendements.

1.3.1 Définitions

Le concept de mémoire longue ne vient pas de la littérature économique, mais d'autres disciplines telles que l'hydrologie, la géophysique, la climatologie et autres sciences naturelles.

Voir Hurst (1951) et Hurst (1957), Mandelbrot et Wallis (1968), Mandelbrot (1972), McLeod et Hipel (1978). Hurst a d'ailleurs donné son nom à un coefficient mesurant la mémoire longue, l'exposant de Hurst.

Il existe plusieurs définitions de la mémoire longue, que ce soit dans le domaine temporel ou dans le domaine fréquenciel. Dans le domaine temporel, la mémoire longue se traduit par une décroissance hyperbolique de la fonction d'autocorrélation. Dans le domaine fréquenciel, la même information est contenue sous la forme spectrale, ce qui a l'avantage selon Banerjee et Urga (2005) de concentrer toute l'information dans l'intervalle $[0, \pi]$. Plusieurs définitions de la mémoire longue sont à notre disposition, toutefois, nous nous contenterons de rappeler les définitions classiques (Banerjee et Urga 2005).

Définition 1 (Mémoire longue dans le domaine temporel)**.** *Dans le domaine temporel, un processus discret et stationnaire* $\{X_t\}_{t\in\mathbb{Z}}$ *est dit à mémoire longue s'il existe un réel* $\alpha \in]0,1[$ *ainsi qu'une constante* $c_\rho > 0$ *tels que*

$$\lim_{k\to+\infty} k^\alpha \rho(k) = c_\rho, \tag{1.16}$$

où $\rho(k)$ *est une fonction d'autocorrélation.*

Cette définition signifie que la corrélation entre deux observations successives décroît lentement lorsque le nombre de lags tend vers l'infini. McLeod et Hipel (1978) proposent une définition plus générale :

Définition 2 (Mémoire longue dans le domaine temporel)**.** *Dans le domaine temporel, un processus discret et stationnaire* $\{X_t\}_{t\in\mathbb{Z}}$ *est dit à mémoire longue si*

$$\lim_{n\to\infty} \sum_{k=-n}^{n} |\rho(k)| = \infty. \tag{1.17}$$

Dans le domaine spectral, commençons par rappeler la définition de la densité spectrale [17] :

Définition 3. *La densité spectrale d'un processus stationnaire* $\{X_t\}_{t\in\mathbb{Z}}$ *vérifiant*

$$\lim_{n\to\infty} \sum_{h=-n}^{n} |\gamma(h)| < \infty, \tag{1.18}$$

où $\gamma(\cdot)$ *est sa fonction d'autocovariance, est donnée par*

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}\lambda h}\gamma(h). \tag{1.19}$$

Définition 4 (Mémoire longue dans le domaine fréquenciel)**.** *Dans le domaine spectral, un processus discret et stationnaire* $\{X_t\}_{t\in\mathbb{Z}}$ *est dit à mémoire longue s'il existe un réel* $\beta \in]0,1[$ *et une constante* $c_f > 0$ *tels que*

$$\lim_{\lambda\to 0} |\lambda|^\beta f(\lambda) = c_f, \tag{1.20}$$

où $f(\lambda)$ *est une fonction de densité spectrale.*

17. Pour plus de détails, le lecteur est invité à se référer à Brockwell et Davis (1990).

Cette définition implique que la densité spectrale n'est pas bornée pour de basses fréquences et diverge comme $\underset{\lambda \to 0}{\sim} \lambda^{-\beta}$.

Banerjee et Urga (2005) rappellent que les définitions 1 et 4 ne sont pas équivalentes mais que Beran (1994) a montré qu'il existe une connexion entre elles. En effet, si $0 < d < 1/2$ alors $\alpha = 1 - \beta$ pour autant que $\alpha = 1 - 2d$.

1.3.2 Modèles appliqués à la représentation de la volatilité

Par le passé, des modèles appliqués à la représentation de la volatilité ont été développés parallèlement aux estimateurs évoqués dans la section 1.2.2. Les premiers modèles étaient assez simples et ne prenaient pas en compte la mémoire longue. Par la suite des chercheurs leur ont apporté diverses améliorations pour palier leurs inconvénients, et notamment la prise en compte de la mémoire longue. Dans cette section, nous allons rappeler les définitions de deux célèbres modèles, le modèle ARMA de Box et Jenkins (1970) et le modèle ARCH de Engle (1982), ainsi leurs développements les plus intéressants parmi lesquels les modèles fractionnairement intégrés capables de prendre en compte la mémoire longue.

Modèles à mémoire courte

Les années 70 ont vu apparaître la classe des modèles ARMA (AutoRegressive Moving Average) de Box et Jenkins, combinaison des modèles autorégressif (AR) et moyenne mobile (MA).

Définition 5. *(Processus ARMA(p,q))*
Un processus $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ *satisfait une représentation ARMA(p,q) si* $\{X_t\}$ *est stationnaire et si pour tout* t*,*

$$b(L)X_t = a(L)\varepsilon_t, \tag{1.21}$$

où $a(L) = 1 + a_1 L + a_2 L^2 + \cdots + a_q L^q$ *et* $b(L) = 1 + b_1 L + b_2 L^2 + \cdots + b_p L^p$*, avec* L *l'opérateur de décalage* ($LX_t = X_{t-1}$) *et où* ε_t *est un bruit blanc de variance finie.*

Parallèlement au modèle ARMA, Engle (1982) a introduit les modèles ARCH (AutoRegressive Conditionaly Heteroscedastic) qui permettent de prendre en compte les regroupements de volatilité, c'est-à-dire le fait que les fortes (respectivement faibles) variations de prix soient suivies par d'autres fortes (faibles) variations de prix mais dont le signe n'est pas prévisible.

Définition 6. *(Processus ARCH(q))*
Un processus X_t *satisfait une représentation ARCH(q) si :*

$$X_t = z_t h_t, \tag{1.22}$$

avec

$$h_t^2 = \alpha_0 + \alpha(L)X_t^2, \tag{1.23}$$

avec $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_q L^q$ *où* L *est l'opérateur de décalage et* $\alpha_i \geq 0$ *pour tout* $i = 1, \cdots, q$. z_t *est indépendant de* h_t *et désigne un bruit blanc faible tel que* $E[z_t] = 0$ *et* $E[z_t^2] = \sigma_z^2$.

De nombreux modèles sont dérivés du modèle ARCH. Chacun de ces modèles tentent de corriger ou de prendre en compte un aspect des séries financières négligé par le modèle ARCH. Citons notamment Bollerslev (1986) qui a introduit les valeurs retardées de la variance dans son équation en définissant le modèle GARCH (Generalized ARCH).

Définition 7. *(Processus GARCH(p,q))*
Un processus X_t *satisfait une représentation GARCH(p,q) si :*

$$X_t = z_t h_t, \tag{1.24}$$

avec

$$h_t^2 = \alpha_0 + \alpha(L) X_t^2 + \beta(L) h_t^2, \tag{1.25}$$

avec $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_q L^q$, $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$ *où* L *est l'opérateur de décalage et où* z_t *désigne un bruit blanc faible,* $\alpha_0 > 0$, $\alpha_i \geq 0$ *pour tout* $i = 1, \ldots, q$ *et* $\beta_j \geq 0$ *pour tout* $j = 1, \ldots, p$. *Pour des raisons de stabilité et de stationnarité, toutes les racines de* $(1 - \alpha(L) - \beta(L))$ *et de* $(1 - \beta(L))$ *doivent être en dehors du cercle unité.*

Le modèle GARCH a connu un très grand succès aussi bien dans le monde académique que professionnel du fait de sa simplicité de mise en œuvre et de sa capacité à rendre compte de la majorité des caractéristiques des séries financières.

Nous pouvons également citer Nelson et son modèle EGARCH (Exponential GARCH) qui prend en compte l'effet de levier.

Définition 8. *(Processus EGARCH(p,q))*
Un processus X_t *satisfait une représentation EGARCH(p,q) si :*

$$X_t = z_t h_t, \tag{1.26}$$

avec

$$\log(h_t^2) = \alpha_0 + \alpha(L) g(z_t) + \beta(L) \log(h_t^2), \tag{1.27}$$

où $\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \cdots + \alpha_q L^q$, $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$ *avec* L *l'opérateur de décalage, et où le résidu normalisé* z_t *désigne un bruit blanc faible et la fonction* $g(.)$ *vérifie :*

$$g(z_{t-i}) = \theta z_{t-i} + \gamma(|z_{t-i}| - E[|z_{t-i}|]). \tag{1.28}$$

Modèles à mémoire longue

Les processus intégrés sont à l'origine des processus fractionnairement intégrés capables de traduire la mémoire longue. Le premier processus intégré est le processus ARIMA développé à partir modèle ARMA. Un paramètre $d \in \mathrm{N}$ est introduit de la manière suivante :

Définition 9. *(Processus ARIMA(p,d,q))*
Si d est un entier non négatif, alors $\{X_t\}$ *est un processus ARIMA(p,d,q) si* $Y_t = (1-L)^d X_t$ *est un processus ARMA(p,q) causal. Autrement dit,* $\{X_t\}$ *vérifie*

$$b(L)(1-L)^d X_t = a(L)\varepsilon_t, \tag{1.29}$$

où $\{\varepsilon_t\}$ *est un bruit blanc et* $a(.)$ *et* $b(.)$ *des polynômes de degré p et q respectivement tel que* $b(Z) \neq 0$ *pour* $|Z| < 1$.

Le processus est intégré si $d = 1$.

Le point de départ des processus fractionnairement intégrés en finance a été de remarquer que les séries temporelles financières telles que la volatilité n'étaient ni I(0), ni I(1) (voir Granger (1980)). En effet, la fonction d'autocorrélation de ces séries décroît lentement de façon hyperbolique (voir figure 1.9). C'est l'illustration du phénomène de "mémoire longue", aussi appelé "persistance". Or, les processus du type ARMA ou GARCH ayant une fonction d'autocorrélation qui décroît de manière exponentielle ou géométrique, c'est bien la preuve que ces processus, dits à mémoire courte, sont inadaptés pour modéliser des séries temporelles financières (voir figure 1.10). Ding, Granger, et Engle (1993), Harvey (1993) et Baillie (1996) montrent que le phénomène de mémoire longue est mieux décrit par des processus fractionnairement intégrés.

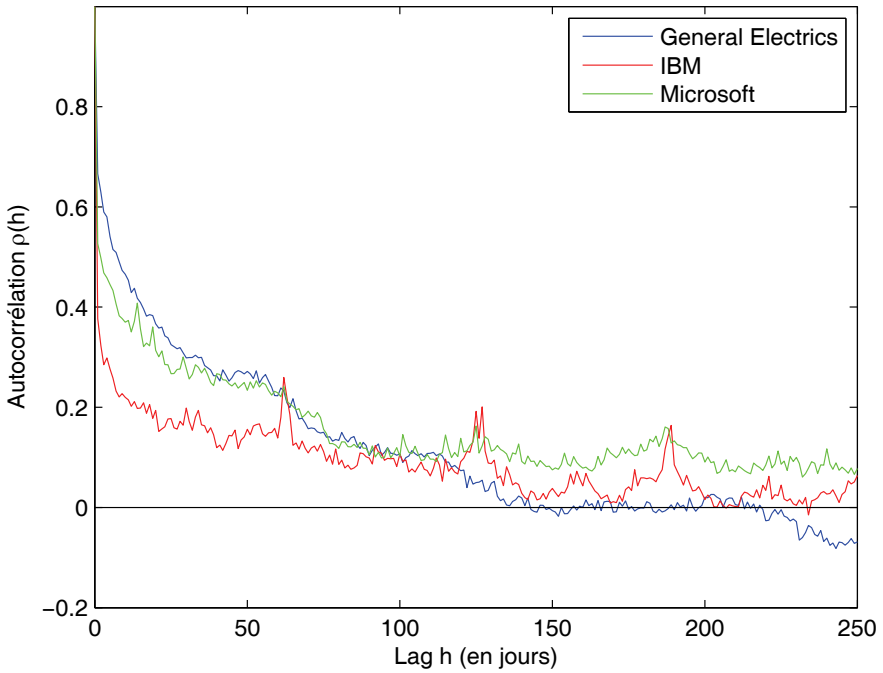

FIGURE 1.9 – Fonctions d'autocorrélation du logarithme de la volatilité réalisée de trois actions américaines à forte capitalisation.

Il existe plusieurs modèles prenant en compte la mémoire longue des séries financières.

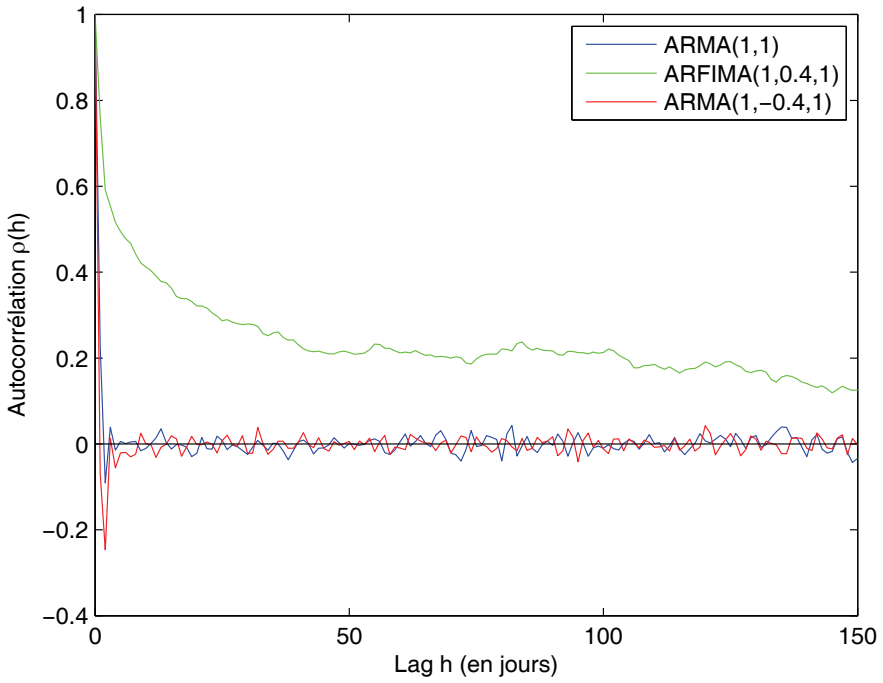

FIGURE 1.10 – Comparaison des fonctions d'autocorrélation empiriques d'un processus ARMA(1,1) vérifiant $(1-0.3L)X_t = (1-0.6L)\varepsilon_t$ et de deux processus ARFIMA(1,d,1) vérifiant $(1-0.3L)(1-L)^d X_t = (1-0.6L)\varepsilon_t$ pour $d = 0.4$ et $d = -0.4$, à horizon de 150 jours.

Deux de ces modèles sont nettement plus connus que les autres. Il s'agit du modèle ARFIMA (AutoRegressive Fractionally Integrated Moving Average) de Granger et Joyeux (1980) (dérivé du modèle ARIMA) et du modèle FIGARCH (Fractionally Integrated GARCH) de Baillie, Bollerslev, et Mikkelsen (1996) dérivé du modèle GARCH de Bollerslev (1986).

Définition 10. *(Processus ARFIMA(p,d,q))*
Si $d \in (-0.5, 0.5)$, alors $\{X_t\}$ satisfait une représentation ARFIMA(p,d,q) si $Y_t = (1-L)^d X_t$ est un processus ARMA(p,q) causal.

La définition implique que $\{X_t\}$ satisfait l'équation

$$b(L)(1-L)^d X_t = a(L)\varepsilon_t, \tag{1.30}$$

où $\{\varepsilon_t\}$ est un bruit blanc et $a(.)$ et $b(.)$ des polynômes de degré p et q respectivement tel que $b(Z) \neq 0$ pour $|Z| < 1$. De plus, $(1-L)^d$ est un opérateur définit par :

$$(1-L)^d = \sum_{k=0}^{+\infty} \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} L^k, \tag{1.31}$$

où Γ est la fonction gamma usuelle.

L'intégrant fractionnaire d représente une mesure de l'intensité de la dépendance à long terme (voir figure 1.11). De plus, la stationnarité au second ordre et l'inversibilité du processus exigent que $|d| < 0.5$. Plus particulièrement,

- si $0 < d < 0.5$, alors le processus est stationnaire et à mémoire longue,
- si $d = 0$, alors le processus est à mémoire courte.
- si $-0.5 < d < 0$, alors le processus présente une mémoire longue négative. On parle d'anti-persistance.

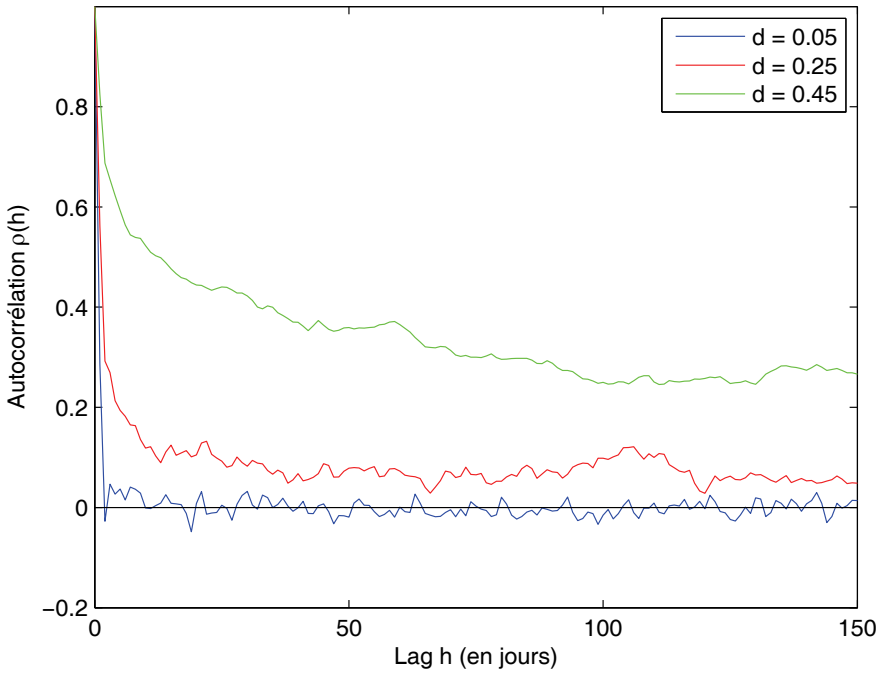

FIGURE 1.11 – Fonctions d'autocorrélation empiriques d'un processus ARFIMA(1,d,1) vérifiant l'équation$(1 - 0.3L)(1 - L)^d X_t = (1 - 0.6L)\varepsilon_t$ pour différentes valeurs de d : $d = 0.0.5$, $d = 0.25$ et $d = 0.45$, à horizon de 150 jours.

Bien qu'un nombre important d'articles concluent qu'un modèle à mémoire longue est plus efficace pour caractériser une série temporelle financière qu'un modèle à mémoire courte, le modèle ARFIMA(p,d,q) souffre de certaines critiques. En effet, selon Man (2003) et de Peretti (2003), une grande variance dans l'estimation des paramètres la rend automatiquement peu efficace.

Un autre modèle a évolué et a fini par prendre en compte le paramètre de mémoire longue : le GARCH. D'autres extensions sont d'abord apparues pour pallier les insuffisances du modèle GARCH tel que les modèles EGARCH (Exponential GARCH) qui généralise l'impact des nouvelles en incluant une relation de dépendance entre la variance et le module et le signe des chocs, TGARCH (Threshold GARCH) et QGARCH (Quadratic GARCH) dont l'objectif est de tenir compte des phénomènes d'asymétrie.

Hélas, la structure de ces modèles impose que la fonction d'autocorrélation décroisse de façon exponentielle. Baillie, Bollerslev, et Mikkelsen (1996) a incorporé le paramètre de mémoire longue dans le modèle GARCH.

Définition 11. *(Processus FIGARCH(p,d,q))*
Un processus X_t satisfait une représentation FIGARCH(p,d,q) si

$$\Phi(L)(1-L)^d X_t^2 = \alpha_0 + (1-\beta(L))v_t, \tag{1.32}$$

où $\Phi(L) = (1-\alpha(L)-\beta(L))(1-L)^{-1}$ et $v_t = \varepsilon_t^2 - \sigma_t^2$. Les racines de $\Phi(L)$ et de $(1-\beta(L))$ doivent être en dehors du cercle unité.

1.3.3 Origine de la mémoire longue dans les séries temporelles financières et économiques

Hormis la mauvaise spécification du modèle, deux facteurs sont à l'origine de la mémoire longue des séries temporelles économiques et financières telles que les séries de volatilités : le premier, mis en évidence par Granger (1980), concerne la nature des séries temporelles financières. En effet, il estime que celles-ci sont les résultantes de l'agrégation de micro-variables, et que cette agrégation est responsable de la mémoire longue. Le second facteur est la présence dans la série temporelle étudiée de changements structurels.

L'agrégation

Nous rappelons les arguments de Granger (1980) en adaptant ses notations à celles de la thèse pour plus de clarté.

Soit X_t une série temporelle centrée, générée par un bruit blanc centré de variance σ, noté ε_t en utilisant le filtre linéaire $a(L)$ où L est l'opérateur de décalage tel que

$$X_t = a(L)\varepsilon_t, \qquad L^k\varepsilon_t = \varepsilon_{t-k}, \tag{1.33}$$

avec $a(L)$, appelé filtre intégrant d'ordre d, pouvant s'écrire sous la forme :

$$a(L) = (1-L)^{-d}a'(L), \tag{1.34}$$

où $a'(z)$ n'a ni pôle ni racine au point $z = 0$. Le processus X_t est alors appelé processus intégré d'ordre d et noté : $I(d)$. Définissons à présent

$$X'_t = (1-L)^d X_t = a'(L)\varepsilon_t. \tag{1.35}$$

Alors $X'_t \sim I(0)$ à cause des propriétés de $a'(L)$. De plus, si $a'(L)$ est le quotient de deux polynômes finis en L d'ordre l et m, et si d est entier, alors X_t est un processus ARIMA(l,d,m) défini par Box et Jenkins (1970). Dans les modèles plus généraux considérés par Granger (1980), d n'a plus besoin d'être entier.

Considérons maintenant deux séries X_{1t} et X_{2t} générées par :

$$X_{jt} = b_j X_{j,t-1} + \varepsilon_{jt}, \qquad j = 1, 2, \tag{1.36}$$

où ε_{1t} et ε_{2t} sont deux bruits blancs centrés. Granger et Morris (1976) et Granger et Newbold (1977) ont montré que la somme $\bar{X}_t = X_{1t} + X_{2t}$ de ces 2 processus $AR(1)$ indépendants est régie par un processus $ARMA(2,1)$ dont la partie autorégressive est $(1 - b_1 L)(1 - b_2 L)$.

Ce résultat peut être généralisé à la somme de N processus $AR(1)$ indépendants du type de l'équation (1.36) avec des b_j différents, qui sera équivalente à un processus $ARMA(N, N-1)$ à condition que les racines des polynômes autorégressif et moyenne mobile n'aient pas de racines communes.

Enfin, Granger (1980), en considérant le spectre de la série $\bar{X}_t$ résultant de l'agrégation de N processus $AR(1)$ indépendants, notés X_{jt}, et en supposant que les b_j sont des variables aléatoires de loi Beta(p,q), montre que

$$\bar{X}_t \sim I(1 - q/2), \tag{1.37}$$

avec $q > 0$. Si $q > 1$, alors $1 - q/2 < 1/2$ et $\bar{X}_t$ a une variance finie. Si $q \in]0, 1]$, alors $1 - q/2 \geq 1/2$ et $\bar{X}_t$ n'a pas une variance finie. De plus, nous observons que la fonction d'autocorrélation de $\bar{X}_t$ décroît lentement.

De nombreuses variables économiques sont des agrégats d'un grand nombre de micro-variables telles que le revenu total personnel, le chômage, la consommation de biens non durables pour ne citer qu'eux. Or, bien que les composantes de ces séries ne soient, en réalité, pas indépendantes, le résultat précédent peut être généralisé. Granger suggère donc que les modèles utilisés pour décrire des séries temporelles financières auront un très grand nombre de paramètres, ce qui n'était pas le cas en pratique jusqu'alors.

Granger (1980) a montré que de l'agrégation d'équations dynamiques (pour reprendre ses propres termes), c'est-à-dire d'équations dont la variable présente est reliée aux variables passées, découle une nouvelle classe de modèles dont les propriétés sont fondamentalement différentes de celles utilisées jusqu'alors en économétrie.

En somme, Granger (1980) affirme (et prouve) que les modèles utilisés jusqu'en 1980, comme par exemple les modèles ARMA, sont mal spécifiés, avec toutes les conséquences que cela entraîne et notamment l'inefficacité des estimations. En outre, cet article ouvre la voie à de nettes améliorations dans la prédiction à long terme des séries temporelles économiques et financières et contribua à l'essor des modèles fractionnairement intégrés.

Les changements structurels

Depuis la fin des années 1990, les ruptures structurelles dans les séries temporelles financières ont fait l'objet de nombreuses études. Gouriéroux et Jasiak (2001) montrent que la mémoire longue peut provenir de la durée des régimes qui suit une loi de puissance. D'autres parmi lesquelles celles de Granger et Hyung (2004), Liu (2000), Diebold et Inoue (2001), ou encore Hyung et Franses (2001) s'accordent à dire que les changements structurels génèrent au sein des séries temporelles des propriétés similaires à celles d'un processus fractionnairement intégré $I(d)$, à savoir une fonction d'autocorrélation décroissant lentement de façon hyperbolique.

Ces ruptures structurelles sont des chocs de très forte importance, tels les deux chocs pétroliers dans les années 1970. Hyung et Franses (2001), lors de l'étude de 23 séries de taux d'inflation américains entre janvier 1967 et août 2000, ont détecté 4 changements structuraux dont daté de juillet 1973 pouvant correspondre au choc pétrolier de 1973-1974. Quant à Lobato et Savin (1998), en étudiant des séries de rendements journaliers et des séries de volatilité journalière de 1964 à 1994, ont pris janvier 1973 comme date de changement. Il faut préciser qu'il n'ont pas réussi à prouver que la mémoire longue était due au changement de 1973.

Pour modéliser une série temporelle financière grâce à un modèle du type changement de structure, il faut préalablement identifier ces changements. Pour cela, nous disposons de la méthode séquencielle de Bai et Perron (1998) et la procédure d'estimation de répartition de Bai (1997). Dans leur étude, Hyung et Franses (2001) remarquent que le nombre de changements identifiés varie selon la méthode utilisée. Toutefois, ces deux méthodes de détermination endogène des changements tendent à surestimer leur nombre en présence ou d'hétéroscédasticité.

Parmi les modèles de changement de structure les plus connus, citons le modèle "Occasional-Break" utilisé par Granger et Hyung (2004) et Hyung et Franses (2001) ainsi que le modèle "STOPBREAK" (Stochastic Permanent Break) d'Engle et Smith (1999) qui inclut une fonction de transition lissée endogène pour indiquer les ruptures strucutrelles.

Hyung et Franses (2001) ont comparé les modèles de changement de structure avec des modèles à mémoire longue et des modèles linéaires autorégressifs. Leurs travaux ont montré que les modèles linéaires ne sont pas aussi bons que les deux autres. De plus, les modèles de changement de structure et les modèles à mémoire longue fournissent des résultats équivalents. Par ailleurs, Granger et Hyung (2004) ont montré que les modèles de changement de structure traduisent parfois mieux les séries temporelles "in sample", mais sont en général moins bons en matière de prédiction (mais pas de manière significative) que les processus fractionnairement intégrés $I(d)$.

Il existe une autre classe de modèles capables de prendre en compte les ruptures structurelles, basés sur un changement de régime. Parmi eux, citons le modèle à changement de régime markovien [18] proposé par Hamilton (1989) et repris par Diebold et Inoue (2001) dans leur étude de la mémoire longue et du changement de régime. Ces derniers rappellent que ce modèle étant $I(0)$, sa capacité à décrire des séries temporelles financières à mémoire longue ne doit pas être bonne.

1.3.4 Signification

La mémoire longue en finance a été l'objet de nombreux travaux de recherche depuis le début des années 80 et les articles de Granger (1980) et Granger et Joyeux (1980). Bien qu'il

18. Le nom original est : Markov-switching model.

y ait une très riche littérature sur la mémoire longue, aussi bien théorique [19] qu'empirique [20], très peu d'articles débattent finalement de ce que traduit concrètement ce phénomène.

Di Matteo, Aste, et Dacorogna (2005) établissent une relation forte entre l'exposant de mémoire longue des indices boursiers et l'état de développement du marché correspondant. Pour cela, ils étudient plusieurs marchés financiers ayant des états de développement différents (des marchés matures et très liquides, et des marchés émergents et moins liquides). Leurs résultats leur permettent de conclure que le paramètre de mémoire longue d d'un marché mature est plus petit que celui d'un marché émergent. En particulier, lorsque les marchés sont matures (par exemple Nasdaq, Nikkei, CAC 40), $d < 0$ et, au contraire, lorsque ils sont émergents (par exemple Russian AK&M, Indonesian JSXC) $d > 0$.

Or précédemment, Dacorogna, Gençay, Müller, Olsen, et Pictet (2001) ont proposé que l'efficience [21] du marché résulte de l'interaction entre un grand nombre d'agents aux caractéristiques hétérogènes. Ainsi, la présence d'un grand nombre d'agents différents caractériserait un marché mature tandis que l'absence de certaines catégories d'agents serait le signe d'un marché moins développé. Pour ce test de l'hypothèse de l'hétérogénéité des agents de marché, Dacorogna, Müller, Olsen, et Pictet (2001) ont considéré la fréquence à laquelle les agents opèrent sur le marché comme marque d'hétérogénéité.

En conclusion, puisqu'un marché efficient est un marché très liquide (il le faut pour que l'arrivée d'information soit immédiatement répercutée dans le prix de l'action), plus un marché est liquide, plus d est petit (négatif). En outre, puisque l'efficience résulte de l'interaction d'un grand nombre d'agents hétérogènes, alors nous en déduisons que plus l'hétérogénéité est grande parmi les agents plus d est petit. En d'autres termes, moins il y a de mémoire longue. En particulier, sur des marchés parfaitement efficients il n'y a pas du tout de mémoire longue ($d = 0$).

Puisque la mémoire longue semble être reliée à l'hérérogénéité des agents, il convient de s'interroger sur la manière de modéliser cette hétérogénéité.

1.4 La prise en compte de l'hétérogénéité des agents

Les agents qui interviennent sur le marché se distinguent les uns des autres par de nombreuses caractéristiques, à savoir

– leurs moyens financiers : des sommes allant de quelques quelques milliers voire quelques dizaine de milliers d'euros pour les particuliers plaçant leurs économies en bourse, jus-

19. Le lecteur est invité à se référer à Granger (1980), Granger et Joyeux (1980), Hosking (1981), Geweke et Porter-Hudak (1983), Lo (1991), Sowell (1992), Cheung et Diebold (1994), Robinson (1995), Engle et Smith (1999), Diebold et Inoue (2001), Dittman et Granger (2002).

20. Le lecteur est invité à se référer à Diebold et Rudebusch (1989, 1991a), Hassler et Wolters (1995), Hyung et Franses (2001), Bos, Franses, et Ooms (2002), Van Dijk, Franses, et Paap (2002)

21. L'hypothèse d'efficience du marché de Fama (1965) stipule que le marché est constitué d'un grand nombre d'agents aux attentes rationnelles, maximiseurs de profit qui essaient de prédire les prix futurs des actions. De plus, les informations importantes sont supposées gratuites et accessibles à tous les participants. Il en résulte que le prix de chaque action, à tout moment, reflète l'information disponible.

qu'à des montants colossaux de l'ordre de plusieurs dizaines de milliards d'euros pour les investisseurs institutionnels en passant par les dizaines de millions d'euros gérés par les traders ou gestionnaires de portefeuille,
- leurs objectifs : certains traders ont des objectifs à très court terme, alors que les particuliers sont prêts à placer leur épargne, que ce soient des PEA, contrats d'assurance vie par exemple, à long voire très long terme : plusieurs années à plusieurs dizaines d'années. Quant aux grands investisseurs institutionnels, leur horizon de gestion est généralement de l'ordre de quelques mois,
- les informations dont ils disposent : les petits épargnants qui gèrent eux même leurs capital boursier n'ont pas forcément accès aux mêmes informations que les professionnels, même s'ils sont abonnés aux revues spécialisées, et de plus le temps de latence entre le moment où paraît l'information et le moment où ils y ont accès fait qu'il auront un temps de réaction beaucoup plus lent,
- ainsi que les outils mathématiques : les connaissances mathématiques et le matériel informatique à la disposition des petits épargnants n'est assurément pas le même que ceux dont bénéficient les investisseurs professionnels.

Bien que tous les acteurs d'un marché financier n'aient qu'un seul et même but : maximiser leur profit et minimiser leur risque, toutes ces caractéristiques engendrent une grande diversité de comportements.

Certes, cette abondance de comportements hétérogènes est atténuée par certains phénomènes. Ceux-ci, très longtemps négligés bien qu'ayant un réel impact sur le comportement des agents, ont suscité l'intérêt d'un nombre grandissant de chercheurs à partir de la fin des années 80. Ainsi on a commencé à entendre parler de rumeurs Banerjee (1993), de mimétisme Orléan (1995) (en résumé, cela signifie que pour un grand nombre d'agents, les opinions sont modifiées de manière endogène comme résultats de l'interaction entre les agents individuels), de comportement de masse Banerjee (1992), Froot, Scharfstein, et Stein (1992), Kirman (1993), Cont et Bouchaud (2000), Corcos, Eckmann, Malaspinas, Malevergne, et Sornette (2002), de modes Shiller (1989). Malgré ces phénomènes, et grâce au grand nombre d'agents sur le marché, l'hétérogénéité des agents demeure.

Or bien qu'aujourd'hui, les agents intervenant sur le marché soient considérés par l'ensemble des chercheurs comme étant hétérogènes, cela n'a pas toujours été le cas. Auparavant et jusqu'au début des années 90, le concept d'agent représentatif était utilisé pour modéliser l'agrégation des comportements des agents. En effet, les chercheurs supposaient que l'agrégation des décisions, des actions de l'ensemble des agents constituant le marché était équivalent aux décisions et actions d'un agent unique qui représentait alors tous les agents. Cette hypothèse, fort pratique pour la modélisation, était peu réaliste.

En effet, ce concept d'agent unique représentatif a été vivement critiqué depuis le début des années 1990, et plus précisément depuis que Kirman (1992) a affirmé et argumenté qu'un agent représentatif n'est pas une bonne solution pour caractériser l'agrégation d'agents hétérogènes. En effet, selon lui, premièrement, rien ne permet de justifier que l'agrégation de différents individus, même s'ils sont tous des maximiseurs de profit, donnent naissance à un maximiseur unique. D'autre part, la rationalité collective ne découle pas nécessairement de la rationalité des individus qui composent le groupe. Plus précisément, il faut non seulement considérer l'action propre à chaque individu, mais également leurs interactions.

C'est pourquoi, de nos jours, les modèles d'agents prennent en compte cette hétérogénéité entre les agents. C'est également la raison pour laquelle le modèle que nous proposons tient compte de la diversité des agents et est capable de capturer différent comportement d'agents.

Conclusion

Cette revue de littérature a été l'occasion d'une part, de nous familiariser avec les faits stylisés des rendements d'action et de leur volatilité et, d'autre part, de nous interroger sur deux problématiques. La première concerne la mémoire longue. En plus de l'erreur de spécification de modèle, deux hypothèses ont été avancées : celle de l'agrégation et celle de changements structuraux au sein des séries de volatilité. Quant à la signification de la mémoire longue en finance, très peu d'articles ont été consacrés. Dacorogna, Gençay, Müller, Olsen, et Pictet (2001), par exemple, ont traité ce problème. Selon eux, la mémoire longue est reliée à l'hétérogénéité des agents. Dans le domaine de la modélisation, la prise en compte de l'hétérogénéité des agents a connu un très grand essor depuis le milieu des années 1990 et les critiques de l'approche par agent représentatif unique (Kirman 1992) pour devenir une approche standard.

En conclusion, la revue de littérature a dessiné les contours de notre modèle de volatilité : un modèle fondé sur l'hypothèse de l'agrégation de Granger, dans lequel les comportements hétérogènes des agents sont pris en compte, devrait pouvoir traduire de la mémoire longue et en fournir une interprétation.

Chapitre 2

Le Modèle d'hétérogénéité

Comme nous l'avons rappelé en introduction, plusieurs hypothèses ont été proposées dans l'abondante littérature consacrée à la mémoire longue des séries temporelles économiques et financières pour expliquer les différentes origines de ce phénomène. La première, que l'on doit à Granger et Joyeux (1980), concerne la nature même de cette série temporelle. En effet, ils considèrent que ce type de série résulte de l'agrégation de micro-variables qui suivent des processus à mémoire courte, et que cette agrégation est responsable de la mémoire longue. La seconde est la présence de rares changements de structure au sein de la série temporelle financière (Hamilton 1989, Diebold et Inoue 2001, Gouriéroux et Jasiak 2001, Granger et Hyung 2004, Gadea, Sabaté, et Serrano 2004, Davidson et Sibbersten 2005, parmi d'autres).

Notre modèle est basé sur la proposition de Granger et Joyeux (1980) ; il justifie la mémoire longue de la série de la log-volatilité par l'agrégation de micro-variables destinées à représenter les attentes individuelles et hétérogènes de chaque investisseur du marché. Ainsi, nous supposons que le marché agrège les anticipations des agents, et que cette agrégation conduit à la log-volatilité observée.

Dans ce chapitre, nous commençons par poser le cadre général de notre modèle (section 2.1). Dans la section 2.2, nous détaillons les propriétés de la (log-)volatilité réalisée et les illustrons au moyen d'un exemple. La section 2.3 est consacrée à l'étude du paramètre d'hétérogénéité. Nous décrivons dans quelles conditions on obtient de la mémoire longue. Enfin, dans la section 2.4, nous discutons de différents comportements d'investisseurs capturés par notre modèle.

2.1 Cadre général

Le rôle central de la log-volatilité dans notre modèle vient de la remarque que, comme nous l'avons rappelé dans le chapitre précédent, il est raisonnable, dans une première approximation, de considérer que la volatilité réalisée suit une distribution log-normale. Comme hypothèse supplémentaire, nous supposons que

(H1) *le processus de log-volatilité* $\{\omega_t\}_{t\in\mathbb{Z}}$ *suit un processus stochastique gaussien,*

ce qui, bien évidemment, assure que la volatilité elle-même possède une distribution log-normale.

A présent, la théorie micro-économique standard nous dit que tout agent rationnel ayant à faire face à un problème de décision, cherche à optimiser l'allocation de ses ressources en se basant sur la totalité des informations dont il dispose. Dans notre contexte, cela signifie que tout investisseur rationnel s'efforce de déterminer la meilleure prédiction de la volatilité réalisée future $\hat{\sigma}_{t+\tau}$, $\tau \geq 1$, basée sur l'ensemble $\{\sigma_1, \ldots, \sigma_t\}$ de ses observations passées, s'il ne dispose que de cette source d'information.

Au sens de l'erreur quadratique moyenne, le meilleur prédicteur est donné par

$$\begin{aligned} \hat{\sigma}_{t+\tau} &= \mathrm{E}\left[\sigma_{t+\tau} | \sigma_1, \ldots, \sigma_t\right], & (2.1)\\ &= \mathrm{E}\left[\mathrm{e}^{\omega_{t+\tau}} | \, \omega_1, \ldots, \omega_t\right], & (2.2)\\ &= \mathrm{e}^{2\sigma_\omega(\tau)^2} \cdot \mathrm{e}^{\hat{\omega}_{t+\tau}}, & (2.3) \end{aligned}$$

où $\hat{\omega}_{t+\tau}$ décrit le meilleur prédicteur de la log-volatilité, basé sur le même ensemble d'observations et $\sigma_\omega(\tau)^2$ est l'erreur de prédiction pour un pas de τ, *i.e.* l'erreur quadratique moyenne $\mathrm{E}\left[(\hat{\omega}_{t+\tau} - \omega_{t+\tau})^2\right]$. Puisque la log-volatilité est supposée suivre un processus gaussien, le meilleur prédicteur $\hat{\omega}_{t+\tau}$ est donné par une combinaison linéaire des observations passées. En particulier (voir Brockwell et Davis (1990, pp. 162-168))

$$\hat{\omega}_{t+1} = \bar{\omega} + \sum_{i=0}^{t-1} \varphi_{t,i} \cdot (\hat{\omega}_{t-i} - \omega_{t-i}), \tag{2.4}$$

où la séquence des coefficients $\{\varphi_{t,i}\}_{i\geq 0}$ et la moyenne à long terme $\bar{\omega}$ dépendent du modèle spécifique auquel fait appel chaque investisseur et de leur méthode particulière de calibrage employée pour estimer les paramètres structuraux du modèle. Ainsi, nous nous attendons à ce que les coefficients φ_i soient spécifiques à chaque acteur du marché de telle façon que chaque investisseur obtient une estimation du niveau futur de la log-volatilité différente. De plus, chaque agent peut incorporer des variables économiques exogènes ou des informations de nature publique ou privée s'il pense que cela peut améliorer les capacités prédictives de son modèle [22].

En outre, en considérant que le marché effectue une agrégation des anticipations des agents [23], nous posons comme principe que

(H2) *la log-volatilité réalisée est la moyenne de toutes les anticipations individuelles de la log-volatilité.*

Ainsi, en notant par $\hat{\omega}_{i,t}$ la prédiction de l'agent i, la log-volatilité réalisée est donnée par

$$\omega_t(n) = \frac{1}{n} \sum_{i=1}^{n} \hat{\omega}_{i,t}, \tag{2.5}$$

22. Le modèle propre a chaque investisseur peut être explicite ou au contraire purement implicite. Il traduit alors l'intuition ou le mode de réflexion d'un agent qui n'utiliserait pas de modèle mathématique.

23. Sur le marché, le prix d'une action est la conséquence des anticipations du prix de cette action par les acteurs du marché. Dans notre modèle, on suppose qu'il en est de même pour la volatilité.

si le marché est composé de n participants. Sous l'hypothèse d'un nombre infiniment grand d'investisseurs la log-volatilité réalisée s'écrit

$$\omega_t = \lim_{n\to\infty} \omega_t(n). \tag{2.6}$$

Notre approche peut paraître hautement simpliste dans la mesure où le véritable processus produisant la volatilité réalisée implique certainement des transformations non-linéaires des anticipations individuelles avant qu'elles soient vraiment agrégées via le processus de formation des prix, qui est bien connu pour reposer sur différents mécanismes de rétroactions positives ou négatives (Shiller 2000, Sornette 2003). Toutefois, dans une première tentative de rendre compte de l'impact de l'hétérogénéité des anticipations des investisseurs sur la dynamique de la volatilité réalisée, et en l'absence de propositions nous permettant de modéliser ces interactions non-linéaires, nous avons préféré nous restreindre à cette hypothèse de linéarité.

De plus, afin de pouvoir résoudre les calculs, nous avons besoin d'autres hypothèses simplificatrices. Pour cela, intéressons nous à un agent i, et notons par $\hat{X}_{i,t}$ son anticipation du niveau futur de l'*excès* de log-volatilité sur sa moyenne à long terme [24]

$$\hat{X}_{i,t} = \hat{\omega}_{i,t} - \bar{\omega}, \tag{2.7}$$

le jour t, et par

$$\bar{X}_{n,t} = \omega_t(n) - \bar{\omega} = \frac{1}{n}\sum_{i=1}^{n} \hat{X}_{i,t}, \tag{2.8}$$

l'excès de log-volatilité réalisée par rapport à sa moyenne de long terme. Nous supposons que l'anticipation $\hat{X}_{i,t}$ dépend de l'anticipation, $\hat{X}_{i,t-1}$, que l'agent a faite la veille et de l'excès de log-volatilité réalisée $\bar{X}_{n,t-1}$, mais également une information publique ε_t et privée $\eta_{i,t}$ selon l'équation récursive

$$\hat{X}_{i,t} = \varphi_i \cdot \hat{X}_{i,t-1} + \psi_i \cdot \bar{X}_{n,t-1} + c_i\varepsilon_t + \eta_{i,t} \qquad t \geq 0, \qquad i = 1, 2, \ldots, n. \tag{2.9}$$

Cette formulation ressemble à celle la définition 8 d'un processus EGARCH dans la section 1.2 du chapitre précédent. De l'équation (2.4), nous devrions nous attendre à ce que $\psi_i = -\varphi_i$ si nous considérons des agents rationnels. Toutefois, afin de généraliser notre modèle au cas où les participants du marché peuvent être des agents à rationalité limitée, nous permettons $\psi_i \neq -\varphi_i$.

Soulignons que la dynamique d'ordre 1 (2.9) peut paraître, là encore, fortement simpliste, et en réalité elle l'est. En réalité, les agents utilisent certainement des procédés de prédiction plus sophistiqués pour obtenir leurs prévisions du futur niveau de la (log)-volatilité réalisée. Ainsi, nous devrions considérer des dynamiques d'ordre supérieur. Toutefois, nous émettons l'hypothèse selon laquelle ces dynamiques d'ordre supérieur peuvent être réduites à une agrégation de dynamiques d'ordre 1. De cette façon, $\hat{X}_{i,t}$ ne caractérise pas exactement la prévision d'un agent individuel , mais plutôt celle d'une classe d'agents.

Dans ce cadre, le vecteur de dimension n, $\hat{X}_t = \left(\hat{X}_{1,t}, \ldots, \hat{X}_{n,t}\right)'$, suit un processus (vectoriel) auto-régressif d'ordre 1. Les bruits gaussiens $\{\varepsilon_t\}$, $\{\eta_{1,t}\}$, $\{\eta_{2,t}\}$,..., sont supposés

24. ω pourrait être spécifique à chaque agent et être noté ω_i, mais cela ne changerait rien à nos résultats.

indépendants, centrés et de variance respective $\text{Var}(\varepsilon_t) = \sigma^2$ et $\text{Var}(\eta_{i,t}) = \mu_i^2$, pour tout i. De plus, les coefficients structurels (φ_i, ψ_i) et c_i, qui caractérisent le comportement des investisseurs, peuvent être considérés comme étant le résultat de tirages indépendants d'une même loi (appelée la distribution d'hétérogénéité). Ces tirages sont indépendants des valeurs des bruits et les variables (φ, ψ) et c sont indépendantes les unes des autres. La distribution de (φ, ψ) doit vérifier quelques hypothèses, qui seront explicitées ci-après, afin de s'assurer de la stationnarité du processus stochastique $\left\{\hat{X}_t\right\}_{t\in\mathbb{Z}}$.

L'intérêt de la dynamique (2.9) réside dans la simplicité de l'interprétation du coefficient structurel (φ, ψ) et de ses propriétés distributionnelles. Il est évident que l'impact de l'anticipation précédente $\hat{X}_{i,t-1}$ sur l'anticipation actuelle $\hat{X}_{i,t}$ dépend de la valeur du coefficient d'hétérogénéité φ_i :

1. quand φ_i est proche de (mais strictement inférieur à) 1, l'impact de l'anticipation précédente sur l'anticipation actuelle est très important, à moins qu'une importante information arrive, *i.e.* à moins que nous observions un grand ε_t et/ou un grand $\eta_{i,t}$, la valeur actuelle de l'excès de log-volatilité escompté $\hat{X}_{i,t}$ sera très proche de la précédente. Globalement, l'agent croit dans la continuité des conditions précédentes du marché en ce qui concerne la volatilité ;
2. quand φ_i est proche de (mais supérieur à) -1, l'impact de l'anticipation précédente reste toujours plus importante que l'arrivée d'une information nouvelle, mais l'agent a tendance systématiquement à croire à un changement du niveau de volatilité puisque son anticipation devient l'opposée de celle qu'il a faite la veille ;
3. enfin, quand $|\varphi_i|$ est proche de 0, l'anticipation est à peine reliée à celle de la veille et repose principalement sur le flux d'informations.

Pour résumer, dans les deux premières situations, nous pouvons remarquer que le marché est fortement auto-référentiel : l'impact d'une information nouvelle reste faible, d'autant plus que φ_i est proche de 1 en valeur absolue. Au contraire, lorsque $|\varphi|$ est proche de 0, le degré de réactivité des participants du marché à l'arrivée d'une information nouvelle est fort. Ainsi, la valeur absolue de φ peut être interprétée comme une manière de quantifier le degré d'efficience du marché. En effet, puisque dans un marché efficient, le prix reflète à tous moment l'information pertinente disponible, dès qu'une nouvelle information arrive, elle doit immédiatement être prise en compte. Ce qui nécessite une une grande réactivité des investisseurs. Par conséquent, si $|\varphi| = 0$, on a affaire à un marché efficient. Les mêmes considérations tiennent évidemment pour ψ_i, mais au lieu de faire référence à l'anticipation passée, il fait référence à la volatilité réalisée passée, publiquement observable.

2.2 Propriétés de la (log-)volatilité

L'équation (2.9) peut s'écrire sous une forme plus compacte comme

$$\hat{X}_t = A\hat{X}_{t-1} + \varepsilon_t \cdot C + \eta_t, \tag{2.10}$$

où $A = D + \frac{1}{n}\Psi \cdot 1_n'$ est une matrice de taille $n \times n$ avec $D = \text{diag}(\varphi_1, \ldots, \varphi_n)$, $\Psi = (\psi_1, \cdots, \psi_n)'$ et $1_n = (1, \ldots, 1)'$ est de taille n, tandis que $C = (c_1, \ldots, c_n)'$ et $\eta_t = (\eta_{1,t}, \ldots, \eta_{n,t})'$. En outre,

il est nécessaire que le rayon spectral de A soit plus petit que 1 si l'on veut s'assurer de la stationnarité du processus vectoriel $\left\{\hat{X}_t\right\}_{t\in\mathbb{Z}}$.

Sous l'hypothèse $||A|| < 1$ [25], *ce qui sera supposé par la suite dans ce chapitre*, la solution stationnaire de l'équation (2.10) est donnée par

$$\hat{X}_t = \sum_{k=0}^{\infty} A^k C \varepsilon_{t-k} + \sum_{k=0}^{\infty} A^k \eta_{t-k}. \tag{2.11}$$

Nous en déduisons le résultats suivant (preuve en annexe 2.A) :

Proposition 1. Dans la limite d'un grand nombre d'agents économiques, $n \to \infty$, l'excès de log-volatilité réalisée $\left\{\bar{X}_t\right\}_{t\in\mathbb{Z}}$, suit un processus auto-régressif à moyenne mobile d'ordre infini

$$\left(1 - \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] L^{k+1}\right) \bar{X}_t = \mathrm{E}\left[c\right] \cdot \left(\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] L^k\right) \varepsilon_t, \tag{2.12}$$

où L décrit l'opérateur de décalage : $L\bar{X}_t = \bar{X}_{t-1}$.

Nous constatons que le flux de l'information privée $\{\eta_{i,t}\}$ n'entre pas en jeu puisque. Nous pouvions nous attendre à ce résultat puisque, sous nos hypothèses, le flux d'information privée ne véhicule aucune information en moyenne au travers des agents.

Comme résultat dérivé de la proposition ci-dessus, nous voyons que la solution de l'équation (2.12) peut être exprimée de façon pratique en termes d'une moyenne mobile d'ordre infini, à condition que $\sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] z^{k+1} \neq 1$ pour tout z dans le cercle unité :

$$\bar{X}_t = \mathrm{E}\left[c\right] \cdot \sum_{k=0}^{\infty} \tilde{\beta}_k \varepsilon_{t-k}, \tag{2.13}$$

où les $\tilde{\beta}_k$ sont donnés formellement par les coefficients de la série

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k := \mathrm{E}\left[\frac{1}{1 - x\cdot\varphi}\right] \left(1 - x\mathrm{E}\left[\frac{\psi}{1 - x\cdot\varphi}\right]\right)^{-1} \tag{2.14}$$

$$= \mathrm{E}\left[\frac{1}{1 - x\varphi}\right] \left(1 - \mathrm{E}\left[\frac{xg(\varphi)}{1 - x\varphi}\right]\right)^{-1}, \tag{2.15}$$

où $g(\varphi) = \mathrm{E}[\psi|\varphi]$.

Revenons à présent sur l'équation (2.12) et étudions la plus en détail. Dans le membre de droite, le processus $Y_t := \left(\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] L^k\right) \varepsilon_t$, est stationnaire si et seulement si

$$\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right]^2 < \infty. \tag{2.16}$$

25. Nous prenons comme norme de la matrice A son rayon spectral, qui correspond à la plus grande norme des valeurs propres λ_i (car elles peuvent être des nombres complexes) : $R(A) = \max_i |\lambda_i|$. Les conditions permettant au rayon spectral de A d'être inférieur à 1 ne seront pas détaillées ici dans la mesure où ce ne sont que des conditions nécessaires pour que la série temporelle agrégée $\bar{X}_t$ soit stationnaire.

Cela nécessite que la loi de φ soit concentrée sur $]-1,1[$, mais n'est pas suffisant, comme l'ont montré Gonçalves et Gouriéroux (1988). En outre, Y_t affiche de la mémoire longue pour autant que

$$\sum_{k=0}^{\infty} |\mathrm{E}\left[\varphi^k\right]| = \infty. \tag{2.17}$$

En nous restreignant au cas où il est dit que Y_t affiche de la mémoire longue si sa densité spectrale

$$f_Y(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi}\left|\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] e^{-ik\lambda}\right|^2 \sim \lambda^{-2d_Y}, \quad d_Y \in \left]-\frac{1}{2},\frac{1}{2}\right[, \tag{2.18}$$

nous concluons que la densité de φ, si elle existe, devrait se comporter comme $(1-\varphi)^{-d_Y}$ lorsque φ tend vers 1. Finalement, Y_t affiche de la mémoire longue si $d_Y \in]0,1/2[$ et de l'anti-persistance si $d_Y \in]-1/2,0[$. La borne supérieure de d_Y est obligatoire pour que le processus soit stationnaire au deuxième ordre.

Quant à la partie gauche de l'équation (2.12), le filtre $A(L) := 1 - \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] L^{k+1}$ est bien défini si

$$\sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right]^2 < \infty, \tag{2.19}$$

ce qui, par l'inégalité de Cauchy-Schwartz, est satisfait si $\mathrm{E}\left[\psi^2\right] < \infty$ et $\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^{2k}\right] < \infty$. Cette condition requise n'est, cependant, pas nécessaire. Comme autre hypothèse, nous supposons que

$$\sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right]^2 x^{k+1} \neq 1, \tag{2.20}$$

pour tout $x \in [-1,1]$ ce qui garantit que $A(z)$ ne s'annule pas pour tout z à l'intérieur du cercle unité. Le comportement de $A(z)$, quand z tend vers un, dépend de l'espérance conditionnelle de ψ sachant φ, à savoir $g(\varphi) := \mathrm{E}\left[\psi|\varphi\right]$. En effet, soit $A(z)$ demeure fini quand $z \to 1$, et X_t affiche de la mémoire longue si et seulement si Y_t lui-même affiche de la mémoire longue, soit $A(z)$ diverge de manière hyperbolique, quand $z \to 1$, et la mémoire (courte ou longue) du processus de la log-volatilité résulte d'un mélange entre les propriétés de la distribution de φ et de l'espérance conditionnelle $g(\varphi)$. Plus précisément, nous pouvons établir que (voir la démonstration en annexe 2.B)

Proposition 2. Sous les hypothèses précédentes, si la densité de φ est $f(\varphi) \sim (1-\varphi)^{-\alpha}$, $\alpha \in (0,1/2)$, et si l'espérance conditionnelle de ψ sachant φ est $g(\varphi) \sim (1-\varphi)^{\beta}$, $\beta > \alpha - 1/2$, quand $\varphi \to 1^-$, la densité spectrale du processus de la log-volatilité se comporte comme

$$f_X(\lambda) \sim \lambda^{-2\min\{\alpha,\beta\}}, \quad \text{quand } \lambda \to 0. \tag{2.21}$$

2.3 Analyse du paramètre d'hétérogénéité du modèle

2.3.1 Absence d'hétérogénéité

Supposons que les investisseurs ne s'intéressent qu'à la volatilité réalisée passée et négligent totalement leur anticipation passée, *i.e.* $\varphi = 0$. Par conséquent, la dynamique des

anticipations individuelles s'écrit

$$\hat{X}_{i,t} = \psi_i \cdot \bar{X}_{n,t-1} + c_i \varepsilon_t + \eta_{i,t}, \tag{2.22}$$

avec $t \geq 0$ et $i = 1, 2, \ldots, n$. En sommant sur tous les agents i, en divisant par n, et en prenant la limite $n \to \infty$, nous obtenons en utilisant l'équation (2.8)

$$\bar{X}_t = \mathrm{E}\left[\psi\right] \cdot \bar{X}_{t-1} + \mathrm{E}\left[c\right] \varepsilon_t \qquad t \geq 0, \tag{2.23}$$

où $\bar{X}_t = \lim_{n \to +\infty} \bar{X}_{n,t}$. Le processus $\bar{X}_t$ est un simple processus AR(1) à mémoire courte uniquement puisque

$$\rho(h) = \mathrm{E}\left[\psi\right]^{|h|}. \tag{2.24}$$

En conclusion, nous remarquons que le coefficient d'hétérogénéité φ est responsable de la mémoire longue tandis que ψ influence principalement la mémoire courte. Cela montre clairement que le phénomène de mémoire longue est enraciné dans le comportement auto-référent des investisseurs, et plus précisément dans la référence à leurs propres anticipations passées et non pas dans la log-volatilité réalisée passée.

2.3.2 Absence de référence à la volatilité passée

A l'opposé du cas précédent, considérons que chaque agent base son anticipation sur son anticipation passée et sur les flux d'informations, mais néglige la volatilité réalisée passée, de sorte que la dynamique des anticipations individuelles s'écrive

$$\hat{X}_{i,t} = \varphi_i \cdot \hat{X}_{i,t-1} + c_i \varepsilon_t + \eta_{i,t}, \tag{2.25}$$

avec $t \geq 0$ et $i = 1, 2, \ldots, n$. L'expression des $\tilde{\beta}_k$ se simplifie alors en

$$\tilde{\beta}_k = \mathrm{E}\left[\varphi^k\right], \tag{2.26}$$

qui nous permet d'écrire l'excès de log-volatilité réalisée et sa fonction d'autocorrélation comme

$$\bar{X}_t = \mathrm{E}\left[c\right] \sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] \varepsilon_{t-k}, \qquad \text{et} \qquad \rho(h) = \frac{\sum_{k=0}^{\infty} \mathrm{E}[\varphi^k]\mathrm{E}[\varphi^{k+h}]}{\sum_{k=0}^{\infty} (\mathrm{E}[\varphi^k])^2}, \tag{2.27}$$

sous réserve que la condition de stationnarité $\sum_{k=0}^{\infty} (\mathrm{E}[\varphi^k])^2 < \infty$ soit vérifiée.

Il s'avère que les propriétés de cette dynamique ont été étudiées par Gonçalves et Gouriéroux (1988). En se basant sur leur résultat, nous pouvons remarquer que cette fonction de corrélation, à savoir la fonction d'autocorrélation de la série temporelle agrégée, est très différente de la fonction de corrélation moyenne $\bar{\rho}$. En effet, avant l'agrégation, la fonction d'autocorrélation de la prédiction de chaque agent individuel vaut

$$\rho_i(h) = \varphi_i^{|h|}, \qquad i = 1, 2, \ldots, \qquad h = 0, 1, \ldots \tag{2.28}$$

de sorte que

$$\bar{\rho}(h) = \mathrm{E}[\varphi^{|h|}] \qquad h = 0, 1, \ldots \tag{2.29}$$

Par ailleurs, la dynamique recèle de la mémoire longue si (et seulement si) la densité de φ diverge en 1 afin d'obtenir une décroissance hyperbolique de la fonction de corrélation (2.27). En d'autres termes, la densité devrait se comporter comme $(1-x)^{-d}$, $d < 1/2$, au voisinage de 1 (comme la densité de la loi Beta par exemple), afin d'obtenir $\rho(h) \sim \lambda \cdot |h|^{2d-1}$, lorsque $|h|$ tend vers l'infini. Dans une telle situation, nous pouvons conclure qu'une part significative des agents basent leur anticipation sur celle qu'ils ont réalisée la veille.

2.3.3 Indépendance entre les paramètres ψ et φ

Nous supposons que les paramètres ψ et φ du modèle sont indépendants et que $\varphi \overset{loi}{=} \mathrm{Beta}(-\alpha, 1+\alpha)$. Nous déduisons que

$$\mathrm{E}\left[\frac{1}{1-x\varphi}\right] = (1-x)^{\alpha}. \tag{2.30}$$

Par conséquent, l'équation (2.14) s'écrit

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{1}{(1-x)^{\alpha} - x\mathrm{E}[\psi]}. \tag{2.31}$$

Sous l'hypothèse $\alpha > 1$ qui assure que le dénominateur de l'équation (2.31) reste borné pour tout $|x| \leq 1$, la série $\sum_{k=0}^{\infty} \tilde{\beta}_k x^k$ se comporte comme $\mathrm{E}[\psi]^{-1}$ au voisinage de 1 et donc ne diverge pas hyperboliquement.

En conclusion, si ψ et φ sont indépendants, la dynamique de la log-volatilité ne recèle pas de mémoire longue.

2.3.4 Dépendance entre les paramètres ψ et φ

Supposons à présent qu'il y ait dépendance entre ψ et φ de sorte que

$$\mathrm{E}[\psi|\varphi] = (1-\varphi)^{\beta}, \tag{2.32}$$

et supposons que $\varphi \overset{loi}{=} \mathrm{Beta}(-\alpha, 1+\alpha)$. Nous déduisons que

$$\mathrm{E}\left[\frac{1}{1-x\varphi}\right] = (1-x)^{\alpha}. \tag{2.33}$$

De plus, obtenons

$$\mathrm{E}\left[\frac{g(\varphi)}{1-x\varphi}\right] = \frac{\Gamma(1+\alpha+\beta)}{\Gamma(1+\alpha)\Gamma(1+\beta)} F(1, -\alpha; 1+\beta; x), \tag{2.34}$$

où F est la fonction hypergéométrique (voir Abramowitz et Stegun (1965, chap. 15)). Ainsi, l'équation (2.14) s'écrit

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{(1-x)^{\alpha}}{1 - \dfrac{\Gamma(1+\alpha+\beta)}{\Gamma(1+\alpha)\Gamma(1+\beta)} x F(1, -\alpha; 1+\beta; x)}. \tag{2.35}$$

Sous l'hypothèse $\beta > 1$ qui assure que le dénominateur de l'équation (2.35) reste borné pour tout $|x| \leq 1$, la série $\sum_{k=0}^{\infty} \tilde{\beta}_k x^k$ se comporte comme $(1-x)^{\alpha}$ au voisinage de 1 et donc diverge hyperboliquement si $\alpha < 0$. L'expression de la densité spectrale est donc déduite des équations (2.13) et (2.35) :

$$f_X(\lambda) = \frac{\mathrm{E}\left[c\right]^2 \sigma_{\varepsilon}^2}{2\pi} \frac{|1 - e^{-i\lambda}|^{2\alpha}}{\left|1 - \mathrm{E}\left[\psi\right] e^{-i\lambda} F(1, -\alpha; 1+\beta, e^{-i\lambda})\right|^2} \tag{2.36}$$

avec $\mathrm{E}\left[\psi\right] = \frac{\Gamma(1+\alpha+\beta)}{\Gamma(1+\alpha)\Gamma(1+\beta)}$. Puisque $f_X(\lambda) \underset{\lambda \to 0}{\sim} C'|\lambda|^{2\alpha}$, la fonction d'autocorrélation s'écrit $\rho(h) \sim K|h|^{-2\alpha-1}$ quand $|h| \to \infty$, sous réserve que la condition de stationnarité de second ordre $\alpha > -1/2$ soit vérifiée.

Pour être complet, il nous reste à discuter du rôle joué par les deux paramètres α et β dans le modèle. Comme φ est supposé suivre une loi Beta$(-\alpha, 1+\alpha)$, alors $\mathrm{E}[\varphi] = -\alpha \in [0, 0.5]$. Cela implique que l'anticipation d'un agent n'est pas principalement basée sur son anticipation passée. Rappelons que $|\varphi|$ devrait être proche de 1 pour pouvoir observer un fort comportement auto-référentiel. En d'autres termes, les agents attachent, en moyenne, plus d'importance à la réalisation de la volatilité passée et aux flux d'informations. Plus α est proche de 0, moins d'importance est donnée à l'anticipation passée. Par conséquent, comme les anticipations des agents sont davantage basées sur des faits communs (volatilité réalisée passée et informations), alors l'hétérogénéité des opinions des agents décroît. En conclusion, la mémoire longue devrait décroître. Ce phénomène est illustré dans les deux graphiques de la figure 2.1, où sont tracées les fonctions d'autocorrélation, calculées par l'utilisation de la FFT (voir annexe 2.C), en mode linéaire (en haut) et en mode log-log (en bas)) pour $\beta = 1.5$ et différentes valeurs de α comprises dans $[-0.45, -0.05]$.

Sur la figure 2.1, nous avons tracé les densités de φ correspondant à différentes valeurs de α. Celui-ci confirme que plus α est proche de -0.5, plus forte est la divergence en 1. Cela signifie que la proportion d'agents attachant de l'importance à leur anticipation augmente et donc la mémoire longue. De plus, nous remarquons que la vitesse de la décroissance de la fonction d'autocorrélation et la divergence de la densité sont étroitement liées. Cela signifie que α a une influence à la fois sur la mémoire courte et sur la mémoire longue, ce qui n'est pas une surprise puisque le paramètre α contrôle à la fois le comportement de la densité de φ autour de 1 et 0.

Quant à β, pour pouvoir illustrer son rôle, nous nous intéressons à l'autre paramètre d'hétérogénéité ψ dont la moyenne $\mathrm{E}[\psi]$ vaut $\frac{\Gamma(1+\alpha+\beta)}{\Gamma(1+\alpha)\Gamma(1+\beta)}$. La limite de $\mathrm{E}[\psi]$ lorsque β tend vers 1 et vers l'infini est respectivement $\mathrm{E}[\psi] \underset{\beta \to 1}{\longrightarrow} 1+\alpha$ et $\mathrm{E}[\psi] \underset{\beta \to +\infty}{\longrightarrow} 0$. Donc, plus β est grand, plus ψ est proche de 0, en moyenne.

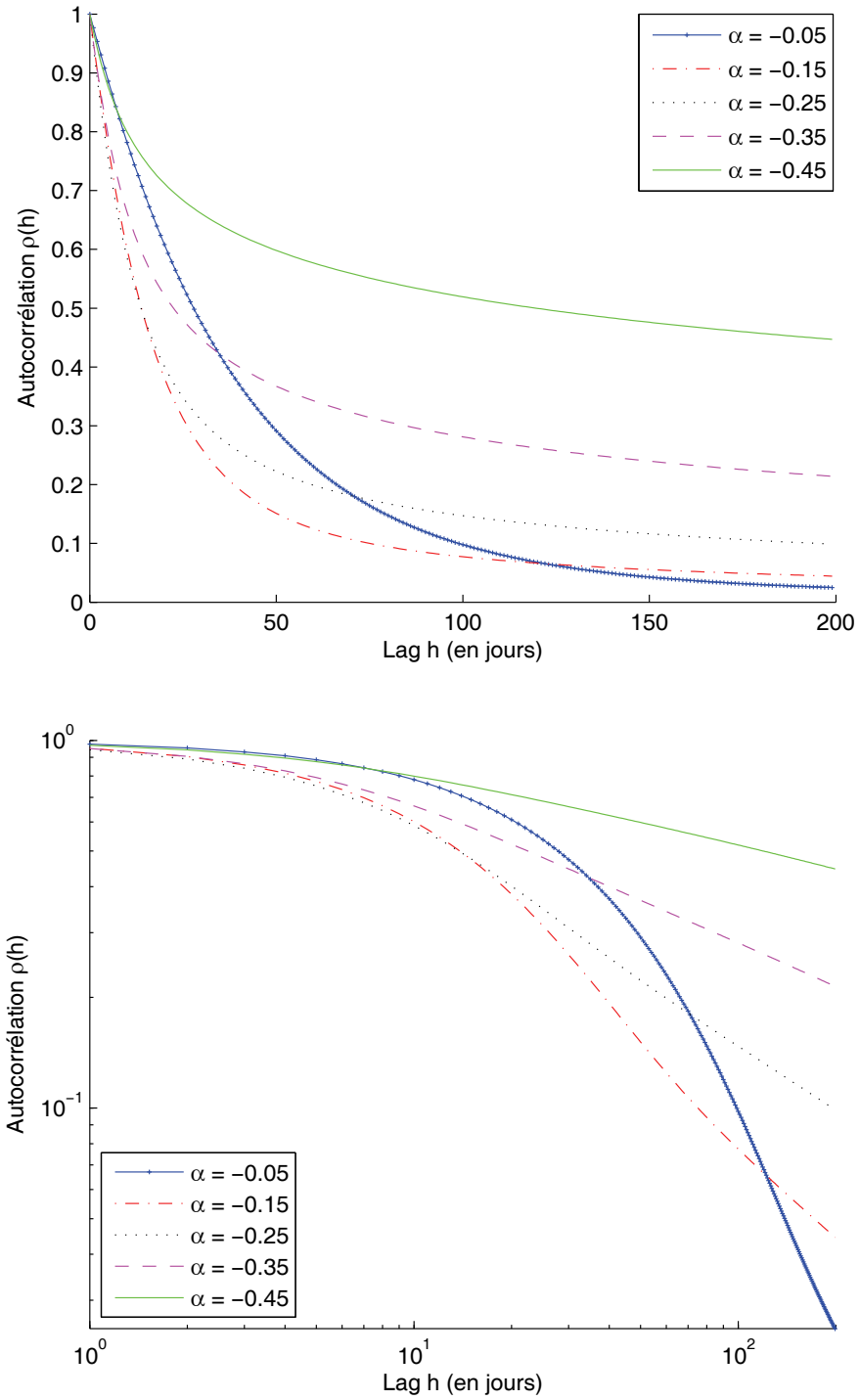

FIGURE 2.1 – Cette figure représente la fonction d'autocorrélation à horizon de 200 jours en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) pour différentes valeurs de α inclues dans $[-0.45, -0.05]$ et $\beta = 1.5$.

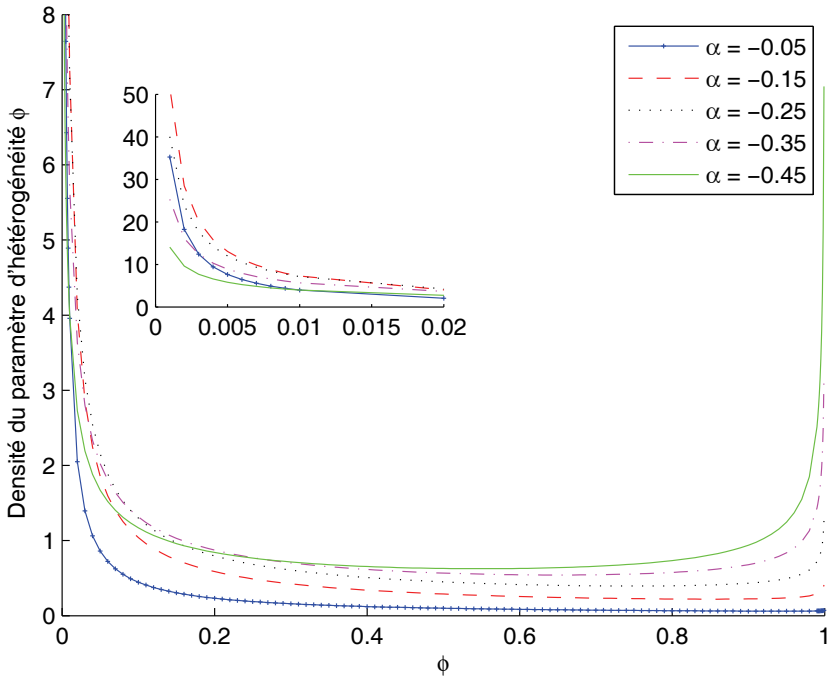

FIGURE 2.2 – Ce graphique représente la densité du coefficient d'hétérogénéité φ pour différentes valeurs de α inclues dans $[-0.45, -0.05]$ et $\beta = 1.5$. La loi de φ est une Beta$(-\alpha, 1+\alpha)$ et sa densité, $f(\varphi) = \frac{1}{B(-\alpha,1+\alpha)}\varphi^{-\alpha-1}(1-\varphi)^{\alpha}$, explique la divergence en 0 et 1.

A présent, en agrégeant le système (2.9), nous obtenons que la réalisation de la volatilité le jour t est égale à une quantité dépendant des anticipations passées, de la réalisation passée de la volatilité multiplié par la moyenne de ψ_i et des flux d'informations. Même si la quantité dépendant des anticipations passées influence la mémoire courte comme nous l'avons précédemment montré, toute chose égale par ailleurs, lorsque $\mathrm{E}[\psi]$ décroît, il y a moins de corrélations à court terme et la mémoire courte diminue. C'est la raison pour laquelle lorsque $\mathrm{E}[\psi]$ se rapproche de 0 (ce qui équivaut à β tend vers l'infini), la mémoire courte chute (voir figure 2.3).

2.4 Exemples de comportements

A présent, discutons en détail de plusieurs exemples typiques de dynamique décrits par notre modèle et relions les aux comportements d'agents.

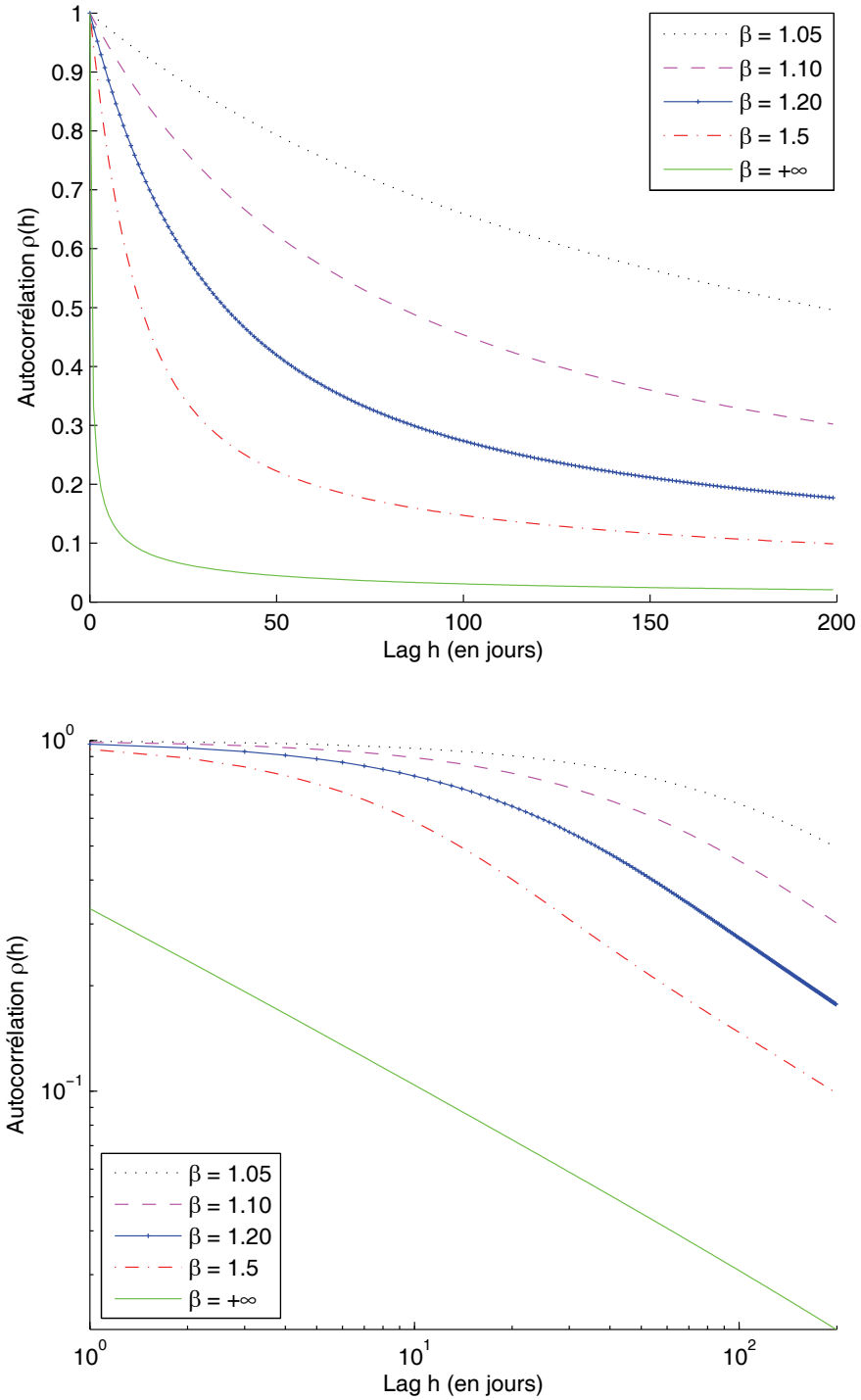

FIGURE 2.3 – Fonction d'autocorrélation à horizon de 200 jours en mode linéaire (graphique du haut) et en mode log-log (graphique du bas), obtenues à partir de $\alpha = -0.25$ et différentes valeurs de β.

2.4.1 Agents rationnels

Comme nous l'avons rappelé dans la section précédente, les agents rationnels basent leurs anticipations au temps t sur les innovations $\hat{X}_{i,s} - \bar{X}_{n,s}$, $s \leq t-1$. Ainsi, en posant

$$\psi_i = -\varphi_i, \tag{2.37}$$

pour tout $i = 1, \ldots, n$, cela nous permet de modéliser le comportement de la volatilité réalisée quand les investisseurs sont tous des agents totalement rationnels puisque cela nous permet de retrouver l'expression du prédicteur optimal (2.4), dans le cas particulier où les agents s'intéressent uniquement à la dernière estimation $\hat{X}_{i,t-1} - \bar{X}_{n,t-1}$. Dans ce cas, la dynamique des anticipations individuelles s'écrit

$$\hat{X}_{i,t} = \varphi_i \cdot \left(\hat{X}_{i,t-1} - \bar{X}_{n,t-1}\right) + c_i \cdot \varepsilon_t + \eta_{i,t}, \tag{2.38}$$

avec $t \geq 0$ et $i = 1, 2, \ldots, n$. Par l'équation (2.14), nous obtenons instantanément $\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = 1$, de sorte que

$$\bar{X}_t = \mathrm{E}\left[c\right] \varepsilon_t. \tag{2.39}$$

Remarquons, au passage, que nous n'avons pas besoin que chaque agent soit rationnel en posant $\psi_i = -\varphi_i$, mais seulement que $\mathrm{E}[\psi|\varphi] = -\varphi$, ce qui signifie que les agents ne sont pas nécessairement individuellement rationnels mais seulement en moyenne. Ainsi, si les agents sont rationnels (individuellement ou en moyenne), l'excès de log-volatilité réalisée reflète seulement l'information agrégée, publiquement disponible au temps t, à savoir $\mathrm{E}\left[c\right] \varepsilon_t$. Ce résultat est logique dans la mesure où les investisseurs rationnels font des anticipations optimales et, par conséquent, la volatilité réalisée peut uniquement transporter une information absente de l'innovation passée $\hat{X}_{i,t-1} - \bar{X}_{n,t-1}$. Alors, puisque le flux d'informations est supposé être un bruit blanc, la volatilité réalisée devrait aussi être un bruit blanc dans le cas où tous les investisseurs seraient rationnels.

2.4.2 Ecart à la rationnalité

Intéressons nous maintenant à un cas particulier d'agent à rationalité limitée. Pour cela, nous supposons [26] que

$$\psi = -\alpha \cdot \varphi + \bar{\varphi}', \tag{2.40}$$

où α et $\bar{\varphi}'$ sont deux constantes et tandis que φ suit une loi Beta(p, q) étirée, plus précisément une loi Beta(p, q) étendue sur l'intervalle $[-1, 1]$. Nous avons établi dans la section précédente que le cas des investisseurs rationnels était traité lorsque $\psi = -\varphi$. Dans un premier temps, pour nous écarter de la rationalité, nous avons ajouté le coefficient α qui a pour vocation de pondérer le recours de l'investisseur à la volatilité réalisée passée. En d'autres termes, lorsque α est proche de zéro, alors la volatilité réalisée passée n'intervient pas dans l'anticipation présente. Au contraire, lorsque α est grand, alors l'importance de l'anticipation passée devient faible par rapport à la réalisation passée. Pour s'éloigner encore un peu plus de la rationalité,

26. Remarque : la plupart des résultats de la section restent valables quand on suppose seulement que $\mathrm{E}\left[\psi|\varphi\right] = -\alpha \cdot \varphi + \bar{\varphi}'$.

nous introduisons un nouveau paramètre $\bar{\varphi}'$. Ainsi, ψ est une fonction affine de φ, et cette modélisation permet de prendre en compte les différents comportements que nous avons étudiés jusqu'à présent, à savoir :

- l'absence d'hétérogénéité, si $\varphi = 0$,
- l'absence de référence à la volatilité passée, si $\alpha = \bar{\varphi}' = 0$,
- la rationalité, si $\alpha = 1$ et $\bar{\varphi}' = 0$.

Comme nous nous attendons à ce que la portion d'investisseurs caractérisés par le paramètre d'hétérogénéité φ_i proche de 1 soit grand et, en réalité, diverge en 1, le paramètre p doit vérifier $p > 1$. Au contraire, nous prévoyons que la portion des agents représentés par un paramètre d'hétérogénéité φ_i proche de -1 reste fini, de sorte que q doive vérifier $0 < q < 1$.

Dans ce cadre, la dynamique de l'anticipation de l'agent i devient

$$\hat{X}_{i,t} = \varphi_i \cdot \hat{X}_{i,t-1} + (\bar{\varphi}' - \alpha\varphi_i) \cdot \bar{X}_{n,t-1} + c_i \cdot \varepsilon_t + \eta_{i,t}, \tag{2.41}$$

avec $t \geq 0$ et $i = 1, 2, \ldots, n$. A partir de l'équation (2.14), l'excès de log-volatilité $\bar{X}_t$ suit la moyenne mobile d'ordre infini (2.13) avec les coefficients $\tilde{\beta}_k$ donnés par la relation

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{(1-x)^{q-1} G(p,q,x)}{1 - \alpha + (\alpha - x\bar{\varphi}')(1-x)^{q-1} G(p,q,x)}, \tag{2.42}$$

où

$$G(p,q,x) = (1+x)^{-q} \cdot F(p+q-1, q; p+q, \frac{2}{1+x^{-1}}), \quad \text{si } x \in \mathbb{C} \setminus \{-1\}, \tag{2.43}$$

avec $F(.,.;.;.)$ est la fonction hypergéométrique. Deux conditions sont nécessaires afin que $G(p,q,x)$ reste fini pour tout $|x| \leq 1 : p > 1$ et $0 < q < 1$. Puisque $0 < q < 1$, le numérateur diverge. A noter qu'au voisinage de -1, $G(p,q,x)$ se comporte comme $(1-x)^{-q} \cdot F(1, q; p+q, \frac{2}{1-x^{-1}})$.

Étudions le comportement du dénominateur, et découvrons les conditions liées à sa convergence.

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{G(p,q,x)}{(1-\alpha)(1-x)^{1-q} + (\alpha - x\bar{\varphi}')G(p,q,x)} \underset{x \to 1}{\longrightarrow} \frac{1}{\alpha - \bar{\varphi}'}, \tag{2.44}$$

ce qui implique la condition de mémoire longue suivante

$$\bar{\varphi}' = \alpha. \tag{2.45}$$

La dynamique de l'anticipation de l'agent i devient alors

$$\hat{X}_{i,t} = \varphi_i \cdot \hat{X}_{i,t-1} + \alpha(1-\varphi_i) \cdot \bar{X}_{n,t-1} + c_i \cdot \varepsilon_t + \eta_{i,t}, \tag{2.46}$$

pour $t \geq 0$ et $i = 1, 2, \ldots, n$. Quand φ_i est proche de 1, alors $\alpha(1-\varphi_i)$ est proche de 0 et inversement, lorsque φ_i est proche de 0, alors $\alpha(1-\varphi_i)$ est proche de α. Finalement, quand φ est proche de -1, alors $\alpha(1-\varphi_i)$ est proche de 2α. Du point de vue d'un agent, il y a comme un compromis entre le poids qu'il accorde à son anticipation passée et celui qu'il attribue à la volatilité réalisée passée. Plus il accorde d'importance à son anticipation passée, moins la réalisation passée importe. Et inversement. Par ailleurs, quand un agent décide de baser son

anticipation sur l'opposé de son anticipation passée (cas correspondant à $\varphi = -1$)), il peut penser par exemple que la veille il a fait erreur, alors il donnera aussi un poids important à la réalisation passée par l'intermédiaire du coefficient 2α. C'est un comportement qui nous semble raisonnable.

Remplacer $\bar{\varphi}'$ par α dans l'équation (2.42) conduit à l'expression simplifiée

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{(1-x)^{q-1} G(p,q,x)}{1-\alpha+\alpha(1-x)^q G(p,q,x)}. \tag{2.47}$$

Donc, sous l'hypothèse que le dénominateur dans l'équation (2.47) reste borné pour tout $|x| \leq 1$, ce qui requiert

$$\frac{1-p}{q} < \alpha < 1, \tag{2.48}$$

la série $\sum_{k=0}^{\infty} \tilde{\beta}_k x^k$ se comporte comme $(1-x)^{q-1}$ au voisinage de 1 et ainsi diverge de manière hyperbolique si $q < 1$. Dans un tel cas, la fonction d'autocorrélation se comporte comme $K|h|^{1-2q}$ quand $|h| \to \infty$, sous réserve que la condition de stationnarité de second ordre $q > 1/2$ soit vérifiée.

Maintenant, nous allons étudier l'influence des paramètres sur l'allure de la fonction d'autocorrélation. D'abord nous nous intéressons au cas de la mémoire courte où $\bar{\varphi}' \neq \alpha$, puis au cas de mémoire longue où $\bar{\varphi}' = \alpha$.

Mémoire courte

Dans le cas $\bar{\varphi}' \neq \alpha$, la condition pour que le dénominateur ne s'annule pas devient

$$\alpha > \max\left(\bar{\varphi}', \frac{\left(1+\frac{q}{p-1}\right)\bar{\varphi}'+2}{1-\frac{q}{p-1}}\right), \quad \text{si } \frac{q}{p-1} > 1, \tag{2.49}$$

$$\bar{\varphi}' < \alpha < \frac{\left(1+\frac{q}{p-1}\right)\bar{\varphi}'+2}{1-\frac{q}{p-1}}, \quad \text{si } \frac{q}{p-1} < 1. \tag{2.50}$$

Nous fixons $p = 5$, $q = 0.75$ et $\alpha = 0.8$. La figure 2.4 montre de quelle façon la fonction d'autocorrélation se comporte lorsque $\bar{\varphi}'$ tend vers α. Nous observons que plus $\bar{\varphi}'$ est proche de α, plus lente est la décroissance de la fonction d'autocorrélation.

Regardons le cas particulier où $\alpha = 1$, comme pour le cas d'un agent rationnel, mais $\bar{\varphi}' \neq 0$. C'est, dans un sens, la manière la plus simple pour représenter l'écart à la rationalité. L'anticipation d'un agent i s'écrit alors

$$\hat{X}_{i,t} = \varphi_i \cdot \left(\hat{X}_{i,t-1} - \bar{X}_{n,t-1}\right) + \bar{\varphi}' \cdot \bar{X}_{n,t-1} + c_i \cdot \varepsilon_t + \eta_{i,t}, \tag{2.51}$$

avec $t \geq 0$ et $i = 1, 2, \ldots, n$. Cela signifie que l'agent effectue son anticipation sur la base de l'innovation passée $\hat{X}_{i,t-1} - \bar{X}_{n,t-1}$ mais, en plus, fait aussi attention à la volatilité réalisée passée, comme le prouve la présence du terme $\bar{\varphi}' \cdot \bar{X}_{n,t-1}$.

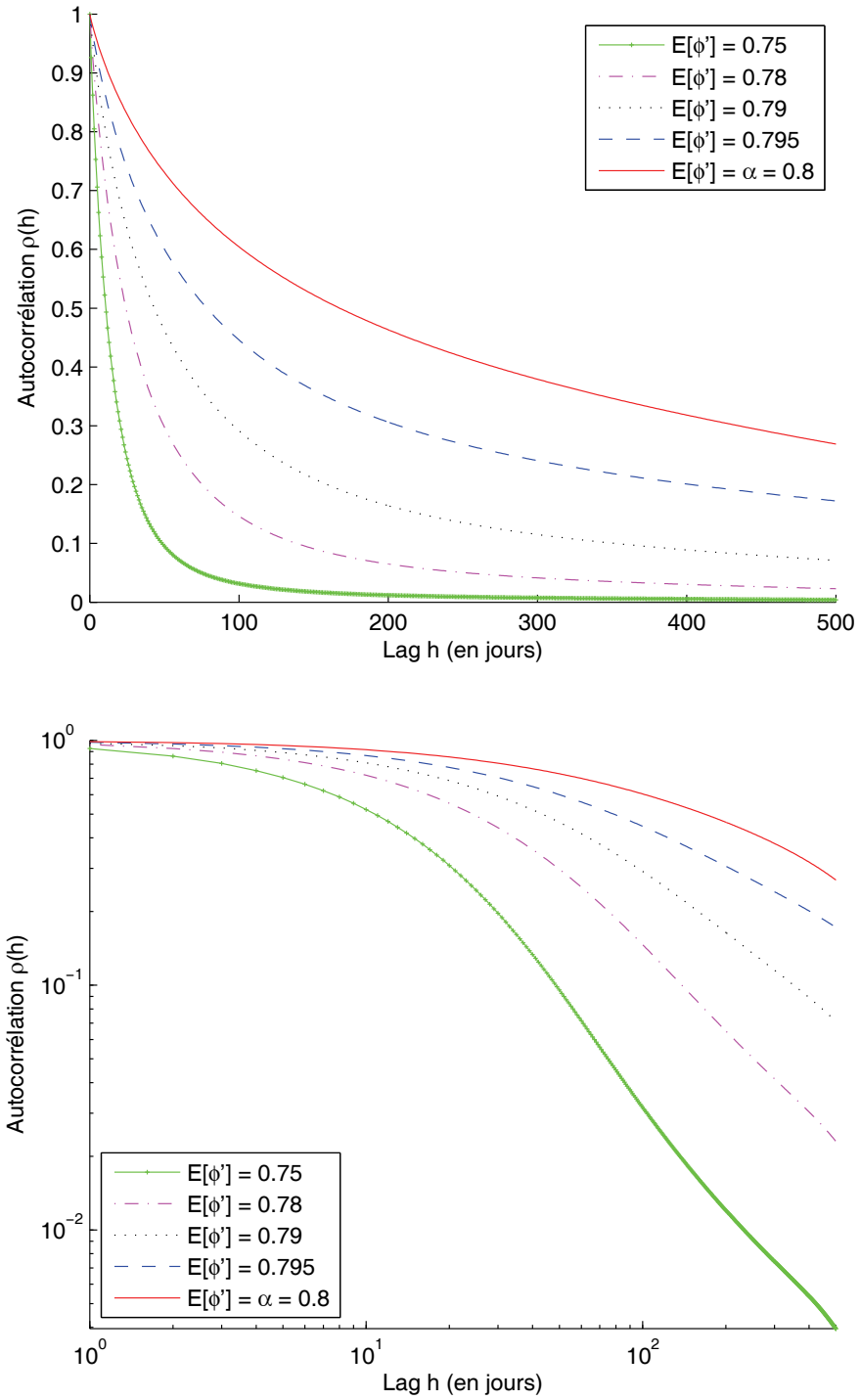

FIGURE 2.4 – Fonctions d'autocorrélation à horizon de 500 jours, obtenues à partir des paramètres $p = 5$, $q = 0.75$, $\alpha = 0.8$ et de différentes valeurs de $\bar{\varphi}'$ comprises dans $[0.75, 0.80]$. Remarque : la notation $\mathrm{E}[\phi]$ dans les légendes vaut pour $\bar{\varphi}'$.

L'équation (2.14) conduit à

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{1}{1 - x \cdot \bar{\varphi}'}, \tag{2.52}$$

ce qui montre que la dynamique de (l'excès de) la log-volatilité réalisée n'est rien d'autre qu'un simple processus AR(1) à mémoire courte

$$\bar{X}_t - \bar{\varphi}' \cdot \bar{X}_{t-1} = \mathrm{E}[c]\varepsilon_t, \tag{2.53}$$

dont la fonction d'autocorrélation est donnée par

$$\rho(h) = \bar{\varphi}'^{|h|} \tag{2.54}$$

Ainsi, si $\bar{\varphi}' = 0$, la volatilité montre exactement le même comportement que dans le cas d'agents complètement rationnels (cas 1), ce qui montre une fois encore que la rationalité au niveau individuel n'est pas une condition nécessaire pour que la volatilité, au niveau agrégé, se comporte comme si les agents étaient individuellement rationnels.

Mémoire longue

A présent, intéressons nous au cas où le modèle affiche de la mémoire longue, concrètement lorsque

$$\bar{\varphi}' = \alpha. \tag{2.55}$$

Nous cherchons à déterminer le rôle des paramètres p, q et α sur la mémoire courte et la mémoire longue.

Les figures 2.5 et 2.6 décrivent le rôle de p. Les deux autres paramètres q et α sont fixés respectivement à 0.75 et 0.3, et p est compris entre 1.05 et 8. Sur la première des deux figures, nous avons tracé les fonctions d'autocorrélation à horizon de 1000 jours en mode linéaire (graphique du haut) et en mode log-log (graphique du bas). Nous observons que plus p et grand, plus la décroissance de la fonction d'autocorrélation est lente sur les premiers lags. Par ailleurs, sur le graphique en mode log-log, les pentes des fonctions d'autocorrélation au delà des 150 premiers jours ne dépendent pas de p. Il semblerait que p n'ait pas d'impact sur la mémoire longue.

Contrairement à ce que l'on pourrait penser en regardant la divergence des densités au voisinage de 1 sur la figure 2.6, celle-ci n'est pas affectée par les valeurs que peuvent prendre p. En effet, seul le paramètre q, via le terme $(1-\varphi)^{q-1}$ contrôle la vitesse de la divergence au voisinage de 1. En résumé la divergence est qualitativement la même au voisinage de 1 et cette illusion d'optique est due à un facteur d'échelle. Il n'en est pas moins vrai que les proportions d'agents auto-référents chutent au profit d'autres catégories telles que les agents rationnels ($\varphi \simeq 0$) ou les agents prenant le contre-pied de leur anticipation précédente ($\varphi \simeq -1$) lorsque se rapproche de 1. En conclusion, plus p influence la mémoire courte dans la mesure où plus p est proche de 1, plus elle est faible.

Le rôle de q est étudié au travers des figures 2.7 et 2.8 où sont tracés respectivement les graphiques des fonctions d'autocorrélation (en mode linéaire sur le graphique du haut,

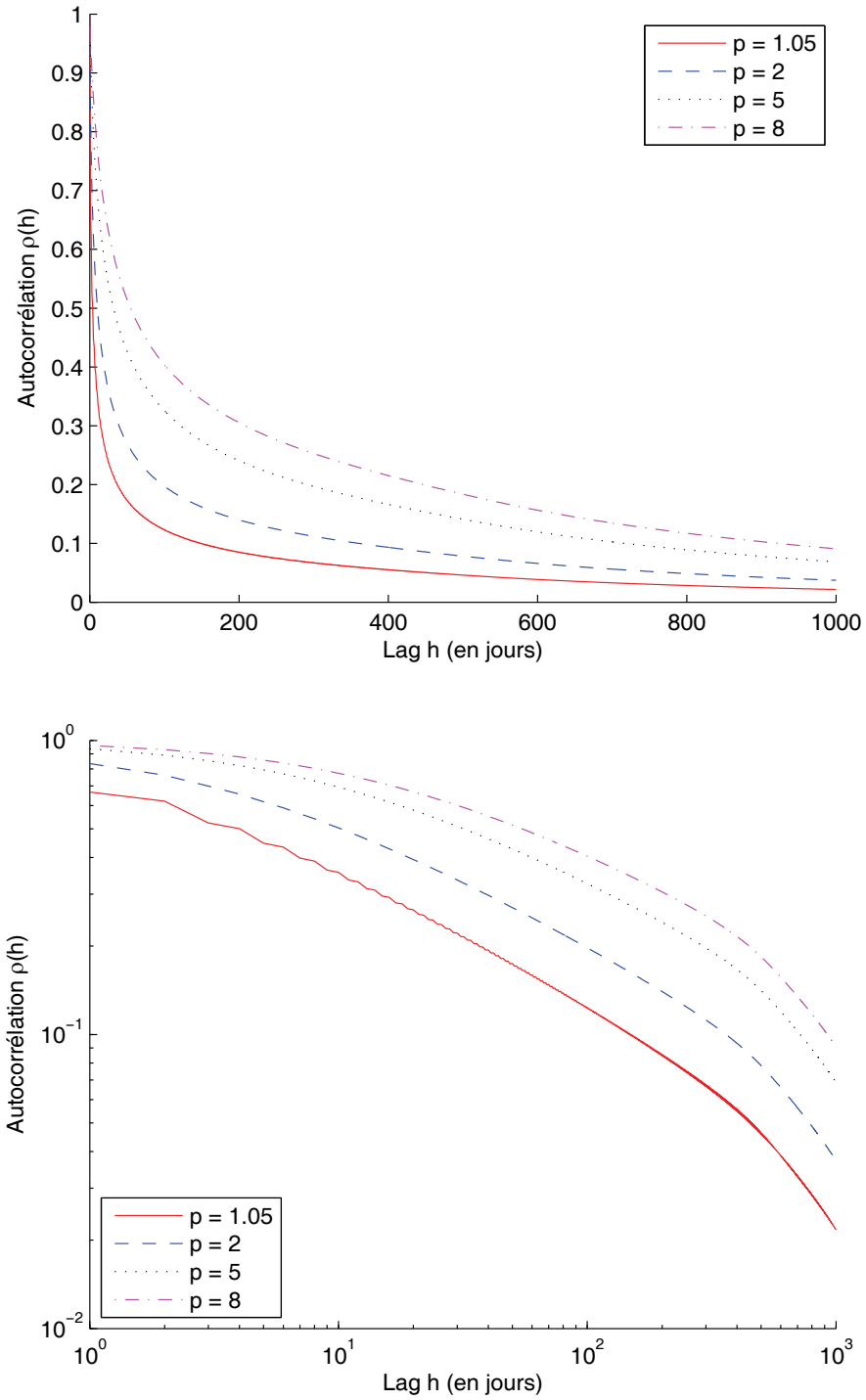

FIGURE 2.5 – Fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) sur 1000 jours obtenues à partir des paramètres $q = 0.75$, $\alpha = 0.3$ et différentes valeurs de p inclues dans $[1.05, 8]$.

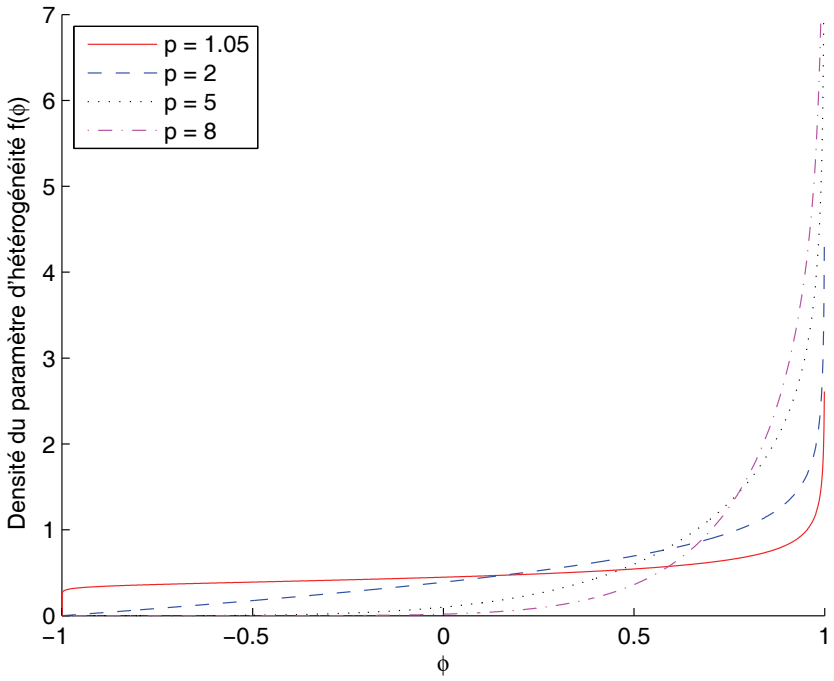

FIGURE 2.6 – Densité du coefficient d'hétérogénéité obtenue à partir du paramètre $q = 0.75$ et de différentes valeurs de p inclues dans $[1.05, 8]$. La loi de φ est une Beta(p,q) étirée sur $[-1, 1]$ et sa densité, égale à $f(\varphi) = \frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1}$, explique la divergence en 0 et 1.

et en mode log-log sur celui du bas) et celui des densités du coefficient d'hétérogénéité. Les deux autres paramètres p et α sont fixés respectivement à 5 et 0.3, tandis que q est compris entre 0.55 et 0.85. En observant la première figure, on remarque sur le graphique en mode linéaire que plus q est proche de 0.5, plus la décroissance de la fonction d'autocorrélation est lente sur les premiers lags, autrement dit, q a un impacte sur la mémoire courte. Par ailleurs, le graphique en mode doublement logarithmique nous permet d'affirmer que ce paramètre a aussi une influence sur la mémoire longue. En effet, au delà du lag 500, les courbes des fonctions d'autocorrélation deviennent des droites. Plus q est proche de 0.5, moins la pente est raide et donc plus le paramètre de mémoire longue est grand. On pouvait logiquement s'attendre à ce résultat étant donné que la divergence au voisinage de 1 est contrôlée par le paramètre q, via le terme $(1-\varphi)^{q-1}$. En conclusion, le paramètre agit de la même manière à la fois sur la mémoire courte et sur la mémoire longue : plus q est proche de 0.5, plus la mémoire courte et la mémoire longue sont grandes.

Quant à la figure 2.8, nous en déduisons que plus q est proche de 0.5, plus la divergence de la fonction de densité au voisinage de 1 est rapide et donc plus le paramètre de mémoire longue est grand. Cela confirme l'analyse de la figure 2.7. Par conséquent, q a une influence sur la mémoire longue.

La figure 2.9 représentant les fonctions d'autocorrélation de la log-volatilité réalisée (en

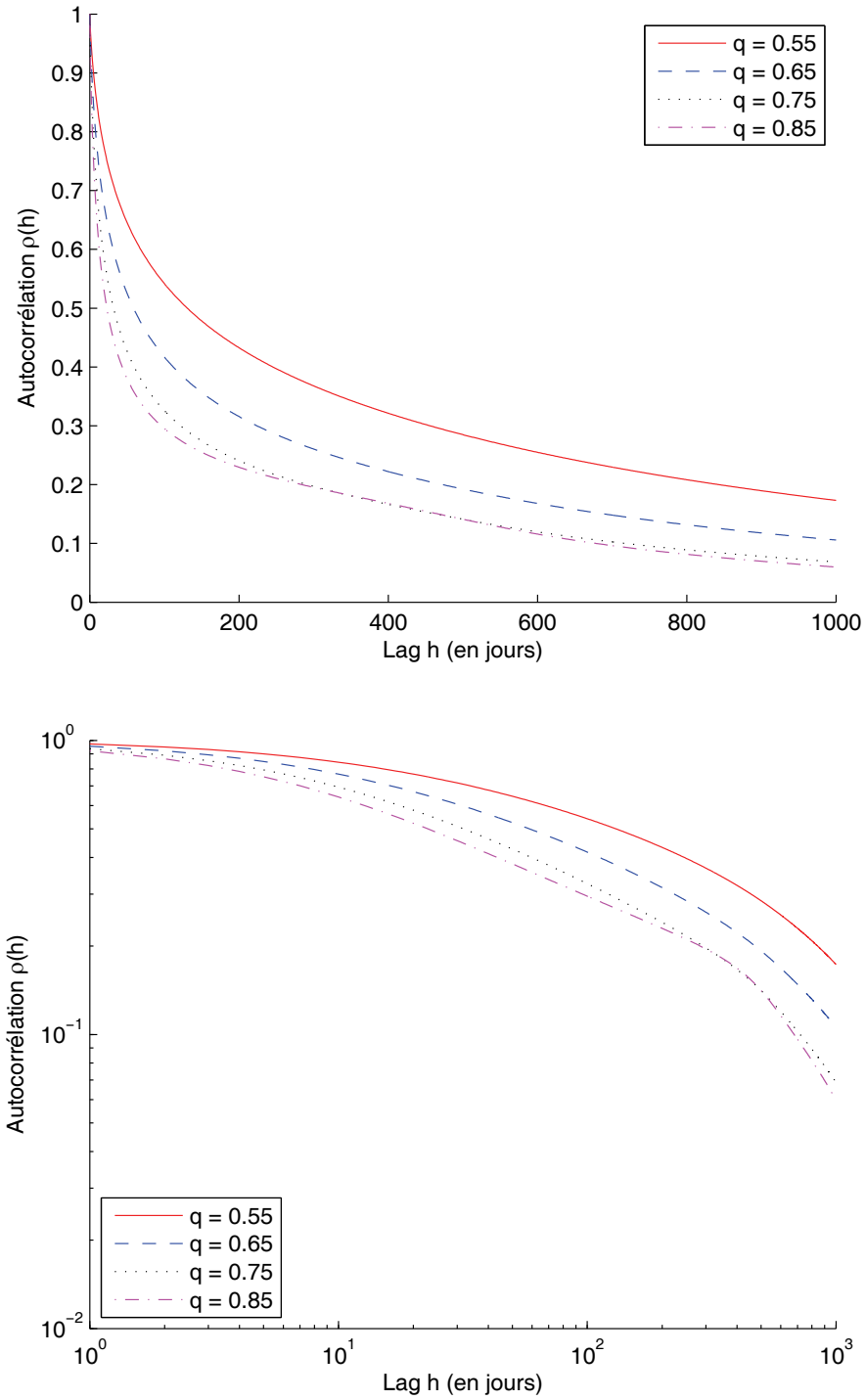

FIGURE 2.7 – Fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) sur 1000 jours obtenues à partir des paramètres $p = 5$, $\alpha = 0.3$ et différentes valeurs de q inclues dans $[0.55, 0.85]$.

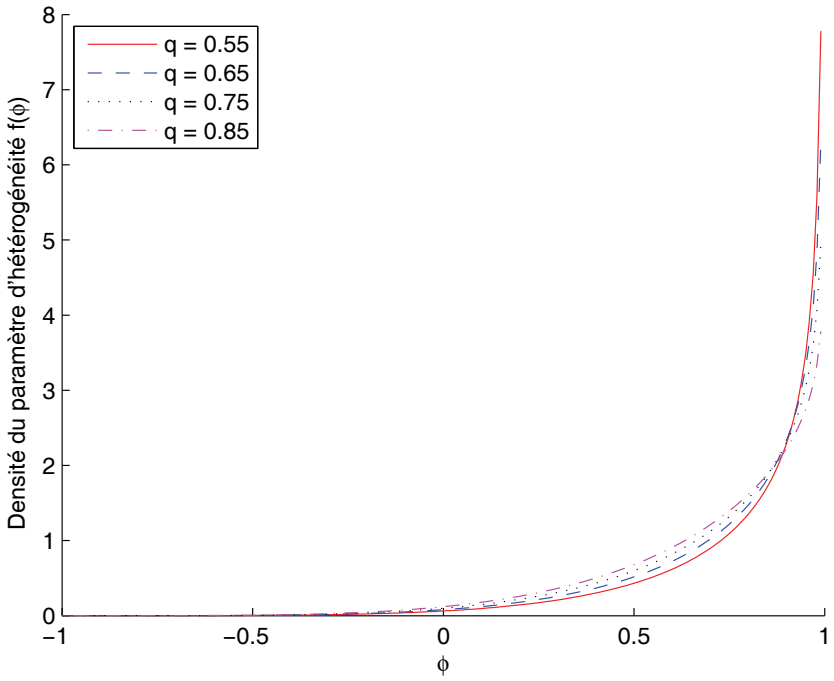

FIGURE 2.8 – Densité du coefficient d'hétérogénéité obtenue à partir du paramètre $p = 5$ et de différentes valeurs de q inclues dans $[0.55, 0.85]$. La loi de φ est une Beta(p,q) étirée sur $[-1, 1]$ et sa densité, égale à $f(\varphi) = \frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1}$, explique la divergence en 0 et 1.

mode linéaire sur le graphique du haut et en mode log-log sur le graphique du bas) montre le rôle joué par α. Nous avons fixé les deux autres paramètres p et q respectivement à 5 et 0.75 et fait prendre à α différentes valeurs comprises entre 0.1 et 0.9. Nous remarquons que α a une grande influence sur la mémoire courte. En effet, plus α est petit, plus la décroissance durant les premiers lags est rapide. Nous n'avons pas tracé la fonction de la densité du paramètre d'hétérogénéité car α n'intervenant pas, nous aurions obtenu une unique courbe et cela n'aurait rien apporté.

2.4.3 Extension de l'écart à la rationalité

L'inconvénient, lorsque l'on analyse le comportement des agents au travers de la densité du paramètre d'hétérogénéité, réside dans le fait qu'on ne balaie pas un éventail assez large de comportements. En effet, il est raisonnable de penser trouver un nombre conséquent d'agents se basant sur les informations, ce qui se traduirait par une bosse autour de $\varphi = 0$. De même, il se pourrait que des agents, pensant s'être trompé la veille, décident de prendre le contre-pied de leur anticipation passée, et dans ce cas, on observerait une bosse aux alentours de $\varphi = -1$. Or, parmi les figures précédentes, pour que la densité du paramètre d'hétérogénéité ne soit pas nul aux alentours de $\varphi = -1$, il n'y a qu'une solution : $p = 1.05$ (voir figure 2.6).

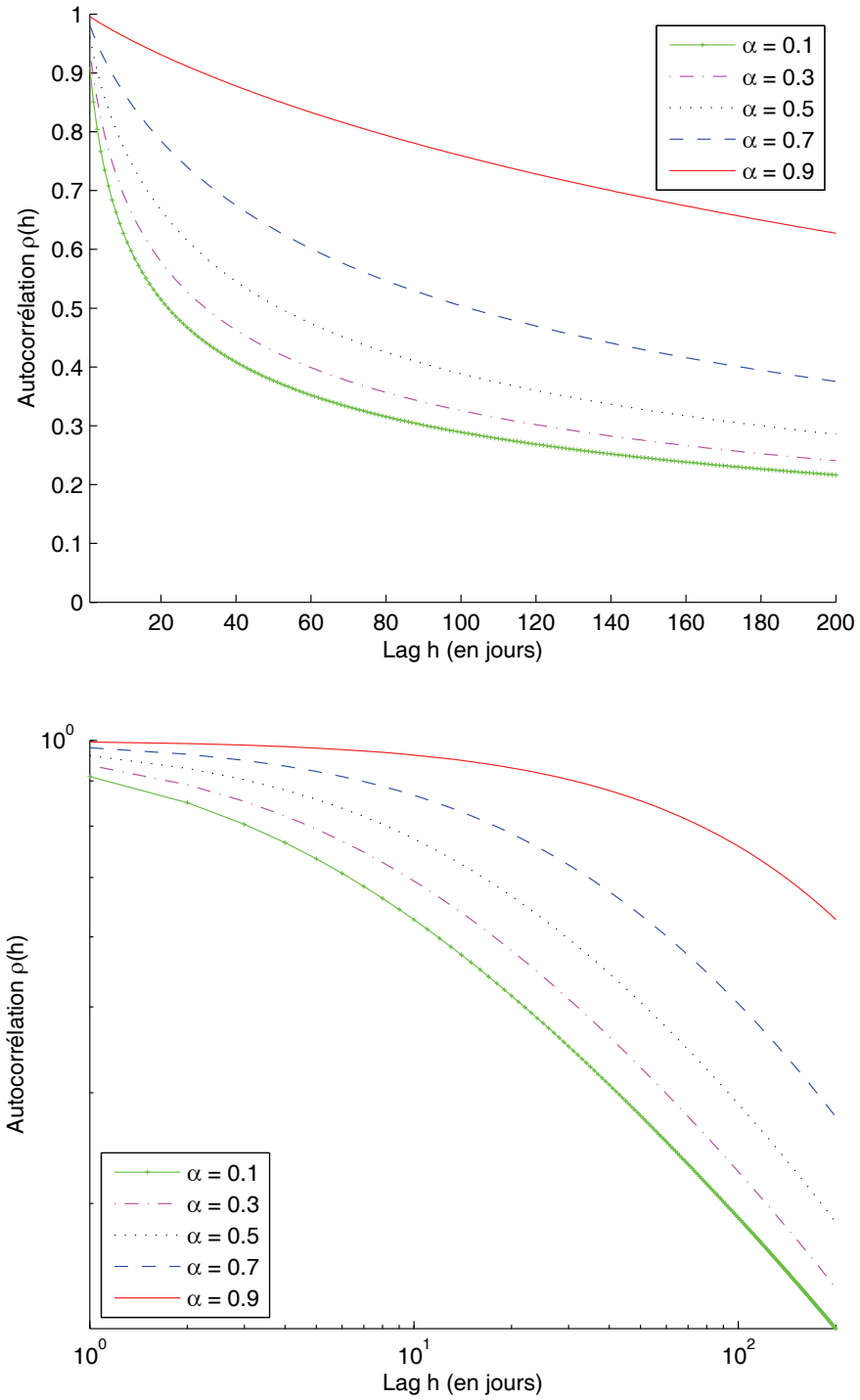

FIGURE 2.9 – Fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) sur 200 jours obtenues à partir des paramètres $p = 5$, $q = 0.75$ et différentes valeurs de α inclues dans $[0.1, 0.9]$.

Le problème est double. D'une part, cette quantité sera toujours faible, et d'autre part, cela impose que la densité soit quasiment la même aux alentours de $\varphi = 0$.

En résumé, cette modélisation manque de souplesse et n'est pas suffisamment aboutie pour prendre en compte la diversité des comportements des investisseurs à laquelle on peut raisonnablement s'attendre. C'est pourquoi, nous allons modifier la fonction de densité du paramètre d'hétérogénéité, en continuant bien sûr à considérer le cas $\bar{\varphi}' = \alpha$ pour qu'il y ait de la mémoire longue. Pour atteindre cet objectif, nous opérons la modification suivante : nous conservons la loi Beta étirée sur $[-1, 1]$, mais y ajoutons une autre loi dont la densité est en forme de cloche. Cette cloche aura la possibilité de se translater sur tout l'intervalle [-1,1]. Par exemple, si la cloche est fixée autour de 0, cela signifie qu'un grand nombre d'agents n'accordent pas d'importance à leur anticipation passée.

$$f(\varphi) = w\frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1} + (1-w)\frac{1}{K(m,\sigma)}(1+\varphi)(1-\varphi)\exp\left(-\frac{1}{2}\frac{(\varphi-m)^2}{\sigma^2}\right), \quad (2.56)$$

où K est la constante de normalisation (qui peut être exprimée sous une forme fermée) de la loi et w dans $[0, 1]$ est le poids donné à la loi Beta étirée. Les conditions de mémoire longue demeurent identiques, seule la condition de stationnarité change. Nous pourrions les exprimer aisément dans une forme fermée, mais puisque leur expression est vraiment lourde, nous ne les donnerons pas.

Etudions le rôle de w. Pour y parvenir, les paramètres p et q sont fixés respectivement à 5 et 0.75, $\alpha = 0.3$, $m = 0$, $\sigma = 0.2$ et w prends différentes valeurs inclues dans $[1/4, 3/4]$. Nous commençons par tracer sur la figure 2.10 les fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas). Nous observons sur le graphique du haut que le poids accordé à chaque densité a un impact sur la mémoire courte et ne semble pas en avoir sur la mémoire longue. Le graphique du bas, confirme cette première impression. En effet, à l'oeil nu, les pentes des différentes fonctions d'autocorrélation sont identiques. Pour être plus précis au sujet du rôle de w sur la mémoire courte, plus w est faible, autrement dit moins on accorde d'importance à la loi Beta, plus la décroissance durant les premiers lags est rapide et donc plus la mémoire courte et faible.

En ce qui concerne la densité du paramètre d'hétérogénéité pour ce jeu de paramètres, la figure 2.11 nous informe que plus w est petit, plus la proportion d'investisseurs se basant sur leur anticipation passée et négligeant les informations ou la réalisation passée est faible. Quant à la vitesse de la divergence de la densité au voisinage de 1, comme dans le cas de p, elle n'est nullement affectée par w et la différence provient d'un facteur d'échelle. En conclusion, le poids w accordé à la loi Beta a un impact uniquement sur la mémoire courte.

La figure 2.12 montre les fonctions d'autocorélation de la log-volatité réalisée pour des paramètres p et q fixés respectivement à 5 et 0.75, $\alpha = 0.3$, $m = 0$, $w = 1/2$ et plusieurs valeurs de σ comprises entre 0.1 et 0.3. Le graphique en mode linéaire (en haut) nous indique que σ un un faible impact sur la mémoire courte : plus sigma est petit, plus la mémoire courte est grande. l ne semble pas que ce paramètre ait d'impact sur la mémoire longue. C'est ce que confirme le graphique en mode log-log (en bas) où l'on observe que la pente des trois

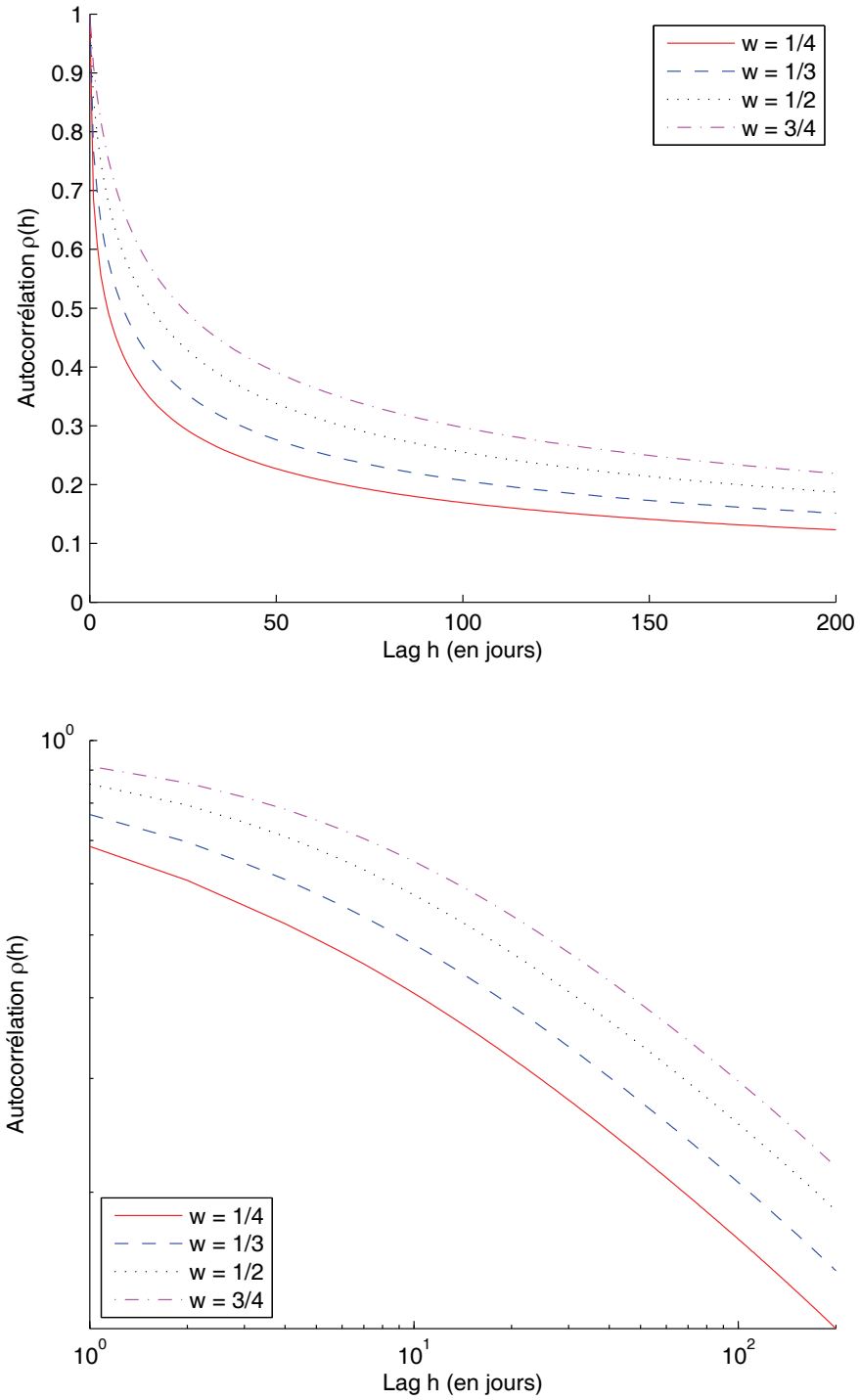

FIGURE 2.10 – Fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) sur 200 jours obtenues à partir des paramètres $p = 5$, $q = 0.75$, $\alpha = 0.3$, $m = 0$, $\sigma = 0.2$ et différentes valeurs de w inclues dans $[1/4, 3/4]$.

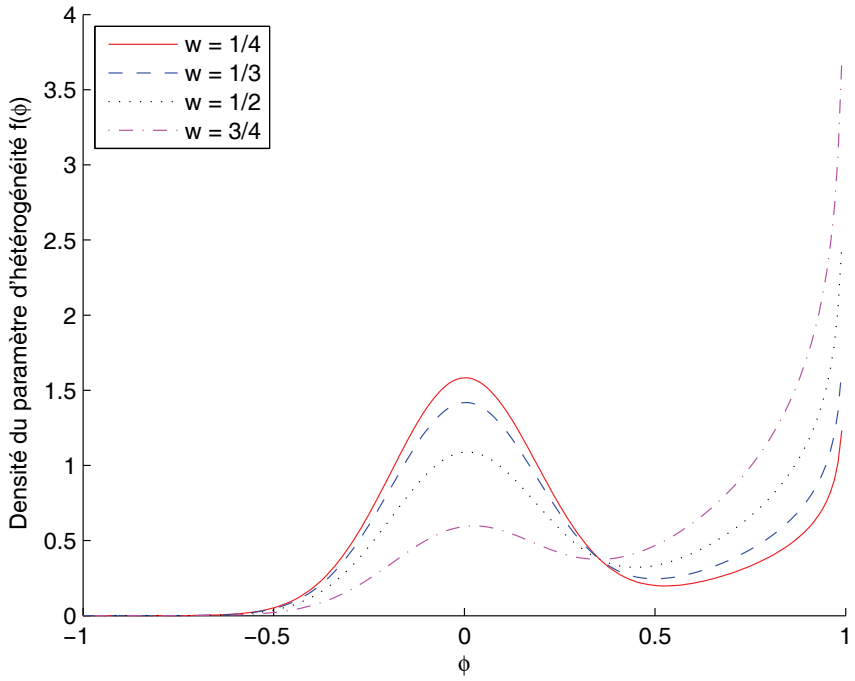

FIGURE 2.11 – Densité du coefficient d'hétérogénéité obtenue à partir du paramètre $p = 5$, $q = 0.75$, $\alpha = 0.3$, $m = 0$, $\sigma = 0.2$ et différentes valeurs de w inclues dans $[1/4, 3/4]$. La loi de φ est une Beta(p,q) étirée sur $[-1, 1]$ et sa densité, égale à $f(\varphi) = w\frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1} + (1-w)\frac{1}{K}(1+\varphi)(1-\varphi)\exp\left(-\frac{1}{2}\frac{(\varphi-m)^2}{\sigma^2}\right)$, explique la divergence en 0 et 1.

fonctions d'autocorrélation est identique.

La figure 2.13 montrant la densité du paramètre d'hétérogénéité pour ces mêmes autres paramètres confirme que σ ne contrôle que la largeur et la hauteur de la bosse. La divergence de la densité à l'approche de 1 n'est absolument pas modifiée. Cela nous conforte dans notre premier jugement concernant son impact nul sur la mémoire longue. Par ailleurs, nous constatons que plus σ est faible, plus la bosse est étroite et haute, ce qui signifie que le nombre d'investisseur se basant sur les informations et la réalisation passée de la log-volatilité est plus grand. Cela engendre une faible augmentation de la mémoire courte.

Il reste à traiter le rôle de m, paramètre qui détermine la position de la bosse sur tout l'intervalle $[-1, 1]$. Nous avons fixé les paramètres p et q respectivement à 5 et 0.75, $\alpha = 0.3$, $\sigma = 0.2$, $w = 1/3$. Puis nous avons tracé les fonctions d'autocorrélation de la log-volatilité réalisée pour plusieurs valeurs de m comprises dans $[-0.8, 0.8]$ (voir figure 2.14). D'après le graphique en mode linéaire, on peut affirmer que la position de la cloche a une influence sur la mémoire courte. Durant les premiers lags, nous observons que plus m est proche de -1, c'est à dire plus la proportion d'agents se basant sur l'opposée de leur anticipation passée combinée à la réalisation passée est importante, plus plus la décroissance de la fonction d'autocorrélation sur les premiers lags est lente, donc plus il y a de mémoire courte. Mais après quelques

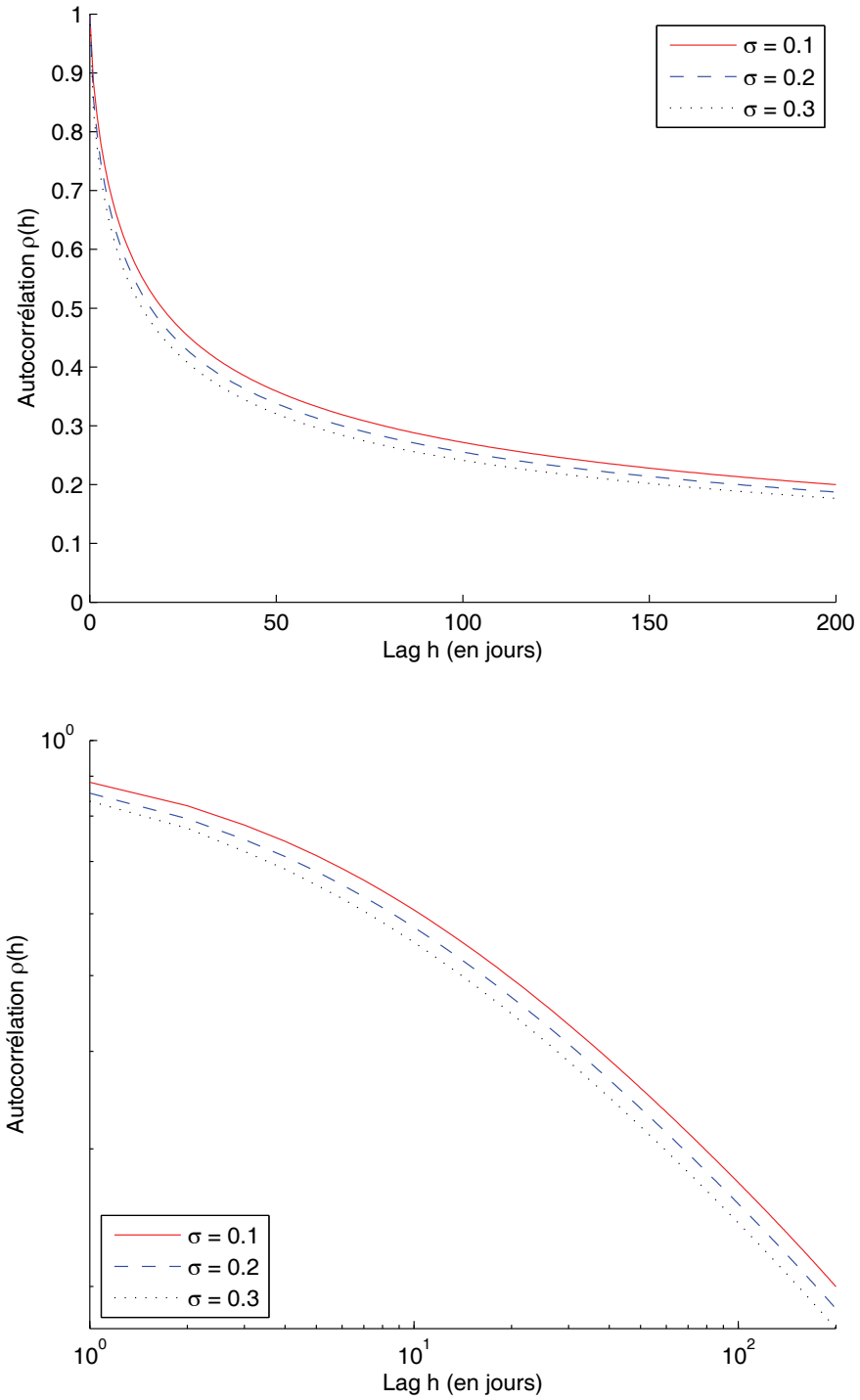

FIGURE 2.12 – Fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) sur 200 jours obtenues à partir des paramètres $p = 5$, $q = 0.75$, $\alpha = 0.3$, $m = 0$, $m = 1/2$ et différentes valeurs de σ inclues dans $[0.1, 0.3]$.

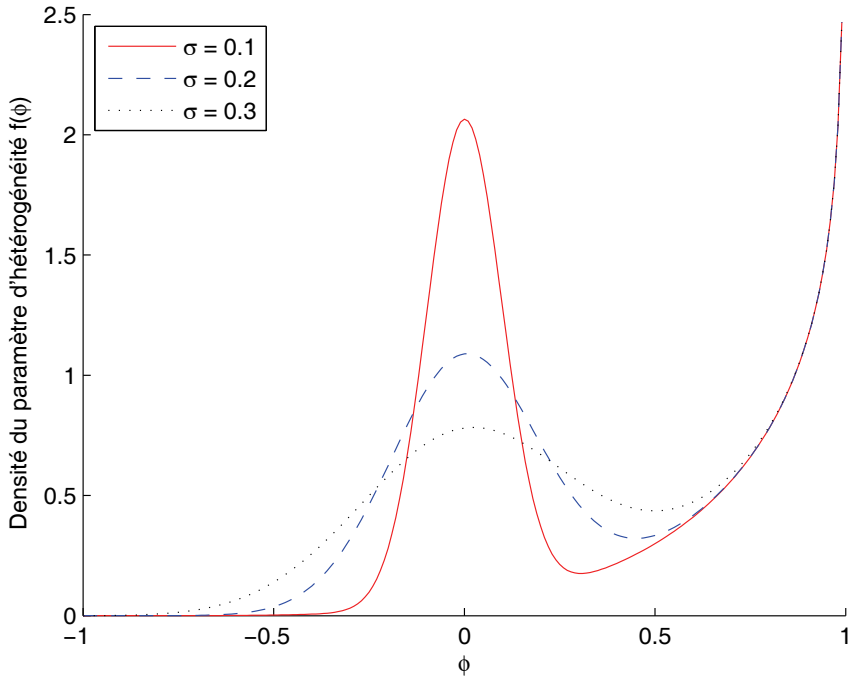

FIGURE 2.13 – Densité du coefficient d'hétérogénéité obtenue à partir du paramètre $p = 5$, $q = 0.75$, $\alpha = 0.3$, $m = 0$, $w = 1/2$ et différentes valeurs de σ inclues dans $[0.1, 0.3]$. La loi de φ est une Beta(p,q) étirée sur $[-1, 1]$ et sa densité, égale à $f(\varphi) = w\frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1} + (1-w)\frac{1}{K}(1+\varphi)(1-\varphi)\exp\left(-\frac{1}{2}\frac{(\varphi-m)^2}{\sigma^2}\right)$, explique la divergence en 0 et 1.

une cinquantaine de lags, cette situation n'est plus. Le graphique en mode log-log montre qu'après 100 jours, les pentes des trois fonctions d'autocorrélation sont identiques et que pas conséquent la valeur de m n'affecte pas la mémoire longue.

Observons à présent la densité du paramètre d'hétérogénéité pour ces mêmes jeux de paramètres (voir figure 2.15). La vitesse de la divergence de la densité au voisinage de 1 n'est en aucun cas modifiée par m. Par conséquent, on s'attend à ce que m n'ait pas d'impact dur la mémoire longue, même si lorsque $m = 0.8$, la proportion d'investisseurs accordant de l'importance à leur anticipation passée et négligeant la réalisation passée augmente. Or, nous avons établi que ce comportement auto-référent contribue à la présence de mémoire longue. C'est sans doute la raison pour laquelle on observe sur le graphique du bas de la figure que contrairement aux fonctions d'autocorélation pour $m = -0.8$ et $m = 0$, qui au bout de 20 jours ont une allure similaire, il faut attendre une centaine de jours pour que ce soit le cas pour celle où $m = 0.8$. Entre 20 et 100 jours, la pente de cette fonction d'autocorrélation est plus forte, ce qui indique que le paramètre de mémoire longue est plus grand.

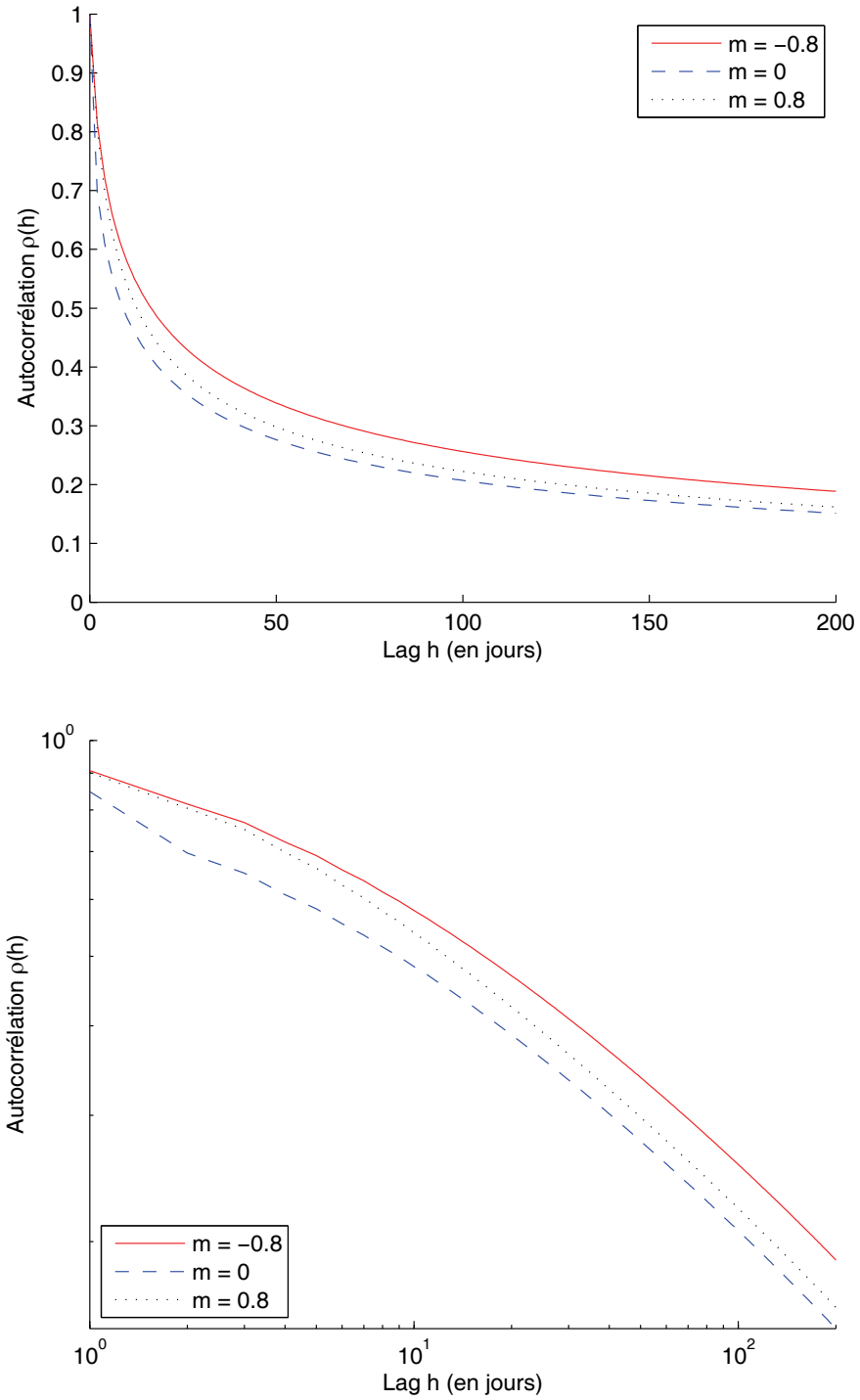

FIGURE 2.14 – Fonctions d'autocorrélation en mode linéaire (graphique du haut) et en mode log-log (graphique du bas) sur 200 jours obtenues à partir des paramètres $p = 5$, $q = 0.75$, $\alpha = 0.3$, $\sigma = 0.2$, $w = 1/3$ et différentes valeurs de m inclues dans $[-0.8, 0.8]$.

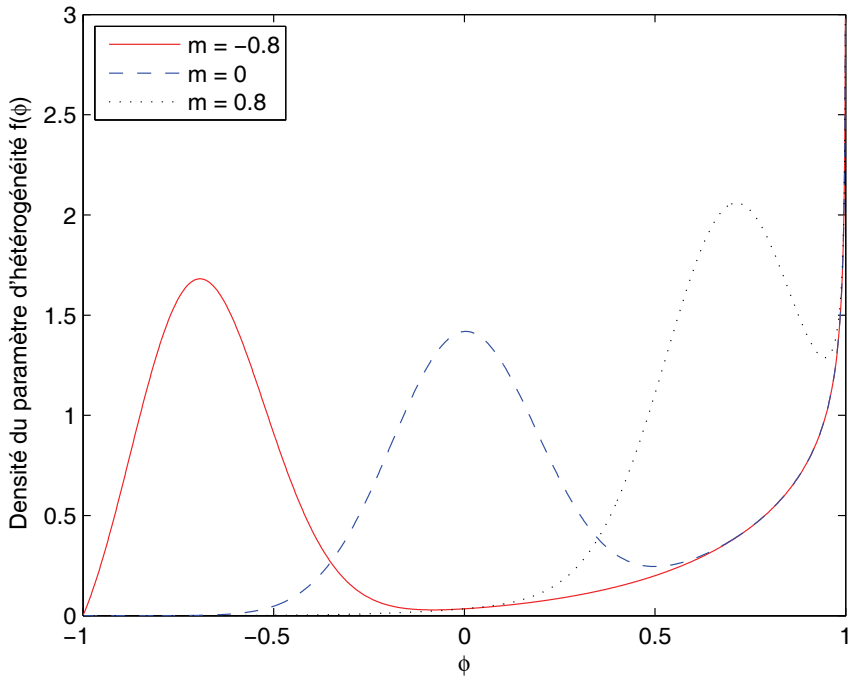

FIGURE 2.15 – Densité du coefficient d'hétérogénéité obtenue à partir du paramètre $p = 5$, $q = 0.75$, $\alpha = 0.3$, $\sigma = 0.2$, $w = 1/3$ et différentes valeurs de m inclues dans $[-0.8, 0.8]$. La loi de φ est une Beta(p,q) étirée sur $[-1, 1]$ et sa densité, égale à $f(\varphi) = w\frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1} + (1-w)\frac{1}{K}(1+\varphi)(1-\varphi)\exp\left(-\frac{1}{2}\frac{(\varphi-m)^2}{\sigma^2}\right)$, explique la divergence en 0 et 1.

Conclusion

En résumé, le modèle de (log-)volatilité réalisée que nous avons proposé est basé sur les hypothèses suivantes :

- Le processus de log-volatilité suit un processus stochastique gaussien,
- Le marché effectue une agrégation des anticipations des investisseurs,
- La log-volatilité réalisée est la moyenne de toutes les anticipations individuelles de la log-volatilité.

De plus, nous supposons que l'anticipation du futur niveau de l'excès de log-volatilité sur sa moyenne à long terme de l'investisseur i le jour t, noté $\hat{X}_{i,t}$, dépend de quatre facteurs, à savoir

- son anticipation de la veille, $\hat{X}_{i,t-1}$,
- la réalisation de l'excès de log-volatilité réalisée de la veille, $\bar{X}_{t-1,n}$,
- une information publique, ε_t,
- et une information privée, $\eta_{i,t}$.

Cela se traduit par l'équation récursive

$$\hat{X}_{i,t} = \varphi_i \cdot \hat{X}_{i,t-1} + \psi_i \cdot \bar{X}_{n,t-1} + c_i \cdot \varepsilon_t + \eta_{i,t}, \tag{2.57}$$

pour $t \geq 0$ et $i = 1, \ldots, n$. Nous avons établi que le paramètre dit d'hétérogénéité φ est responsable de la mémoire longue. La condition nécessaire pour avoir de la mémoire longue est que sa densité diverge au voisinage de 1 (Gonçalves et Gouriéroux 1988). Plus la divergence est rapide, plus grande est le nombre d'agents auto-référents ce qui implique une plus grande hétérogénéité de comportements et plus la mémoire longue est importante. Quant au paramètre ψ, il contrôle la mémoire courte.

Nous avons également montré que pour un agent i rationnel, nous devions poser $\psi_i = -\varphi_i$. Or, nous considérons que les acteurs du marché sont à rationalité limité, en d'autres termes que leur comportement a tendance à s'écarter d'un comportement rationnel. C'est pourquoi, nous avons finalement posé

$$\psi_i = -\alpha\varphi + \bar{\varphi}'. \tag{2.58}$$

La condition nécessaire et suffisante pour qu'il y ait mémoire longue est de poser

$$\bar{\varphi}' = \alpha. \tag{2.59}$$

Le modèle s'écrit donc

$$\hat{X}_{i,t} = \varphi_i \cdot \hat{X}_{i,t-1} + \alpha(1 - \varphi_i) \cdot \bar{X}_{n,t-1} + c_i \cdot \varepsilon_t + \eta_{i,t}, \tag{2.60}$$

pour $t \geq 0$ et $i = 1, \ldots, n$. Le paramètre d'hétérogénéité est en relation directe avec ce qui se passe sur le marché :

- si $\varphi \approx 1$, l'impact de l'anticipation précédente est très importante et à moins qu'une information de grande ampleur n'intervienne, la valeur de l'excès de volatilité escompté sera très proche de celle de la veille. En somme, il croit à la continuité des conditions du marché concernant la volatilité,
- si $\varphi \approx -1$, l'impact de l'anticipation précédente est toujours plus importante que les informations, mais cette fois-ci, l'anticipation de l'investisseur devient l'opposée de celle de la veille. La réalisation de la log-volatilité de la veille peut aussi entrer en ligne de compte à condition que 2α soit suffisamment grand. Si c'est le cas, cela peut signifier qu'il s'est trompé dans son anticipation précédente et qu'il la rectifie en se basant à la fois sur la réalisation passée et l'opposée de son anticipation passée. Par contre si $2\alpha \approx 0$, l'investisseur aurait tendance à croire à un changement du niveau de la volatilité,
- si $|\varphi| \approx 0$, l'investisseur ne tient pas compte de son anticipation précédente et repose sur les informations et la réalisation passée.

Pour que, d'une part, la densité du paramètre d'hétérogénéité φ puisse diverger en 1, et que d'autre part, cette densité puisse traduire la diversité des comportements des investisseurs, nous l'avons définie comme étant une somme pondérée d'une loi Beta étirée sur $[-1, 1]$ et d'une loi en forme de cloche également définie sur $[-1, 1]$:

$$f(\varphi) = w\frac{1}{2^{p+q-1}B(p,q)}(1+\varphi)^{p-1}(1-\varphi)^{q-1} + (1-w)\frac{1}{K(m,\sigma)}(1+\varphi)(1-\varphi)\exp\left(-\frac{1}{2}\frac{(\varphi-m)^2}{\sigma^2}\right), \tag{2.61}$$

où,

- seul le paramètre q, coefficient de la loi Beta qui contrôle la divergence au voisinage de 1 (pour cela il doit vérifier $0 < q < 1$), a un impact sur la mémoire longue. Il joue aussi un rôle sur la mémoire courte. Plus q est proche de 0.55, plus la mémoire courte et la mémoire longue sont grandes.
- tous les autres paramètres ont un impact plus ou moins grand sur la mémoire courte :
  - α, la valeur accordé à $\bar{\varphi}$, a une très grande influence,
  - p, l'autre coefficient de la loi Beta qui, pour que le problème soit correctement défini, doit vérifier $p > 1$ a une grande influence sur la mémoire courte.
  - w, le poids accordé à la loi Beta étirée sur $[-1, 1]$, a un impact moyen sur la mémoire courte,
  - σ, la largeur du pic de la bosse n'a qu'un impact faible sur la mémoire courte,
  - m, la position de la bosse, a également un impact faible sur la mémoire courte.

Annexes

2.A Démonstration de la proposition 1

Sous l'hypothèse $||A|| < 1$, la solution stationnaire de l'équation

$$\hat{X}_t = A\hat{X}_{t-1} + \varepsilon_t \cdot C + \eta_t, \tag{2.A.1}$$

est donnée par

$$\hat{X}_t = \sum_{k=0}^{\infty} A^k C \varepsilon_{t-k} + \sum_{k=0}^{\infty} A^k \eta_{t-k} \tag{2.A.2}$$

et l'excès de log-volatilité réalisée s'écrit

$$\bar{X}_{n,t} = \sum_{k=0}^{\infty} \left(\frac{1}{n} 1_n' A^k C \right) \varepsilon_{t-k} + \sum_{k=0}^{\infty} \frac{1}{n} 1_n' A^k \eta_{t-k}. \tag{2.A.3}$$

Tenant compte du fait que $A = D + \frac{1}{n}\Psi \cdot 1_n'$ (où $D = \text{diag}(\varphi_1, \ldots, \varphi_n)$), A^k est solution de

$$A^k = A^{k-1} \left(D + \frac{1}{n} \Psi \cdot 1_n' \right), \tag{2.A.4}$$

de telle sorte que l'on doit résoudre l'équation de récurrence

$$1_n' A^k = \left(1_n' A^{k-1} \right) D + \beta_{k-1} 1_n' \tag{2.A.5}$$

où $\beta_k = \frac{1}{n} 1_n A^k \psi$. Il est alors question de manipulations algébriques simples pour montrer que

$$\begin{aligned}
\frac{1}{n} 1_n' A^k C &= \frac{1}{n} 1_n' D^k C + \frac{1}{n} \sum_{i=1}^{k} \beta_{i-1} \cdot 1_n' D^{k-i} C, && (2.A.6)\\
&\underset{n\to\infty}{=} \mathrm{E}[c] \cdot \mathrm{E}\left[\varphi^k\right] + \mathrm{E}[c] \sum_{i=1}^{k} \beta_{i-1} \mathrm{E}\left[\varphi^{k-i}\right] \quad \text{(p.s)}, && (2.A.7)
\end{aligned}$$

où

$$\begin{aligned}
\beta_k &= \frac{1}{n} 1_n' D^k \Psi + \sum_{i=1}^{k} \beta_{i-1} \cdot \left(\frac{1}{n} 1_n' D^{k-i} \Psi \right), && (2.A.8)\\
&\underset{n\to\infty}{=} \mathrm{E}\left[\psi \varphi^k\right] + \sum_{i=1}^{k} \beta_{i-1} \cdot \mathrm{E}\left[\psi \varphi^{k-i}\right] \quad \text{(p.s)}, && (2.A.9)
\end{aligned}$$

alors que

$$\begin{aligned}
\frac{1}{n} 1_n' A^k \eta_{t_k} &= \frac{1}{n} 1_n' D^k \eta_{t_k} + \frac{1}{n} \sum_{i=1}^{k} \beta_{i-1} \cdot 1_n' D^{k-i} \eta_{t_k}, && (2.A.10)\\
&\underset{n\to\infty}{=} \mathrm{E}\left[\varphi^k\right] \mathrm{E}[\eta_t] + \sum_{i=1}^{k} \beta_{i-1} \mathrm{E}\left[\varphi^{k-i}\right] \mathrm{E}[\eta_t] \quad \text{(p.s)}, && (2.A.11)\\
&\underset{n\to\infty}{=} 0, && (2.A.12)
\end{aligned}$$

étant donné que $\mathrm{E}[\eta_t] = 0$.

Donc, $\bar{X}_t = \lim_{n\to\infty} \bar{X}_{n,t}$ est égal à

$$\bar{X}_t = \mathrm{E}[c] \cdot \sum_{k=0}^{\infty} \left(\mathrm{E}\left[\varphi^k\right] + \sum_{i=1}^{k} \beta_{i-1} \mathrm{E}\left[\varphi^{k-i}\right] \right) \varepsilon_{t-k}. \tag{2.A.13}$$

Simplifions cette expression en fixant

$$\tilde{\beta}_k = \mathrm{E}\left[\varphi^k\right] + \sum_{i=1}^{k} \beta_{i-1} \mathrm{E}\left[\varphi^{k-i}\right], \tag{2.A.14}$$

de sorte que

$$\bar{X}_t = \mathrm{E}[c] \cdot \sum_{k=0}^{\infty} \tilde{\beta}_k \varepsilon_{t-k}. \tag{2.A.15}$$

A présent, en multipliant l'équation (2.A.14) par x^k et en sommant sur k de zéro à l'infini, on obtient, pour tout x tel que les sommes ci-dessous existent

$$\sum_{k=0}^{\infty} \tilde{\beta}_{k+1} x^{k+1} = \sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^{k+1}\right] x^{k+1} + \sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} \beta_i \mathrm{E}\left[\varphi^{k-i}\right] \right) x^{k+1}. \tag{2.A.16}$$

Comme il est bien connu que

$$\sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} \beta_i \mathrm{E}\left[\varphi^{k-i}\right] \right) x^{k+1} = x \left(\sum_{k=0}^{\infty} \beta_{k+1} x^k \right) \left(\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] x^k \right), \tag{2.A.17}$$

on obtient

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] x^k \left\{ 1 + x \sum_{k=0}^{\infty} \beta_k x^k \right\}. \tag{2.A.18}$$

Ensuite, en se concentrant sur $\sum_{k=0}^{\infty} \beta_k x^k$ et en utilisant l'équation (2.A.9), le même argument entraîne

$$\sum_{k=0}^{\infty} \beta_{k+1} x^{k+1} = \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^{k+1}\right] x^{k+1} + \sum_{k=0}^{\infty} \left(\sum_{i=0}^{k} \beta_i \mathrm{E}\left[\psi\varphi^{k-i}\right] \right) x^{k+1}, \tag{2.A.19}$$

$$\sum_{k=0}^{\infty} \beta_k x^k = \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] x^k + x \left(\sum_{i=0}^{\infty} \beta_i x^i \right) \left(\sum_{i=0}^{\infty} \mathrm{E}\left[\psi\varphi^i\right] x^i \right). \tag{2.A.20}$$

En d'autres termes

$$\sum_{k=0}^{\infty} \beta_k x^k = \frac{\sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] x^k}{1 - x \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] x^k}, \tag{2.A.21}$$

qui, en remplaçant dans l'équation (2.A.18), mène à

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{\sum_{k=0}^{\infty} \mathrm{E}\left[\varphi^k\right] x^k}{1 - x \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\varphi^k\right] x^k} \tag{2.A.22}$$

et pour autant que la permutation de l'espérance et des signes somme soit autorisée

$$\sum_{k=0}^{\infty} \tilde{\beta}_k x^k = \frac{\mathrm{E}\left[\dfrac{1}{1 - x \cdot \varphi}\right]}{\mathrm{E}\left[1 - \dfrac{\psi}{1 - x \cdot \varphi}\right]}, \tag{2.A.23}$$

Ce qui conclue la démonstration. □

2.B Démonstration de la proposition 2

L'expression de la densité spectrale de X_t découle de la proposition 1 (Brockwell et Davis 1990) et s'exprime

$$f_X(\lambda) = \frac{\mathrm{E}\left[c\right]^2 \cdot \sigma_\varepsilon^2}{2\pi} \left| \frac{\sum_{k=0}^{\infty} \mathrm{E}\left[\phi^k\right] e^{-ik\lambda}}{1 - e^{-i\lambda} \sum_{k=0}^{\infty} \mathrm{E}\left[\psi\phi^k\right] e^{-ik\lambda}} \right|^2 . \tag{2.B.1}$$

Pour autant que la permutation entre l'espérance et le signe somme soit autorisée, nous pouvons récrire cette relation ainsi :

$$f_X(\lambda) = \frac{\mathrm{E}\left[c\right]^2 \cdot \sigma_\varepsilon^2}{2\pi} \left| \frac{\mathrm{E}\left[\dfrac{1}{1 - e^{-i\lambda} \cdot \varphi}\right]}{1 - e^{-i\lambda} \cdot \mathrm{E}\left[\dfrac{g(\varphi)}{1 - e^{-i\lambda} \cdot \varphi}\right]} \right|^2 . \tag{2.B.2}$$

A présent, concentrons nous sur le terme

$$N(x) = \mathrm{E}\left[\frac{1}{1 - x \cdot \varphi}\right]. \tag{2.B.3}$$

D'après le théorème de Karamata (Bingham, Goldie, et Teugels 1989), nous déduisons que $N(x) \sim (1-x)^{-\alpha}$, quand $x \to 1$, étant donné que la densité de φ satisfait $f(\varphi) \sim (1-\varphi)^{-\alpha}$, $\alpha \geq 0$.

De façon similaire, étant donné que la densité de φ satisfait l'hypothèse précédente et que l'espérance conditionnelle satisfait $g(\varphi) \sim (1-\varphi)^{\beta}$, $\beta \geq \alpha - 1/2$, le terme

$$D(x) = \mathrm{E}\left[\frac{g(\varphi)}{1 - x \cdot \varphi}\right] \sim \begin{cases} (1-x)^{-\alpha+\beta}, & \alpha > \beta \\ \mathrm{E}\left[\frac{g(\varphi)}{1-\varphi}\right], & \alpha < \beta \end{cases} \tag{2.B.4}$$

quand $x \to 1$.

donc, quand α est positif et supérieur à que β, le rapport $N(x)/D(x) \sim (1-x)^{-\beta}$ de telle sorte que la densité spectrale $f_X(\lambda) \sim \lambda^{-2\beta}$. Au contraire, quand β est supérieur ou égal à α, le rapport $N(x)/D(x) \sim (1-x)^{-\alpha}$ et la densité spectrale $f_X(\lambda) \sim \lambda^{-2\alpha}$. □

2.C Calcul de la fonction d'autocorrélation par l'utilisation de la FFT

La fonction d'autocorrélation $\gamma_X(h)$ d'une série temporelle $\{X_t\}$ peut être obtenue par la densité spectrale $f_X(\lambda)$ de $\{X_t\}$ selon

$$\gamma_X(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f_X(\lambda) d\lambda = \int_0^{2\pi} e^{i\lambda h} f_X(\lambda) d\lambda \tag{2.C.1}$$

L'équation (2.C.1) en termes d'une somme d'intégrales s'écrit

$$\gamma_X(h) = \sum_{k=0}^{N-1} \int_{\frac{2\pi}{N}k}^{\frac{2\pi}{N}(k+1)} e^{i\lambda h} f_X(\lambda) d\lambda \tag{2.C.2}$$

Par la méthode des trapèzes nous pouvons exprimer une intégrale sur un petit domaine tel que

$$\sum_{k=0}^{N-1} \int_{\frac{2\pi}{N}k}^{\frac{2\pi}{N}(k+1)} e^{i\lambda h} f_X(\lambda) d\lambda = \frac{\pi}{N}\left(e^{i\frac{2\pi}{N}k} f_X\left(\frac{2\pi}{N}k\right) + e^{i\frac{2\pi}{N}(k+1)} f_X\left(\frac{2\pi}{N}(k+1)\right)\right) \tag{2.C.3}$$

Etant donné la forme de la densité spectrale, nous devrions rencontrer des problèmes en 0 et 2π, ré-écrivons l'équation (2.C.2) comme

$$\begin{aligned}\gamma_X(h) = \int_0^{\frac{2\pi}{N}} e^{i\lambda h} f_X(\lambda) d\lambda + \int_{\frac{2\pi}{N}(N-1)}^{2\pi} e^{i\lambda h} f_X(\lambda) d\lambda \\ + \frac{\pi}{N} \sum_{k=1}^{N-2} \left(e^{i\frac{2\pi}{N}kh} f_X\left(\frac{2\pi}{N}k\right) + e^{i\frac{2\pi}{N}(k+1)h} f_X\left(\frac{2\pi}{N}(k+1)\right)\right)\end{aligned} \tag{2.C.4}$$

$$\begin{aligned}\gamma_X(h) = 2K(\alpha,\beta) \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{2\pi}{N}h\right)^{2k}}{(2k)!(2k+2\alpha+1)} - \frac{2\pi}{N} \mathcal{R}e\left\{e^{i\frac{2\pi}{N}h} f_X\left(\frac{2\pi}{N}\right)\right\} \\ + \frac{2\pi}{N} \sum_{k=1}^{N-1} e^{i\frac{2\pi}{N}kh} f_X\left(\frac{2\pi}{N}k\right)\end{aligned} \tag{2.C.5}$$

où l'expression de la densité spectrale est déduite à partir des équations (2.13) et (2.35) :

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{|1-e^{-i\lambda}|^{2\alpha}}{\left|1 - \frac{\Gamma(1+\alpha+\beta)}{\Gamma(1+\alpha)\Gamma(1+\beta)} e^{-i\lambda} F(1,-\alpha;1+\beta,e^{-i\lambda})\right|^2} \tag{2.C.6}$$

avec $F(.,.;.;.)$ la fonction Hypergéométrique, σ^2 la variance de ε_t, et

$$K(\alpha,\beta) = \frac{\sigma^2}{2\pi}\left(\frac{2\pi}{N}\right)^{2\alpha+1} \frac{1}{\left|1-\frac{\Gamma(\alpha+\beta)}{\Gamma(1+\alpha)\Gamma(\beta)}\right|^2} \tag{2.C.7}$$

Le dernier terme de l'équation (2.C.5) peut être facilement calculé grâce à l'utilisation de la Transformée de Fourier Rapide inverse (iFFT). □

Chapitre 3

Résultats empiriques

Maintenant que nous avons défini notre modèle et analysé ses paramètres, nous nous intéressons, dans ce chapitre, aux résultats empiriques qu'il nous fournit.

Nous entamons ce chapitre par une section consacrée à l'inférence statistique. Nous y définissons la densité du paramètre d'hétérogénéité du modèle, puis discutons du choix de l'estimateur. Nous achevons cette première section par des commentaires sur la mise en œuvre de l'optimisation où nous décrivons notamment l'algorithme que nous employons pour obtenir les estimations des paramètres du modèle. Dans la section 2, nous présentons les résultats empiriques obtenus à partir de notre modèle. Nous en profitons pour vérifier certains faits stylisés des rendements et de la volatilité réalisée, puis passons au calibrage du modèle et discutons des résultats obtenus. Enfin, dans la troisième section, nous menons une étude sur le comportements des agents et les différences que l'on observe en fonction de la phase (prospérité ou crise) du marché.

3.1 Inférence statistique

Nous avons montré dans le chapitre précédent le rôle crucial de la densité du paramètre d'hétérogénéité φ dans la modélisation de la log-volatilité réalisée. En effet, premièrement, le phénomène de mémoire longue peut-être attribué à la divergence hyperbolique[27] de la densité de φ au voisinage de 1^-. Deuxièmement, l'interprétation de sa densité va permettre de témoigner du comportement des investisseurs sur le marché.

Dans le chapitre précédent, nous avons obtenu à une densité, composée d'une densité de loi Beta pour sa capacité à reproduire la divergence hyperbolique au voisinage de 1^- et d'une densité en forme de bosse pour traiter une plus grande gamme de comportements. Nous allons reprendre ici ce principe. Ainsi, nous allons modéliser la loi du paramètre d'hétérogénéité

27. Comme nous l'avons vu au chapitre précédent, ceci assure que la décroissance de la fonction d'autocorrélation de la log-volatilité est hyperbolique.

comme un mélange d'une loi régulière définie sur $[-1,1]$, de densité f_1, et d'une loi Beta(p,q) définie sur $[0,1[$, de densité f_2.

En ce qui concerne l'estimation de la densité de la loi régulière, nous disposons de plusieurs méthodes couramment utilisées, notamment l'estimateur par noyau et l'estimateur par série de Fourier [28]. Dans notre situation, l'estimateur par noyau présente un sérieux inconvénient dans la mesure où la densité doit être définie sur $\mathbb{R}$, ce qui n'est pas du tout le cas de la densité de φ, définie sur $[-1,1]$. Puisque l'autre méthode ne possède pas cet inconvénient, nous estimons la densité de φ via un développement en série de Fourier. Nous émettons l'hypothèse suivante

$$f_1(x) = \frac{1}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi x) + \sum_{n=1}^{\infty} b_n \sin(n\pi x), \quad x \in [-1,1]. \tag{3.1}$$

Le terme $\frac{1}{2}$ dans l'équation (3.1) correspond au premier coefficient de Fourier a_0. En effet, c'est la condition nécessaire pour que f_1 soit une densité [29].

Quant à la loi Beta, nous avons montré que les paramètres p et q sont soumis à des contraintes pour que la densité soit correctement définie. Premièrement, pour obtenir la divergence au voisinage de 1^-, $0 < q < 1$, et pour qu'il n'y ait pas de divergence en zéro, $p > 1$. Nous posons $q = 1-d$ pour avoir une fonction d'autocorrélation se comportant comme $K|h|^{2d-1}$ quand $|h| \to +\infty$ et fixons arbitrairement [30] $p = 2$, ce qui conduit à la densité de la loi Beta$(2,1-d)$:

$$f_2(x) = (1-d)(2-d)\frac{x}{(1-x)^d}\mathbb{1}_{[0;1]}(x). \tag{3.2}$$

Au final, la densité de φ peut s'écrire

$$f(x) = wf_1(x) + (1-w)f_2(x), \quad w \in [0,1], \tag{3.3}$$

où w est le poids accordé à la densité régulière.

L'expression des fonctions d'autocovariance et d'autocorrélation de la log-volatilité réalisée peuvent être calculées en utilisant les expressions des moments d'ordre k de la variable d'hétérogénéité φ

$$\mathrm{E}\left[\varphi^k\right] = w\mathrm{E}_1\left[\varphi^k\right] + (1-w)\mathrm{E}_2\left[\varphi^k\right], \tag{3.4}$$

où $\mathrm{E}_1[\cdot]$ et $\mathrm{E}_2[\cdot]$ représentent les espérances par rapport, respectivement à f_1 et f_2, ce qui conduit à

$$\mathrm{E}\left[\varphi^k\right] = w\left(\frac{1}{2}A_k + \sum_{n=1}^{\infty} a_n B_{n,k} + \sum_{n=1}^{\infty} b_n C_{n,k}\right) + (1-w)\cdot\frac{\Gamma(k+2)}{\Gamma(k+3-d)}\Gamma(3-d), \tag{3.5}$$

où les expressions de A_k, B_k et C_k sont données en annexe 3.A.

28. Pour plus de détails sur ces deux estimateurs, le lecteur est invité à se référer à Pagan et Ullah (1999)

29. Plus précisément, $a_0 = \frac{1}{2}$ découle de la condition de normalisation de la densité : $\int_{-1}^{1} f_1(x)dx = 1$, avec $f_1(x) \geq 0$ pour tout $x \in [-1,1]$.

30. Fixer $p = 2$ n'a pas d'importance parce que le comportement régulier (notamment en zéro) de la densité f est capturé par f_1.

Pour plus de simplicité et parce qu'il est impossible d'estimer un nombre infini de paramètres, nous supposons que les coefficients de Fourier (a_n, b_n) de l'équation (3.1) deviennent négligeables au-delà du rang q, de telle sorte que la densité f_1 s'écrit

$$f_1(x) = \frac{1}{2} + \sum_{n=1}^{q} a_n \cos(n\pi x) + \sum_{n=1}^{q} b_n \sin(n\pi x), \quad x \in [-1, 1]. \tag{3.6}$$

La discussion de la valeur de troncature q appropriée est abordée dans la section 3.1.2.

Notons par $\theta = (a_1, \ldots, a_q, b_1, \ldots, b_q, \alpha, w, \sigma, d)'$ le vecteur de dimension $2q + 4$ des paramètres de notre problème : les a_i et b_i sont les coefficients de Fourier, α est le paramètre qui caractérise en partie l'écart à la rationalité, w est le poids accordé à la densité f_1 (autrement dit la densité de la loi régulière), σ est l'écart-type de ε_t (l'information publique) et, enfin, d est le paramètre de mémoire longue. Notons également par $\gamma(h; \theta)$ la valeur de la fonction d'autocovariance au lag h de la distribution d'hétérogénéité définie par les équations (3.3), (3.2) et (3.6) et pour la valeur θ du paramètre.

3.1.1 Choix de l'estimateur

Étant donné l'échantillon de taille T des données observées du logarithme de la volatilité réalisée $\{\omega_1, \ldots, \omega_T\}$, il semble naturel d'estimer θ en minimisant la différence pondérée entre la fonction d'autocovariance empirique

$$\hat{\gamma}_T(h) = \frac{1}{T} \sum_{i=1}^{T-h} (\omega_i - \bar{\omega}_T)(\omega_{i+h} - \bar{\omega}_T), \qquad h = 0, 1, \ldots, \tag{3.7}$$

et $\gamma(h; \theta)$, la fonction d'autocovariance pour la valeur du paramètre θ. Nous pourrions ainsi considérer l'estimateur de la distance minimum classique[31] $\hat{\theta}_{T,L}$ (Newey et McFadden 1994) solution de

$$\min_{\theta \in \Theta} \left[\hat{\gamma}_{T,L} - \gamma_L(\theta)\right]^t W_L^{-1} \left[\hat{\gamma}_{T,L} - \gamma_L(\theta)\right], \tag{3.8}$$

où $\gamma_L(\theta) = (\gamma(1; \theta), \ldots, \gamma(L; \theta))^t$, $\hat{\gamma}_{T,L} = (\hat{\gamma}_T(1), \ldots, \hat{\gamma}_T(L))^t$ et W_L est une matrice symétrique, définie positive quelconque de taille $L \times L$ dont le but est de pondérer l'importance des lags (en particulier, elle accorde plus de poids aux points dont l'estimation est plus fiable), tandis que Θ est l'espace des paramètres

$$\Theta = \left\{\theta = (a_1, \ldots, a_q, b_1, \ldots, b_q, \alpha, w, \sigma, d)';\ f_1 \geq 0,\ 0 < d < 1 \text{ et } 0 \leq w \leq 1\right\}. \tag{3.9}$$

Toutefois, lorsque l'on a affaire à des séries temporelles à mémoire longue, *i.e* telles que $\sum_{h=0}^{\infty} |\gamma(h)| = \infty$, les propriétés asymptotiques des estimations $\hat{\gamma}_{T,L}$ de l'échantillon ne sont pas vraiment satisfaisantes. En effet, comme le rappelle Hosking (1996), la distribution limite de $\gamma_T(h)$ lorsque $d > 0$ est telle que

$$T^{1-2d} \left(\hat{\gamma}_T(h) - \gamma(h)\right) \xrightarrow{\mathcal{L}} \mathcal{R}, \tag{3.10}$$

31. Dans la littérature, cet estimateur porte le nom de *classical minimum distance estimator*.

où $\mathcal{R}$ décrit la distribution modifiée de Rosenblatt [32]. En particulier, la moyenne de la distribution de Rosenblatt n'est pas égale à zéro et peut même être plus grande que son écart-type pour une valeur de d supérieure à, ou de l'ordre de $1/4$.

Par conséquent, il est tentant d'avoir recours à la minimisation d'un autre critère avec des propriétés asymptotiques plus intéressantes. En fait, indépendamment de la valeur de $d \in]-\frac{1}{2}, \frac{1}{2}[$, la distribution limite de n'importe quel sous-ensemble de variables

$$D_h = \sqrt{T}\left[(\hat{\gamma}_T(h) - \hat{\gamma}_T(0)) - (\gamma(h) - \gamma(0))\right], \quad h \geq 1, \tag{3.11}$$

est une loi normale multivariée de moyenne 0 et de matrice de covariance (voir Hosking (1996, th. 5))

$$\begin{aligned} \left[\Sigma^2\right]_{kl} &= \lim_{T\to\infty} \mathrm{Cov}\left(D_k, D_l\right), & (3.12)\\ &= \frac{1}{2}\sum_{s=-\infty}^{\infty}\left[\gamma(s) - \gamma(s-k) - \gamma(s-l) + \gamma(s-k+l)\right]^2. & (3.13) \end{aligned}$$

Étant donné ce résultat asymptotique, un estimateur adéquat du paramètre θ est donné par la solution de

$$\min_{\theta\in\Theta}\left[\hat{\eta}_{T,L} - \eta_L(\theta)\right]^t W_L^{-1}\left[\hat{\eta}_{T,L} - \eta_L(\theta)\right], \tag{3.14}$$

où

$$\eta_L(\theta) = \left(\gamma(1;\theta) - \gamma(0;\theta), \ldots, \gamma(L;\theta) - \gamma(0;\theta)\right)^t, \tag{3.15}$$

et

$$\hat{\eta}_{T,L} = \left(\hat{\gamma}_T(1) - \hat{\gamma}_T(0), \ldots, \hat{\gamma}_T(L) - \hat{\gamma}_T(0)\right)^t. \tag{3.16}$$

Propriété asymptotique de l'estimateur

La convergence et la normalité asymptotique de l'estimateur $\hat{\theta}_{T,L}$ découlent des propriétés asymptotiques générales de l'estimateur de la distance minimum classique. En ce qui concerne la normalité asymptotique, nous pouvons établir le résultat suivant

Proposition 3. Nous supposons que $\theta_{T,L} \xrightarrow{P} \theta_0$, pour tout $L \geq 2q+4$, quand $T \longrightarrow \infty$,

$$\sqrt{T}\left(\hat{\theta}_{T,L} - \theta_0\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, A_L(\theta_0)\right), \tag{3.17}$$

avec

$$A_L(\theta_0) = \left[G_L(\theta_0)^t W_L G_L(\theta_0)\right]^{-1} G_L(\theta_0)^t W_L \Sigma_L^2(\theta_0) W_L G_L(\theta_0)\left[G_L(\theta_0)^t W_L G_L(\theta_0)\right]^{-1}, \tag{3.18}$$

où

$$G_L(\theta_0) = \mathrm{grad}_\theta \eta_L(\theta)|_{\theta=\theta_0}, \tag{3.19}$$

et

$$\left[\Sigma_L^2(\theta_0)\right]_{lk} = \frac{1}{2}\sum_{s=-\infty}^{\infty}\left[\gamma\left(s;\theta_0\right) - \gamma\left(s-k;\theta_0\right) - \gamma\left(s-l;\theta_0\right) + \gamma\left(s-k+l;\theta_0\right)\right]^2, \tag{3.20}$$

$l, k = 1, \ldots, L$.

32. Le lecteur est invité à se référer à Hosking (1996).

Démonstration. Par le théorème 5 dans Hosking (1996), $\sqrt{T}\left(\hat{\eta}_{T,L} - \eta_L(\theta_0)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \Sigma_L^2(\theta_0)\right)$, ainsi, le résultat provient directement du théorème 3.2 dans Newey et McFadden (1994). □

Estimation du lag optimal

Pour un lag fixé L, le minimum de la variance asymptotique $A_L(\theta_0)$ est atteint lorsque $W_L = \left[\Sigma_L^2(\theta_0)\right]^{-1}$, de sorte que

$$\sqrt{T}\left(\hat{\theta}_{T,L} - \theta_0\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, G_L(\theta_0)^t \left[\Sigma_L^2(\theta_0)\right]^{-1} G_L(\theta_0)\right). \tag{3.21}$$

D'un autre côté, étant donné W_L, nous pouvons obtenir une valeur optimale L^* de L comme solution de

$$L^* = \arg \min_{L \in \{2q+4, \ldots, T-1\}} ||A_L(\theta_0)||^2. \tag{3.22}$$

Proposition 4. Sous les hypothèses de la proposition 3, en notant par $D(\theta)$ le vecteur

$$D(\theta) = \begin{bmatrix} w\cos(\pi x) \\ \vdots \\ w\cos(q\pi x) \\ w\sin(\pi x) \\ \vdots \\ w\sin(q\pi x) \\ 0 \\ f_1(x) - f_2(x) \\ 0 \\ (w-1)f_2(x)\left[\ln(1-x) + \frac{3-2d}{(1-d)(2-d)}\right] \end{bmatrix}, \tag{3.23}$$

l'estimateur $\hat{f}(x)$ de la densité $f(x)$ est asymptotiquement gaussien

$$\sqrt{T}\left(\hat{f}(x) - f(x)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, D(\theta_0)^t A(\theta_0) D(\theta_0)\right). \tag{3.24}$$

Démonstration. $D(\theta)$ n'est rien d'autre que le gradient de $f(x)$ par rapport à θ. Ainsi, en utilisant la méthode Delta (Van der Vaart 2000), le résultat découle directement de la proposition 3. □

3.1.2 Commentaires sur la mise en œuvre

Nature convexe du problème

Pour des échantillons de tailles finies, la distance $\left(\hat{\eta}_{T,L} - \eta_L(\theta)\right)^t W_L^{-1} \left(\hat{\eta}_{T,L} - \eta_L(\theta)\right)$ peut posséder plusieurs minima locaux, et en pratique, c'est le cas. Par conséquent, il parait nécessaire d'utiliser un algorithme de minimisation qui soit capable de gérer ce problème, en

évitant de se faire piéger dans un minimum local, et ainsi trouver le minimum global. Dans de telles situations, les algorithmes génétiques fournissent des solutions pertinentes. C'est pourquoi ils ont été retenus pour résoudre notre problème.

L'idée sous-jacente des algorithmes génétiques est basée sur l'imitation du processus de sélection naturel et les principes génétiques. L'algorithme génétique débute avec une population de vecteurs d'essai – appelés gènes – contenant le paramètre θ à optimiser et procède comme suit :

- La première étape consiste à *répliquer* les vecteurs d'essai initiaux selon leur adéquation, c'est à dire les gènes qui satisfont au mieux le critère d'optimisation [33] ont la plus grande probabilité d'être reproduit. Ainsi, en moyenne, la nouvelle population a une distance plus faible que la population initiale, mais sa diversité est aussi plus faible puisque les gènes les plus justes apparaissent deux fois plus (ou plus encore) dans la nouvelle population.
- La deuxième étape est le *croisement* qui conduit à combiner les différents paramètres dans différents vecteurs tirés dans la nouvelle population afin de mélanger leurs caractéristiques.
- La troisième et dernière étape est la *mutation*, où certains gènes subissent des changements aléatoires, *i.e.*, certains paramètres des vecteurs nés du croisement sont modifiés de manière aléatoire. Cette étape est essentielle pour maintenir la diversité de la population qui assure l'exploration de l'espace d'optimisation tout entier.

Les vecteurs obtenus après cette troisième étape sont alors utilisés comme population initiale et le processus est réitéré afin d'obtenir une nouvelle génération de gènes, et ainsi de suite. La convergence de cet algorithme vers le minimum global du problème est assuré par le théorème fondamental des algorithmes génétiques (Golberg 1989). Un exemple d'algorithmes génétiques particulièrement efficaces est l'algorithme génétique différentiel avancé [34] de Price et Storn (1997) ou l'algorithme de Dorsey et Mayer (1995).

Inconvénient de l'algorithme génétique

Comme l'algorithme génétique est très coûteux en temps de calcul, nous nous sommes orientés vers la méthode du simplexe [35] de Nelder et Mead (1965), qui est un algorithme de minimisation multidimensionnel, non linéaire et sans contrainte. Cette méthode présente toutefois un sérieux inconvénient dans notre situation : à partir du point de départ choisi pour initialiser la procédure, il trouve le minimum local le plus proche. Or, dans le cas qui nous intéresse, nous savons pertinemment qu'il en existe plusieurs. Ainsi, un mauvais choix dans la valeur initiale risque très probablement de nous conduire à un minimum local plutôt que le minimum global que nous recherchons.

33. Dans le cas qui nous concerne, il s'agit de la distance $(\hat{\eta}_{T,L} - \eta_L(\theta))^t W_L^{-1} (\hat{\eta}_{T,L} - \eta_L(\theta))$.
34. Le nom d'origine de cet algorithme est : Differential Evolutionary Genetic Algorithm.
35. voir aussi Lagarias, Reeds, Wright, et Wright (1998) pour une discussion récente sur la convergence de la méthode.

Choix de l'algorithme

Afin de surmonter ce problème, nous avons développé une procédure itérative, ci-après appelée méthode étape par étape, basée sur la méthode de Nelder Mead, qui s'est trouvée être très efficace et bien plus rapide que l'algorithme génétique, lorsque les coefficients (a_n, b_n) de l'équation (3.6) décroissent au moins comme $1/n^2$. Cette méthode consiste à commencer par restreindre l'optimisation à $q = 1$. Dans cette première étape, l'optimisation peut être réalisée au moyen de la méthode de Nelder Mead. Elle fournit une première estimation $\left(\hat{a}_{1(1)}, \hat{b}_{1(1)}, \hat{\alpha}_{(1)}, \hat{w}_{(1)}, \hat{\sigma}_{(1)}, \hat{d}_{(1)}\right)$. Dans une deuxième étape, nous posons $q = 2$, et lançons l'algorithme de Nelder Mead pour la valeur initiale $\left(\hat{a}_{1(1)}, 0, \hat{b}_{1(1)}, 0, \hat{\alpha}_{(1)}, \hat{w}_{(1)}, \hat{\sigma}_{(1)}, \hat{d}_{(1)}\right)$ qui conduit à une nouvelle estimation des paramètres $\left(\hat{a}_{1(2)}, \hat{a}_{2(2)}, \hat{b}_{1(2)}, \hat{b}_{2(2)}, \hat{\alpha}_{(2)}, \hat{w}_{(2)}, \hat{\sigma}_{(2)}, \hat{d}_{(2)}\right)$, et ainsi de suite jusqu'à ce que la valeur fixée de q soit atteinte.

La figure 3.1 illustre la convergence de cette procédure dans le cas de l'expérience numérique suivante :

1. nous générons une densité de référence avec un q choisi,
2. nous appliquons les deux procédures utilisant la méthode de Nelder Mead,
3. nous comparons la précision des deux densités estimées à la densité de référence,
4. les étapes (1) et (2) sont itérées 1000 fois.

Les deux graphiques montrent l'efficacité de la méthode étape par étape. Sur le graphique du haut, la densité de référence a été tracée pour $q = 5$ et la densité estimée pour $q = 5$ aussi. La meilleure densité estimée est sans conteste celle obtenue avec la procédure itérative. Sur le graphique de droite, nous prenons en compte que la densité devrait avoir un nombre infini de paramètres (en d'autres termes, un nombre infini de coefficients de Fourier, voir équation (3.1)) ou bien, au moins, le véritable ordre q dans l'équation (3.6) devrait être inconnu. C'est la raison pour laquelle nous avons généré une densité de référence avec $q = 10$ et lancé l'estimation pour $q = 5$, pour étudier spécialement l'effet de troncature sur la précision des résultats des deux approches. Le résultat obtenu à partir d'une simulation tirée aléatoirement est montré sur le graphique du bas de la figure 3.1. Une fois encore, la méthode itérative fournit de meilleurs résultats que l'approche globale.

Hypothèse sur l'estimateur

Bien que nous n'ayons pas été capable de prouver la convergence de notre procédure étape par étape, nos simulations numériques montrent qu'elle fournit des estimations qui sont toujours proches de la vrai valeur des paramètres (sous l'incertitude prédite par la proposition 3). En outre, ces estimations sont presque toujours plus précises que celles obtenues en utilisant l'algorithme génétique, à cause de la convergence très lente, dans notre situation, de ce type d'algorithme.

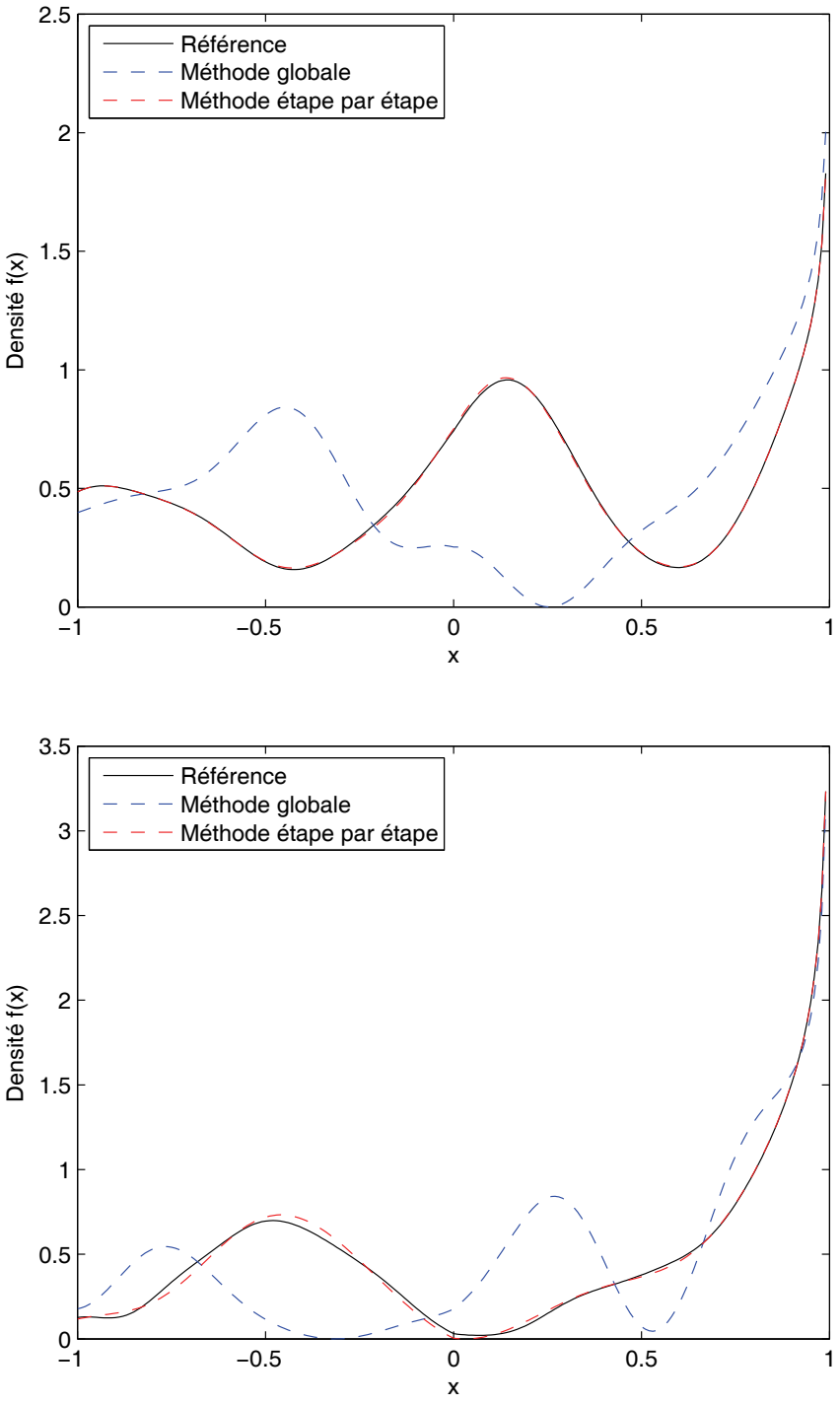

FIGURE 3.1 – Comparaison de l'efficacité des algorithmes de minimisation basés sur la méthode de Nelder Mead. La densité estimée est toujours calculée pour $q = 5$, alors que la densité de référence est calculée pour $q = 5$ (en haut) et $q = 10$ (en bas).

Impossibilité de mettre en œuvre le calcul du lag optimal

Le temps de calcul a été un obstacle majeur dans la mise en œuvre de la procédure économétrique. Nous avons d'abord été contraints d'abandonner l'algorithme génétique, pourtant particulièrement adapté à notre problème, mais cela ne s'arrête pas là. En effet, nous avons également été obligé de mettre de côté le calcul du lag optimal, extrêmement coûteux en temps. Nous avons donc fixé arbitrairement le lag optimal $L^* = 120$, soit environ 6 mois, ce qui nous a semblé être un bon compromis étant donné les caractéristiques de la volatilité d'actions.

3.2 Résultats empiriques

Nous estimons la volatilité réalisée journalière au moyen de l'estimateur (1.10), comme nous l'avons discuté au préalable dans la section 1.2.3 (chapitre 1). Nous soulignons que cette variable sera considérée comme une variable observée dans toute la suite. C'est une hypothèse classique de la littérature sur la volatilité réalisée.

Dans cette section, nous commençons par valider les faits stylisés des rendements et de la volatilité des rendements d'actions. En particulier, nous vérifions si notre hypothèse de base au sujet de la log-normalité de la volatilité réalisée est valable. Cette question a été discutée dans la section 1.2.4 (chapitre 1) sur la base d'une impression visuelle qui demande, à présent, à être formalisée. Puis, nous présentons les conclusions tirées à partir du calibrage de notre modèle. Enfin, nous étudions le comportement des agents en fonction de la phase dans laquelle se trouve le marché, à savoir une phase de croissance ou au contraire une crise.

Dans ce but, nous utilisons des prix hautes-fréquences de 24 actions américaines échangées sur le New York Stock Exchange (ci-après NYSE) ou sur le Nasdaq entre le 01/01/1994 et le 31/12/2003 (soit 2518 jours ouvrés). Cette période présente l'intérêt d'inclure des évènements importants :

- la bulle internet de la fin des années 1990 (apogée en mars 2000),
- la crise obligataire de fin 1998.

Les données, fournies par TickData [36], sont détaillées dans le tableau 3.1. Il regroupe 10 actions à moyenne capitalisation (dont les capitalisations, en moyenne au cours de la période considérée, se situent entre 1.62 et 3.90 milliards de dollars) et 14 actions à forte capitalisation (entre 108.01 et 290.44 milliards de dollars, à l'exception des deux plus faibles établies à 31.84 et 84.30 milliards de dollars).

Ces compagnies exercent dans des secteurs variés : la distribution commerciale (5) et non commerciale (5), la finance (5), les technologies (4), l'énergie (2), la communication (2) et l'industrie (1). Nous constatons d'une part que ces secteurs ne sont pas équitablement représentés, et d'autre part qu'il n'y a pas assez de compagnie par secteur pour pouvoir tirer des généralités par secteur.

36. L'adresse du site internet de l'entreprise est http ://www.tickdata.com/.

Symbole	Compagnie	Marché	Secteur	Cap. (Md US $)
BKS	Barnes & Noble Inc.	NYSE	dist. (c)	1.62
VLO	Valero Energy corp (new)	NYSE	énergie	1.70
DHI	DR Horton inc.	NYSE	dist. (c)	1.83
LEN	Lennar corp CL a common	NYSE	dist. (c)	1.85
TCB	TCF Financial Corp.	NYSE	financier	1.92
NYB	New York Bancorp inc.	NYSE	financier	2.17
FNF	Fidelity Natl Financial inc.	NYSE	financier	2.27
MCHP	Microchip Technology inc.	Nasdaq	tech.	2.99
WPO	Washington post co clb	NYSE	com.	3.71
GILD	Gilead Science inc.	Nasdaq	dist. (nc)	3.90
GM	General Motors co.	NYSE	dist. (c)	31.81
PG	Procter & Gamble co.	NYSE	dist. (nc)	84.30
IBM	Intel Business Machines corp.	NYSE	tech.	108.01
CSCO	Cisco Systems inc.	Nasdaq	com.	112.26
MRK	Merck & co inc.	NYSE	dist. (nc)	112.73
KO	Coca-Cola co.	NYSE	dist. (nc)	118.97
AIG	American intl group inc.	NYSE	financier	125.58
INTC	Intel corp.	Nasdaq	tech.	140.77
WMT	Wall-Mart Stores inc.	NYSE	dist. (c)	144.61
C	Citigroup	NYSE	financier	148.79
PFE	Pfizer inc.	NYSE	dist. (nc)	198.20
XOM	Exxon mobile corporation	NYSE	énergie	199.19
MSFT	Microsoft corp.	Nasdaq	tech.	239.81
GE	General Electrics co.	NYSE	industriel	290.44

TABLE 3.1 – Descriptif des actions constituant notre base de données : nom, marché, secteur et capitalisation moyenne (en milliards de dollars US) sur toute la période (du 01/01/1994 au 31/12/2003).

Avant d'aller plus loin, il est préférable de signaler que nous devrions obtenir 390 cotations par jours puisque la séance de cotation débute à 9h30 et s'achève à 16h. Toutefois, nous disposons d'un nombre de données bien inférieur, en particulier pour les moyennes capitalisations : 133 par jour en moyenne. De plus, le nombre de données journalières varie au cours du temps : sur la période allant de 1994 à 1997 inclu, le nombre de données disponibles est très inférieur à celui du reste de la période (respectivement 45 et 192 pour les moyennes capitalisations, et 282 et 374 pour les grandes capitalisations). Néanmoins, puisque nous n'avons pas observé de différence de comportement des séries de volatilité sur la prériode 1994 - 1997 et sur la période 1997 - 2003, nous avons décidé d'estimer la volatilité réalisée sur toute la période allant du 01/01/1994 au 31/12/2003.

3.2.1 Validation des faits stylisés des rendements et de la volatilité des rendements d'actions

Les faits stylisés des rendements et de la volatilité réalisée que nous allons tester concernent la normalité, ou l'absence de normalité de ces séries. Pour pouvoir accepter ou rejeter les hypothèses de normalité, nous allons recourir à un test statistique de normalité. Notre choix

s'est porté sur le test de Jarque-Bera. Par la suite, les hypothèses testées le seront avec un niveau de confiance à 95%.

En ce qui concerne la volatilité réalisée, nous avons détaillé dans la section 1.2.3 (chapitre 1) deux estimateurs de la volatilité réalisée : celui défini par Andersen, Bollerslev, Diebold, et Ebens (2001), noté $\hat{\sigma}_{ABDL}$ et décrit dans l'équation (1.10) et celui développé par Martens (2002) à partir de ce dernier pour tenir compte de la volatilité overnight, noté $\hat{\sigma}_M$ et présenté dans l'équation (1.13). Nous avons testé les hypothèses à partir des séries de volatilité obtenues au moyen de ces deux estimateurs. Il en résulte que l'apport de l'estimateur $\hat{\sigma}_M$ n'est pas concluant. Par conséquent, nous nous contentons de présenter les résultats obtenus à partir de $\hat{\sigma}_{ABDL}$.

Non normalité des rendements journaliers

L'hypothèse de normalité d'un échantillon aléatoire est habituellement testé par le test skewness-kurtosis. La skewness est un nombre sans dimension qui caractérise le degré d'asymétrie d'une distribution autour de sa moyenne. Sa définition est

$$S(x_1, \ldots, x_n) = \frac{1}{N} \sum_{j=1}^{N} \left[\frac{x_j - \bar{x}}{\sigma} \right]^3, \tag{3.25}$$

où $\sigma = \sigma(x_1, \ldots, x_n)$ est l'écart-type de l'échantillon. Une valeur positive (resp. négative) de la skewness signifie que la distribution a tendance à être asymétrique à droite (resp. à gauche).

La kurtosis, quant à elle, est aussi une quantité sans dimension qui mesure l'aplatissement de la distribution relativement à une distribution normale. Si la kurtosis est supérieure (resp. inférieure) à 3 alors la distribution est plus puiquée (resp. plate) que celle d'une loi normale. La kurtosis est donnée par

$$K(x_1, \ldots, x_n) = \frac{1}{N} \sum_{j=1}^{N} \left[\frac{x_j - \bar{x}}{\sigma} \right]^4. \tag{3.26}$$

Le test skewness-kurtosis est couramment appelée test de Jarque-Bera. Sa statistique est la somme du carré de la skewness de l'échantillon et de l'excès de kurtosis, correctement normalisées par leur variance asymptotique dans le cas d'un bruit blanc :

$$JB_n = n \left(\frac{S^2}{6} + \frac{(K-3)^2}{24} \right), \tag{3.27}$$

où K est la kurtosis de la série, S sa skewness et n sa taille.

Les rendements journaliers sont calculés à partir des prix de fermeture à fermeture. Ainsi, le rendement du jour j, prendra en compte le rendement overnight (entre le cours d'ouverture

du jour j et le cours de clôture de la veille $j-1$), noté $r_{j-1,co}$ et le rendement de la journée de cotation j (entre le cours d'ouverture et le cours de clôture), noté $r_{j,oc}$:

$$r_j = \underbrace{\ln P_{j,0} - \ln P_{j-1,T}}_{r_{j-1,co}} + \underbrace{\ln P_{j,T} - \ln P_{j,0}}_{r_{j,oc}} = \ln P_{j,T} - \ln P_{j-1,T}. \tag{3.28}$$

La figure 3.2 montre les rendements journaliers de l'action TCF Financial du 01/01/1994 au 31/12/2003. La majorité des rendements sont compris entre -2.5% et $+2.5\%$. Les rendements prennent des valeurs extrêmes aussi bien en positif qu'en négatif à la fin de l'année 1998 et au début de l'année 2000, ce qui correspond respectivement à la crise obligataire et à la crise des nouvelles technologies.

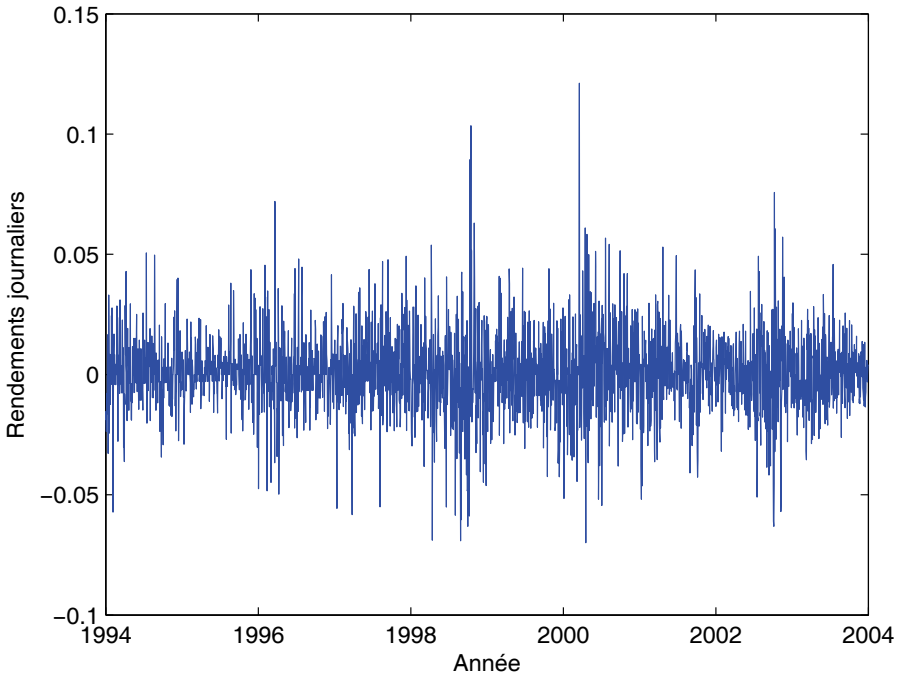

FIGURE 3.2 – Rendements journaliers de l'action TCF Financial sur une période de 10 ans, allant du 01/01/1994 au 31/12/2003.

Nous avons regroupé les différentes statistiques de la série des rendements journaliers dans le tableau 3.2. Celui-ci nous indique que les valeurs de la statistique JB_n sont toutes très grandes et largement supérieures au seuil critique de non-rejet de l'hypothèse de normalité (5.99) : les statistiques sont dans le meilleur des cas de l'ordre de plusieurs centaines (327 pour PFE) et les p-values sont, logiquement, toutes égales à zéro. Nous pouvons donc conclure à la non-normalité des rendements journaliers de toutes les actions de notre échantillon.

En regardant plus attentivement ce tableau, nous constatons que la skewness est proche de zéro en valeur absolue : $|S| < 0.3$ dans 18 cas sur 24. A noter 2 valeurs extrêmes : LEN (-8.73) et PG (-3.62), qui présentent également une kurtosis corrigée ($K-3$) anormalement

	$\bar{m}$ (en millièmes)	$\bar{\sigma}$	S	K-3	JB_n stat.	p-value
Statistique	-	-	0	0	5.992	-
BKS	0.39	0.030	-0.97	23.47	58181	0
VLO	0.49	0.023	0.29	2.72	808	0
DHI	0.93	0.030	0.08	2.06	436	0
LEN	0.62	0.031	-8.73	243.21	$6.2\ 10^6$	0
TCB	0.73	0.017	0.25	3.15	1063	0
NYB	1.30	0.018	-0.11	3.43	1206	0
FNF	0.50	0.025	0.06	4.69	2298	0
MCHP	0.93	0.042	0.08	2.80	823	0
WPO	0.45	0.011	0.40	4.65	2331	0
GILD	1.20	0.042	0.03	3.03	961	0
GM	-0.01	0.021	-0.31	4.41	2084	0
PG	0.50	0.018	-3.62	73.15	5.710^5	0
IBM	0.74	0.023	-0.10	7.81	6403	0
CSCO	1.00	0.034	-0.03	4.15	1807	0
MRK	0.40	0.018	-0.08	2.54	679	0
KO	0.38	0.017	-0.05	3.58	1343	0
AIG	0.62	0.019	0.19	2.69	770	0
INTC	0.84	0.030	-0.41	5.26	2968	0
WMT	0.56	0.021	0.12	2.13	481	0
C	0.78	0.023	0.12	4.27	1917	0
PFE	0.72	0.020	-0.17	1.74	327	0
XOM	0.37	0.015	0.08	2.83	843	0
MSFT	0.95	0.024	-0.13	4.21	1867	0
GE	0.50	0.019	0.07	3.07	989	0

TABLE 3.2 – Statistiques de la série des rendements journaliers du 01/01/1994 au 31/12/2003. L'hypothèse de normalité pour un niveau de confiance à 95 % est acceptée si la statistique $JB_n < 5.99$ ou si la p-value ≥ 0.05.

élevée en comparaison de celle des autres actions : 243.21 pour LEN et 73.15 pour PG, alors que dans les autres cas ($1.74 < K - 3 < 5.26$) dans 20 cas sur 24. Ces valeurs de la kurtosis corrigée, toutes nettement positives, signifient que les densités des rendements journaliers des actions sont plus piquées que des densités d'une loi normale.

Pour illustrer cette non-normalité des rendements nous avons tracé sur la figure 3.3 la densité de la loi des rendements, obtenue au moyen de l'estimateur par noyau, ainsi que la densité de la loi normale la plus proche. Les tests statistiques sont immédiatement confirmés par notre impression visuelle. En effet, la densité des rendements est beaucoup plus piquée que celle de la loi normale, et on note des queues épaisses.

Normalité de la log-volatilité réalisée

Le test skewness-kurtosis est adapté pour les séries contenant des observations indépendantes. Par conséquent, ils ne sont pas pertinents pour mesurer le degré de normalité des séries temporelles financières et économiques dont les observations sont corrélées, telles que la série de la log-volatilité des rendements journaliers d'actions.

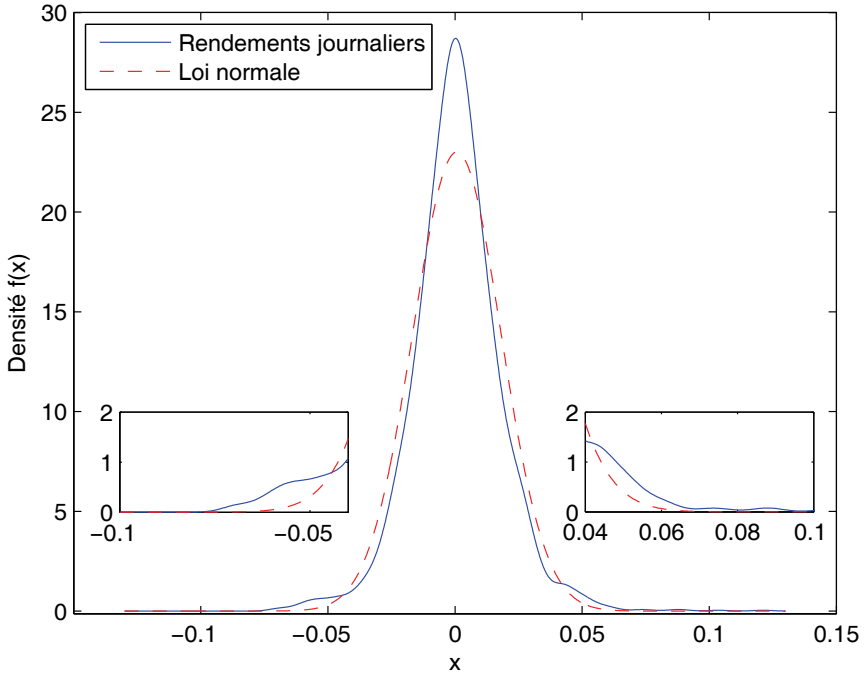

FIGURE 3.3 – Densité de la loi des rendements journaliers de l'action TCF Financial calculée sur une période de 10 ans, allant du 01/01/1994 au 31/12/2003 à partir de l'estimateur par noyau.

Pour surmonter cette lacune, Lobato et Velasco (2004) suggèrent d'employer la statistique du test standard, basé sur la skewness et la kurtosis de l'échantillon mais normalisées de manière adéquate : la normalisation tiendra compte de la dépendance sérielle des observations qui entraîne une augmentation de l'écart-type asymptotique de la skewness et de la kurtosis de l'échantillon pour des observations positivement corrélées.

Donc, soit $\mu_k = \mathrm{E}(\omega_t - \bar{\omega}_n)^k$, pour k entier naturel, les moments centrés du logarithme de la série de la volatilité réalisée, avec $\hat{\mu}_k = n^{-1}\sum_{t=1}^{n}(\omega_t - \bar{\omega}_n)^k$ les moments de l'échantillon correspondant, tandis que $\bar{\omega}_n$ décrit la moyenne de l'échantillon et n sa taille.

Le test skewness-kurtosis généralisé proposé par Lobato et Velasco (2004) est donné par :

$$JB_n^* = n\left(\frac{\hat{\mu}_3^2}{6\hat{F}^{(3)}} + \frac{(\hat{\mu}_4 - 3\hat{\mu}_2^2)^2}{24\hat{F}^{(4)}}\right), \tag{3.29}$$

avec

$$F^{(k)} = \sum_{j=1-n}^{n-1} \hat{\gamma}(j)^k, \tag{3.30}$$

où $\hat{\gamma}(j) = n^{-1}\sum_{t=1}^{n-|j|}(\omega_t - \bar{\omega}_n)(\omega_{t+|j|} - \bar{\omega}_n)$ est l'autocovariance de l'échantillon qui doit vérifier

$$\sum_{j=0}^{\infty} |\gamma(j)| < \infty. \tag{3.31}$$

Par ailleurs, pour que $F^{(k)} < +\infty$, il suffit que $|\gamma(j)|^{k(2d-1)} < +\infty$. Autrement dit,

$$d < \frac{1}{2}\left(1 - \frac{1}{k}\right). \tag{3.32}$$

Nous en déduisons que pour $k = 3$, $d < 1/3$ et pour $k = 4$, $d < 3/8$. En conclusion, il faudrait que

$$d < 0.\bar{3}, \tag{3.33}$$

pour que $F^{(3)}$ et $F^{(4)}$ convergent. C'est la condition minimale pour que JB_n^* converge vers quelque chose, même si ce n'est pas suffisant pour conserver cette statistique. Or, rien n'indique que ce soit le cas pour les séries temporelles à mémoire longue. Pour ce qui est des séries de log-volatilité réalisée, comme nous l'illustrons dans le tableau 3.6 de la section 3.2.2, c'est même bien souvent le contraire. Par conséquent, ce test n'est pas, pour ainsi dire, approprié dans notre situation. Malheureusement, à notre connaissance, aucun test n'a encore été développé pour tester l'hypothèse de normalité de séries temporelles en présence de mémoire longue. Toutefois, nous pouvons raisonnablement nous attendre à ce que la statistique du test (3.29) fournisse une borne inférieure de la valeur critique réelle d'un test de normalité basé sur la skewness et la kurtosis de l'échantillon tenant compte de la présence de mémoire longue, dans la mesure où l'action de la mémoire longue implique une augmentation de l'écart-type asymptotique de la skewness et de la kurtosis de l'échantillon. Ainsi, le non-rejet de l'hypothèse de normalité en utilisant (3.29), devrait être interprété en faveur de l'acceptation de cette hypothèse. Au contraire, le rejet devrait être considéré comme non-concluant.

Pour pouvoir calibrer le modèle comme nous l'avons proposé dans la section 3.1, nous avons à vérifier que la volatilité réalisée ne s'éloigne pas trop d'une loi log-normale, autrement dit que la log-volatilité réalisée suit une loi normale. Les résultats des tests sont présentés dans le tableau 3.3.

L'hypothèse de normalité avec un niveau de confiance à 95 % est très largement accepté pour 3 moyennes capitalisations sur 10 (TCB, MCHP et GILD avec une p-value respectivement égale à 0.41, 0.85 et 0.72) et pour 3 grandes capitalisations sur 14 (CSCO, KO, WMT avec une p-value respectivement égale à 0.24, 0.60 et 0.61). Nous avons tracé, sur la figure 3.4, la densité de la loi de la log-volatilité réalisée de KO obtenue au moyen de l'estimateur par noyau ainsi que la densité d'une loi normale. Nous constatons que les deux courbes sont très proches l'une de l'autre. Contrairement à la figure 3.3 de la densité des rendements journaliers, les queues épaisses ont disparu et les pics sont similaires. On note toutefois une légère asymétrie vers la droite mais qui n'est pas suffisante pour faire rejeter l'hypothèse pour un niveau de confiance à 95 %.

Pour les autres, les statistiques sont, en grande partie, de l'ordre de quelques dizaines, et pour quelques cas extrêmes elles peuvent atteindre plusieurs centaines. La kurtosis corrigée est dans 23 cas sur 24 positive (à l'exception de WMT où elle vaut -0.11) ce qui signifie que les distributions des séries de log-volatilité réalisée sont légèrement plus piquées que celle d'une loi normale. Au sujet de la skewness, elle est positive dans 20 cas sur 24 et inférieure à 0.6 dans 19 cas sur 20. Cela signifie que la distribution a des queues asymétriques qui s'étendent vers les x positifs.

Puisque nous avons expliqué précédemment que le rejet de l'hypothèse de normalité était non concluante en raison de la sous-estimation de l'écart-type asymptotique de la kurtosis et

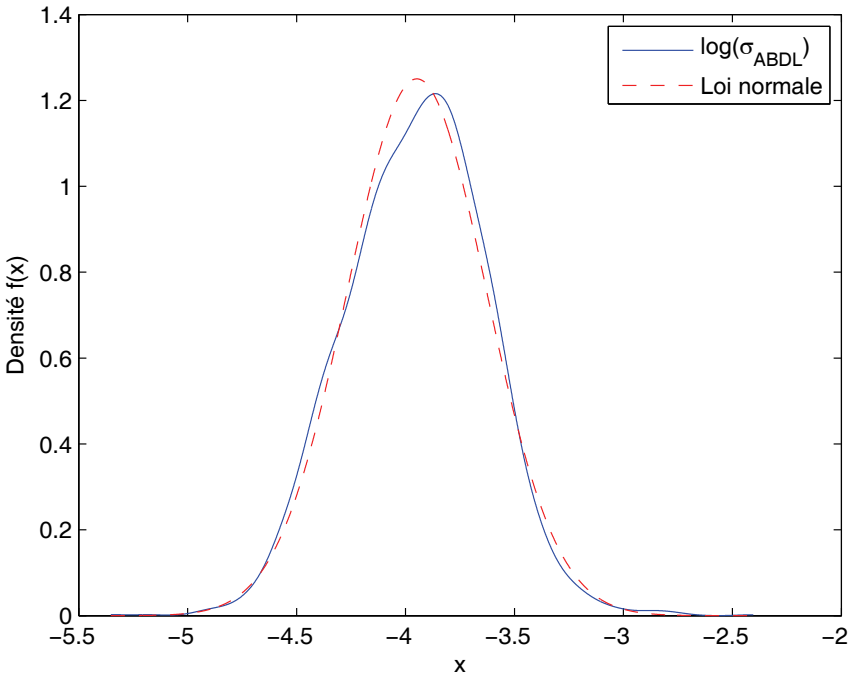

FIGURE 3.4 – Densité de la loi de la log-volatilité réalisée journalière de l'action Coca-Cola calculée sur une période de 10 ans, allant du 01/01/1994 au 31/12/2003 au moyen de l'estimateur par noyau.

de la skewness, nous montrons sur la figure 3.5 la densité de la log-volatilité réalisée et celle de la loi normale la plus proche dans le cas d'un rejet de l'hypothèse de normalité. Globalement, nous remarquons que cette figure, consacrée à l'action Fidelity National Financial, est similaire à la figure 3.4 concernant l'action Coca-Cola où l'hypothèse de normalité avait été acceptée.

En conclusion, nous pouvons affirmer deux choses. Premièrement, la statistique de Jarque-Bera modifiée est très sévère pour juger de la normalité de série temporelles à mémoire longue. Et deuxièmement, on constate graphiquement que d'une manière générale, la densité de la log-volatilité réalisée ne s'éloigne pas trop d'une densité de loi normale.

Normalité des rendements journaliers normalisés par la volatilité réalisée journalière

A présent, nous allons vérifier un dernier fait stylisé : la normalité des rendements journaliers normalisés par la volatilité réalisée journalière, définis par

$$r_j^* = \frac{r_j}{\hat{\sigma}_j}, \tag{3.34}$$

	$\bar{m}$	$\bar{\sigma}$	S	K-3	JB_n^* stat.	p-value
Statistique	-	-	0	0	5.992	-
BKS	-3.80	0.435	0.32	1.23	42.9	0.00
VLO	-4.02	0.371	0.14	1.33	107.4	0
DHI	-3.48	0.469	0.39	0.20	7.2	0.03
LEN	-3.84	0.389	0.40	4.09	652.9	0
TCB	-4.29	0.438	0.07	0.42	1.8	0.41
NYB	-3.88	0.529	-0.32	0.50	19.7	0.00
FNF	-3.78	0.410	0.09	0.49	14.2	0.00
MCHP	-2.98	0.405	0.13	0.05	0.3	0.85
WPO	-4.99	0.511	-0.44	1.65	45.0	0.00
GILD	-3.06	0.497	0.21	0.03	0.7	0.72
GM	-4.02	0.321	0.52	1.91	249.0	0
PG	-4.08	0.335	0.28	4.77	252.1	0
IBM	-4.01	0.378	0.84	2.61	254.7	0
CSCO	-3.36	0.442	0.46	0.28	2.9	0.24
MRK	-3.96	0.340	0.32	0.41	7.1	0.03
KO	-3.95	0.319	-0.01	0.35	1.4	0.50
AIG	-4.17	0.347	0.37	1.15	28.6	0.00
INTC	-3.54	0.378	0.58	1.05	21.1	0.00
WMT	-3.70	0.427	-0.28	-0.11	1.0	0.61
C	-3.74	0.348	0.28	2.09	69.4	0.00
PFE	-3.88	0.329	0.54	0.92	26.3	0.00
XOM	-4.12	0.306	0.43	1.03	22.7	0.00
MSFT	-3.74	0.348	0.41	0.97	24.3	0.00
GE	-3.91	0.326	0.36	0.84	14.9	0.00

TABLE 3.3 – Statistiques de la série de la log-volatilité réalisée journalière du 01/01/1994 au 31/12/2003 obtenue par l'estimateur $\hat{\sigma}_{ABDL}$. L'hypothèse de normalité pour un niveau de confiance à 95 % est acceptée si la statistique $JB_n < 5.99$ ou si la p-value ≥ 0.05.

où r_j est le rendement journalier de clôture à clôture défini par l'équation (3.28) et $\hat{\sigma}_j$ est l'estimation de la volatilité journalière. Les résultats obtenus à partir de la série des estimations calculées grâce à l'estimateur $\hat{\sigma}_{ABDL}$ sont résumées dans le tableau 3.4.

Les résultats sont meilleurs que ceux de la log-volatilité réalisée. En effet, l'hypothèse de normalité avec un niveau de confiance à 95 % est acceptée pour 9 actions (3 de moyenne capitalisation et 6 de grande capitalisation). En outre, elle est rejetée de très peu (la p-value est ≥ 0.1 pour 9 autres autres actions (là encore 3 de moyenne capitalisation et 6 de grande capitalisation). D'un point de vue graphique, ceci est illustré par la figure 3.6 où sont tracées la densité de la loi des rendements de l'action TCB normalisés par sa volatilité réalisée ainsi que celle de la loi normale la plus proche. Effectivement, les deux courbes sont quasiment confondues.

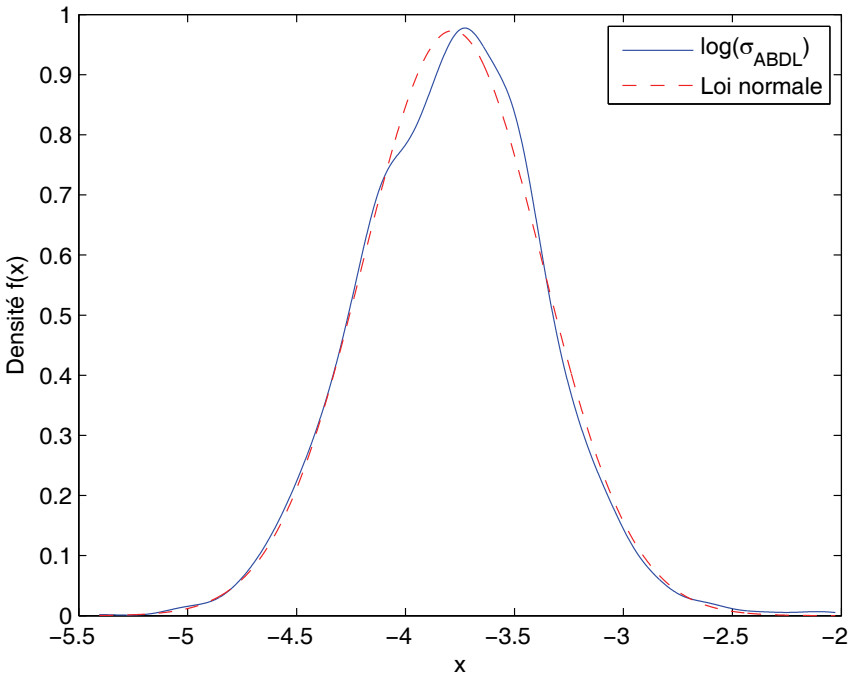

FIGURE 3.5 – Densité de la loi de la log-volatilité réalisée journalière de l'action Fidelity National Financial calculée sur une période de 10 ans, allant du 01/01/1994 au 31/12/2003 à partir de l'estimateur par noyau.

3.2.2 Calibrage du modèle sur l'intervalle de temps tout entier

Soulignons que l'estimation de la densité de la variable d'hétérogénéité φ nous fournit un accès au paramètre d qui caractérise le comportement de la mémoire longue de la série temporelle de la log-volatilité. Il est alors intéressant de comparer les valeurs de d obtenues avec notre modèle à celles obtenues par d'autres méthodes telles que la méthode R/S définie par Hurst (1951) et la méthode de régression introduite par Geweke et Porter-Hudak (1983). Notons que la première méthode a été affinée par Mandelbrot (1972, 1975), Mandelbrot et Taqqu (1979) et plus tard par Lo (1991). Hélas, la généralisation de Lo fait intervenir un paramètre supplémentaire qui rend l'estimation de d très sensible (Teverovsky, Taqqu, et Willinger 1999). Par conséquent, nous avons seulement recours à la statistique du R/S classique et celle de Geweke et Porter-Hudak.

Dans cette section, les résultats graphiques obtenus contiendront, d'une part, les fonctions d'autocorrélation, empirique et estimée, de la série de la log-volatilité réalisée, et d'autre part, la densité du paramètre d'hétérogénéité du modèle ainsi que son intervalle de confiance à 95%, point par point. On pourra déduire la qualité de l'estimation, premièrement grâce aux fonctions d'autocorrélation (si elles sont confondues le résultat est bon, par contre si elles sont éloignées ce n'est plus le cas), et deuxièmement grâce à l'amplitude de l'intervalle de confiance (plus celui-ci sera étroit, meilleure sera l'estimation). Ce n'est évidemment pas un

	$\bar{m}(10^{-2})$	$\bar{\sigma}$	S	K-3	JB_n^* stat.	p-value
Statistique	-	-	0	0	5.992	-
BKS	0.37	1.01	0.18	-0.06	13.2	0.00
VLO	0.49	1.05	0.05	0.20	5.5	0.07
DHI	3.07	0.91	0.10	0.26	10.5	0.01
LEN	4.49	1.00	0.15	-0.01	9.1	0.01
TCB	5.72	1.05	0.09	-0.11	4.7	0.10
NYB	5.38	0.75	0.09	1.16	141.5	0
FNF	2.27	0.89	0.02	0.11	1.4	0.49
MCHP	1.46	0.76	0.11	0.57	38.7	0.00
WPO	4.15	1.28	0.04	-0.25	7.3	0.03
GILD	1.50	0.79	0.07	0.54	32.0	0.00
GM	-0.27	1.00	0.12	-0.04	6.0	0.05
PG	5.47	0.84	0.11	-0.02	5.3	0.07
IBM	2.95	0.97	0.05	-0.27	8.4	0.02
CSCO	5.18	0.87	0.05	0.21	5.9	0.05
MRK	3.76	0.87	0.15	0.04	9.3	0.01
KO	2.81	0.80	0.13	0.60	43.9	0.00
AIG	4.01	1.00	0.13	-0.09	8.4	0.02
INTC	5.20	0.91	0.10	-0.10	5.3	0.07
WMT	2.37	0.81	0.11	0.64	47.6	0.00
C	3.35	0.83	0.04	-0.14	2.6	0.27
PFE	4.68	0.84	0.05	-0.18	4.5	0.11
XOM	3.54	0.83	0.03	0.08	1.1	0.57
MSFT	4.68	0.89	0.10	-0.17	7.7	0.02
GE	3.24	0.84	0.11	0.12	6.2	0.05

TABLE 3.4 – Statistiques de la série des rendements journaliers normalisés par la volatilité réalisée journalière (estimateur ABDL) en tenant compte de la volatilité et du rendement overnight, du 01/01/1994 au 31/12/2003. L'hypothèse de normalité pour un niveau de confiance à 95 % est acceptée si la statistique $JB_n < 5.99$ ou si la p-value ≥ 0.05.

test de validité du modèle, juste une indication.

A propos de l'intervalle de confiance, nous nous trouvons dans un cas particulier. En règle générale, pour une loi normale, on a

$$0.95 = \int_{\mu - q_{\alpha/2}\sigma}^{\mu + q_{\alpha/2}\sigma} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}y^2} dy, \tag{3.35}$$

avec $\alpha = 0.05$ et $q_{\alpha/2} = 1.96$. L'intervalle de confiance à 95% est donné par $[\mu - q_{\alpha/2}\sigma, \mu + q_{\alpha/2}\sigma]$. Ici, puisqu'il s'agit de densité, nous avons une contrainte supplémentaire : une densité est toujours positive. Par conséquent la borne inférieure de l'intervalle de confiance doit l'être également. Lorsque $\mu - q_{\alpha/2}\sigma < 0$, il faut donc remplacer $\mu - q_{\alpha/2}\sigma$ par 0 et $\mu + q_{\alpha/2}\sigma$ par B_{sup} dans l'équation (3.35) et déterminer la valeur de B_{sup}. On obtient

$$B_{sup} = \mu + \sigma\phi^{-1}\left(0.95 + \phi\left(-\frac{\mu}{\sigma}\right)\right), \tag{3.36}$$

ou $\phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}y^2} dy$. Ainsi l'intervalle de confiance à 95% sera donné par

$$\left[\max(0, \mu - q_{\alpha/2}\sigma), \max(B_{sup}, \mu + q_{\alpha/2}\sigma)\right]. \tag{3.37}$$

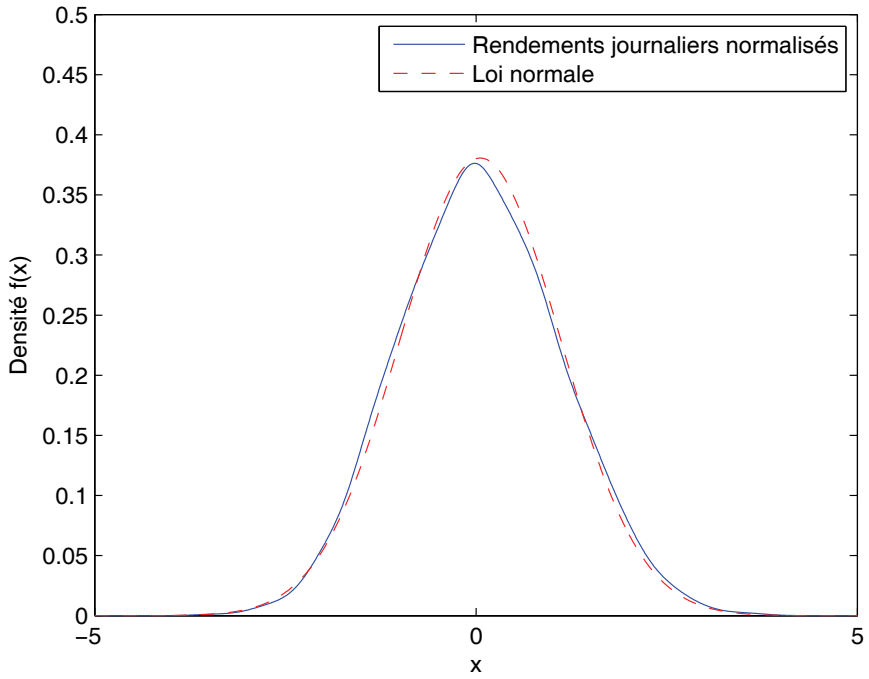

FIGURE 3.6 – Densité de la loi des rendements journaliers normalisés par la volatilité réalisée journalière de l'action TCF Financial calculée sur une période de 10 ans, allant du 01/01/1994 au 31/12/2003 à partir de l'estimateur par noyau.

Résultats graphiques obtenus à partir de deux actions

Nous commençons par présenter les résultats graphiques obtenus pour une moyenne capitalisation américaine : Fidelity National Financial (voir figure 3.7). Sur le graphique du bas, nous observons que la fonction d'autocorrélation estimée colle très bien à la fonction d'autocorrélation empirique. Quant à la densité de la variable d'hétérogénéité φ tracée sur le graphique du haut, nous constatons que l'intervalle de confiance à 95% est également bon (autrement dit suffisamment étroit). De manière générale, les intervalles de confiance sont aussi bons pour presque toutes les actions. Cette forme montrant 3 masses distinctes (une proche de -1, une autre proche de 1 et la dernière autour de 0) est représentatif de la moitié des actions, l'autre moitié étant décrite ci-après. Nous avons également tracé sur ce graphique $E[\psi|\varphi]$ afin de nous rendre compte de l'importance relative de la réalisation passée de la log-volatilité par rapport à l'anticipation passée dans l'anticipation présente de l'investisseur.

En termes de comportement d'investisseur, cela signifie qu'il existe principalement 3 catégories d'agents. Premièrement, lorsque l'on regarde au centre de la densité, nous pouvons conclure que la plupart des agents basent leur anticipation sur les flux d'informations (publique et privée) et sur la réalisation passée de la volatilité. Ils négligent leur anticipation passée et le poids de la réalisation passée est négligeable dans leur décision. D'autres croient

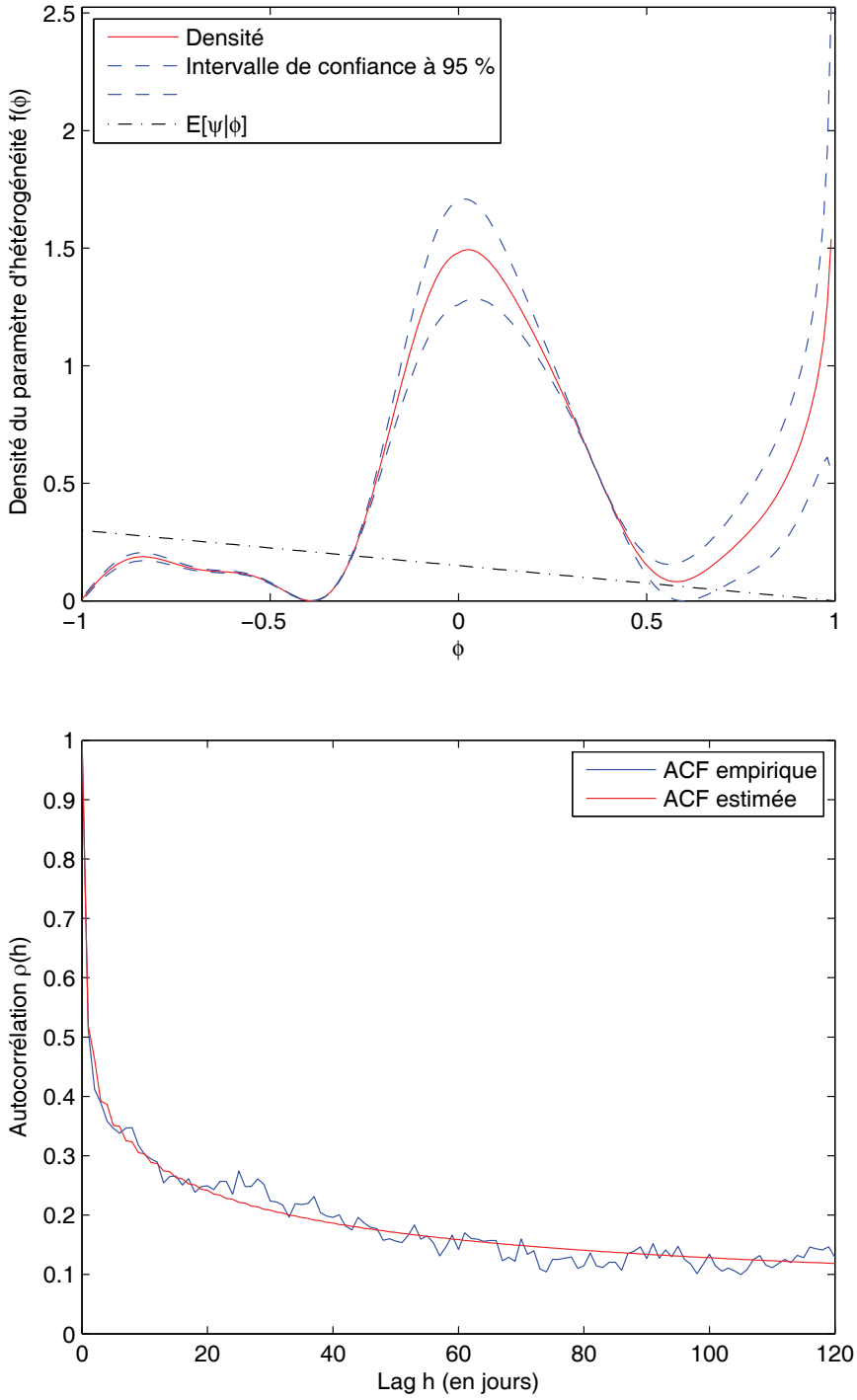

FIGURE 3.7 – Estimation de la densité de la log-volatilité réalisée avec son intervalle de confiance à 95 % (graphique du haut) et de la fonction d'autocorrélation à horizon 120 jours (graphique du bas) de Fidelity National Financial.

fermement en l'anticipation qu'ils ont établie la veille, qui est reliée à la partie divergente de la fonction de la densité au voisinage de 1, et négligent les autres critères. Enfin, un petit nombre d'agents, représentés par la petite masse entre -1 et -0.5, prennent systématiquement le contre-pied de leur anticipation passée et tiennent aussi compte de la réalisation passée. C'est normal que cette dernière catégorie de comportement soit minoritaire parce que l'on ne s'attend pas a priori à ce que les investisseurs aient un comportement erratique.

A présent, nous montrons l'allure de la densité de la variable d'hétérogénéité que nous obtenons pour l'autre moitié des actions, en prenant l'exemple d'une grande capitalisation américaine : Microsoft (ci-après MSFT, voir figure 3.8). Au lieu de distinguer trois masses, nous n'en voyons plus que deux.

Les agents confiants dans leur anticipation passée sont toujours présents (matérialisés par la divergence au voisinage de 1). Le changement réside dans l'autre masse qui englobe un certains nombre de comportements différents. On y retrouve, comme pour FNF des agents prenant le contre-pied de leur anticipation passée et prenant en compte la réalisation passée, dont la proportion a augmenté (vers -1), et ceux, plus rationnels, qui se basent à la fois sur l'information dont ils disposent et sur la réalisation passée de la volatilité réalisée (autour de 0). Le changement majeur réside dans l'apparition d'une masse de taille importante sur $[-1, 0]$ avec une pic en -0.5. Cela correspond à un grand nombre d'agents qui prennent en compte à la fois les informations et la réalisation passée, mais également le contraire de leur anticipation passée, et ceci de manière pondérée. C'est une situation plutôt logique lorsqu'un agent réalise que son anticipation passée est éloignée de la réalisation passée : il prend l'opposé de son anticipation passée et fait plus attention aux informations nouvelles.

Ces deux allures typiques de la distribution de l'hétérogénéité ne semblent pas reliées à la taille des entreprises, pas plus qu'à un secteur industriel spécifique. Toutefois, le petit nombre d'actions par industrie dans notre base de données ne nous permet pas, hélas, de tirer des conclusions définitives.

Comparaison des valeurs de d obtenues par notre modèle avec celles obtenues par d'autres méthodes

Les résultats des estimations du paramètre de mémoire longue d obtenus à l'aide de notre modèle et par les deux méthodes semi-paramétriques : la méthode de Geweke et Porter-Hudak (1983) (notée ci-après GPH) et la méthode classique R/S de Hurst (1951). Elles sont détaillées dans l'annexe 3.B. Ces résultats sont résumés dans le tableau 3.5. Il parait évident que, d'une part chaque estimateur conduit à conclure que la mémoire longue est présente dans quasiment toutes les séries de la log-volatilité réalisée des rendements journaliers d'actions, alors que d'autre part, chaque estimateur produit pour une même action des valeurs différentes de d. La plupart des valeurs obtenues par notre modèle sont comprises entre 0.2 et 0.4 et quelques unes sont négatives, alors qu'avec l'estimateur GPH elles sont comprises entre 0.3 et 0.45. Enfin, le paramètre de mémoire longue obtenu avec la méthode classique R/S est majoritairement compris entre 0.40 et 0.55.

Il est bien reconnu qu'au delà de $d = 0.5$ la série n'est plus stationnaire (au second ordre).

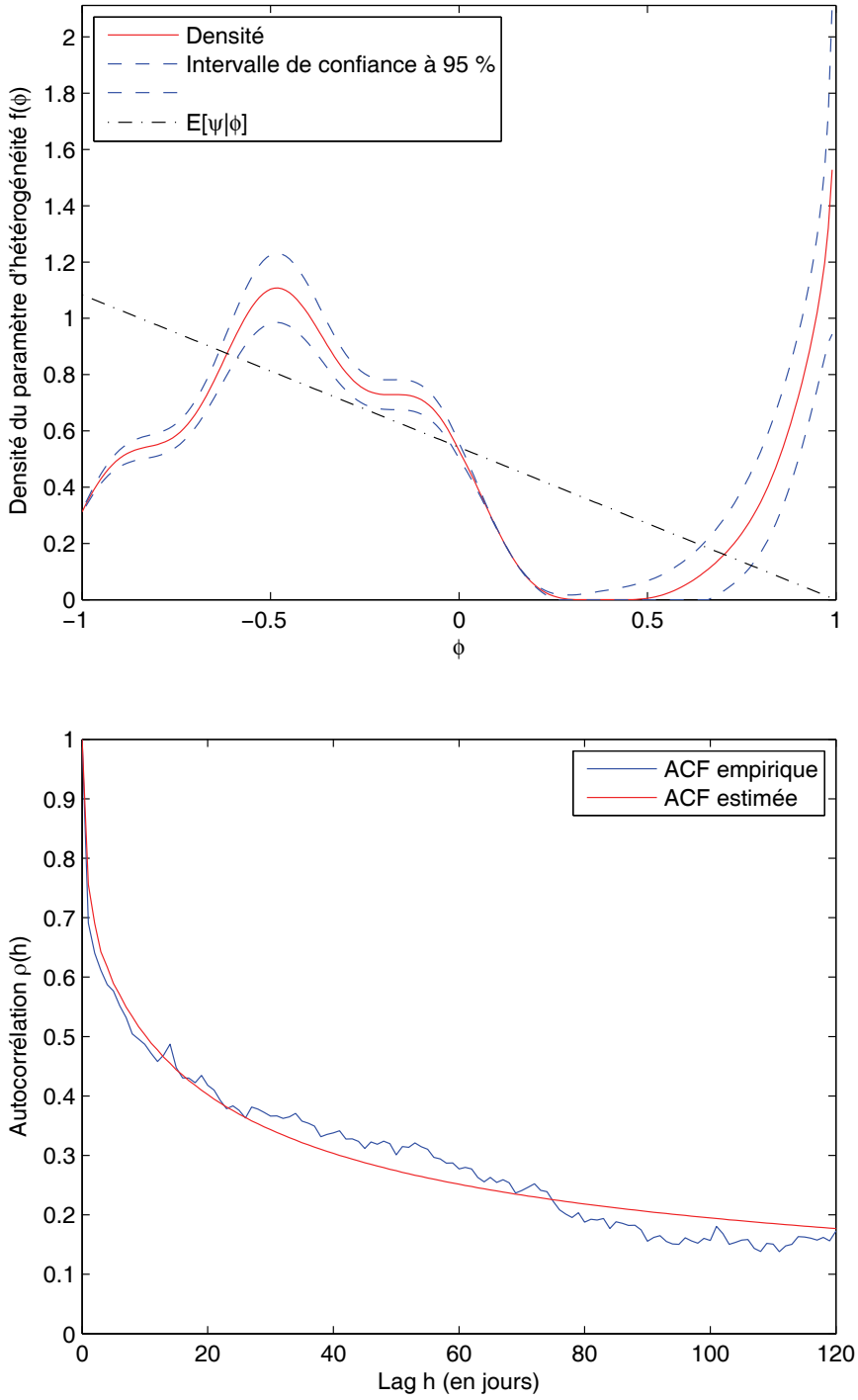

FIGURE 3.8 – Estimation de la densité de la log-volatilité réalisée avec son intervalle de confiance à 95 % (graphique du haut) et de la fonction d'autocorrélation à horizon 120 jours (graphique du bas) de Microsoft.

Action	$\hat{\sigma}_{ABDL}$ d	Action	$\hat{\sigma}_{ABDL}$ d
BKS	0.19 (0.04)	GM	-0.30 (0.04)
VLO	0.23 (0.04)	PG	0.22 (0.05)
DHI	0.34 (0.04)	IBM	0.21 (0.05)
LEN	0.31 (0.05)	CSCO	0.33 (0.06)
TCB	0.31 (0.04)	MRK	0.22 (0.05)
NYB	0.00 (0.06)	KO	0.16 (0.05)
FNF	0.25 (0.04)	AIG	0.10 (0.03)
MCHP	0.19 (0.04)	INTC	0.23 (0.06)
WPO	0.45 (0.04)	WMT	0.31 (0.04)
GILD	0.37 (0.06)	C	-0.02 (0.04)
		PFE	0.16 (0.05)
		XOM	-0.09 (0.04)
		MSFT	0.21 (0.03)
		GE	0.04 (0.06)

TABLE 3.5 – Estimations du paramètre de mémoire longue, d, par notre modèle, pour toutes les actions (les moyennes capitalisation à gauche et les grandes capitalisations à droite) sur l'ensemble de la période allant du 01/01/1994 au 31/12/2003. Les écarts-types sont fournis entre parenthèses.

Comme les estimations par la méthode de Hurst sont supérieures à (mais proches de) 0.5, nous pouvons dire que cela est dû à l'incertitude de la méthode et, peut-être, à la difficulté de la mettre en œuvre efficacement.

Par ailleurs, les différences remarquées entre ces résultats pourraient être expliquées par la difficulté à les paramétrer. Par exemple, lorsque l'on utilise la méthode GPH, en faisant appel à l'équation (3.B.9) de l'annexe 3.B, nous devons fixer une valeur pour m. Dans notre cas, la taille de nos séries est d'environ $N \approx 2500$ observations. Dans la littérature, on trouve $m = \sqrt{N}$ ou $m = N/4$ par exemple. Choisir l'un ($m \simeq 50$) plutôt que l'autre ($m =\simeq 625$) entraîne une réelle différence dans les estimations. Pour décider de la valeur à attribuer à m, nous avons simulé un processus à mémoire longue pour différentes valeurs de $d \in \{0.05, 0.1, 0.15, \ldots, 0.45\}$. Puis nous avons estimé d par la méthode GPH pour un m fixé. Nous avons réitéré ce procédé une centaine de fois pour chaque valeur de d. Nous en avons déduit l'estimation moyenne de d pour un m fixé. Enfin, nous avons recommencé le processus pour différentes valeurs de m et conclu que $m = N/2$ est la meilleure valeur.

3.3 Étude de l'effet d'éclatement de la bulle

L'étude de l'effet bulle est motivée par le fait que nous serons capables d'obtenir des informations sur le comportement des agents durant une période prospère ou supposée comme telle par le marché (ou de gonflement de la bulle spéculative) où la valeur des actions augmente, et durant une période où les prix chutent (dégonflement de la bulle). En particulier nous serons capables de répondre à la question : "Est-ce que les agents se comportent différemment avant et après l'éclatement d'une bulle spéculative ?"

	$\hat{\sigma}_{ABDL}$		
action	d	d_{gph}	d_{hurst}
BKS	0.19 (0.04)	0.36 (0.02)	0.44 (0.04)
VLO	0.23 (0.04)	0.32 (0.02)	0.43 (0.03)
DHI	0.34 (0.04)	0.34 (0.02)	0.44 (0.05)
LEN	0.31 (0.05)	0.34 (0.02)	0.50 (0.05)
TCB	0.31 (0.04)	0.35 (0.02)	0.53 (0.03)
NYB	0.00 (0.06)	0.33 (0.02)	0.37 (0.04)
FNF	0.25 (0.04)	0.34 (0.02)	0.28 (0.04)
MCHP	0.19 (0.04)	0.45 (0.02)	0.42 (0.06)
WPO	0.45 (0.04)	0.28 (0.02)	0.46 (0.05)
GILD	0.37 (0.06)	0.44 (0.02)	0.42 (0.05)
GM	-0.30 (0.04)	0.30 (0.02)	0.30 (0.03)
PG	0.22 (0.05)	0.41 (0.02)	0.53 (0.04)
IBM	0.21 (0.05)	0.36 (0.02)	0.44 (0.05)
CSCO	0.33 (0.06)	0.47 (0.02)	0.51 (0.05)
MRK	0.22 (0.05)	0.37 (0.02)	0.49 (0.05)
KO	0.16 (0.05)	0.41 (0.02)	0.55 (0.06)
AIG	0.10 (0.03)	0.44 (0.02)	0.48 (0.05)
INTC	0.23 (0.06)	0.38 (0.02)	0.53 (0.06)
WMT	0.31 (0.04)	0.42 (0.02)	0.46 (0.08)
C	-0.02 (0.04)	0.40 (0.02)	0.51 (0.05)
PFE	0.16 (0.05)	0.40 (0.02)	0.55 (0.06)
XOM	-0.09 (0.04)	0.41 (0.02)	0.49 (0.06)
MSFT	0.21 (0.03)	0.43 (0.02)	0.53 (0.06)
GE	0.04 (0.06)	0.44 (0.02)	0.50 (0.08)

TABLE 3.6 – Estimations du paramètre de mémoire longue, d, par notre modèle et deux méthodes semi-paramétriques : Geweke et Porter-Hudak (GPH) et Hurst, pour toutes les actions sur l'ensemble de la période allant du 01/01/1994 au 31/12/2003. Les écarts-types sont donnés entre parenthèses.

Afin de pouvoir procéder aux estimations, nous avons sélectionné 5 grandes capitalisations parce que le phénomène de bulle est très prononcé dans la courbe des prix, ce qui n'est pas le cas pour les moyennes capitalisations. Nous divisons la série de la log-volatilité des rendements de chaque action lorsque le prix maximum est atteint en deux sous-séries. La première est la phase précédant l'éclatement de la bulle (ci-après appelée période pré-bulle), et la deuxième celle suivant l'éclatement de la bulle (ci-après appelée période post-bulle). Nous illustrerons cette étude au moyen, d'une part, d'une entreprise dite de nouvelle technologie : Cisco Systems (ci-après CSCO) qui a subi de plein fouet la crise boursière résultant de l'éclatement de la bulle internet en mars-avril 2000, et d'autre part, d'une compagnie de distribution : Coca-Cola (ci-après KO).

3.3.1 Étude d'une action de nouvelle technologie : Cisco Systems

Nous avons tracé, sur la figure 3.9, le cours de l'action Cicsco Systems entre le 01/01/1994 et le 31/12/2003. On observe principalement deux phases : la première de 1995 au 27 mars

2000 et la seconde du 28 mars 2000 au 31 décembre 2003 [37]. La première phase correspond aux années de prospérité boursière où le S&P500 a gagné consécutivement plus de 20 % par an [38]. Le cours de cette action a tout de même été momentanément affecté fin 1998 par la crise obligataire (l'action perdant brutalement environ 40 % de sa valeur), puis a repris de plus belle sa progression : + 800 % en 17 mois (entre novembre 1998 et mars 2000. L'apogée se situe le 27 mars 2000. Dans la deuxième phase, synonyme de crise boursière, en un peu plus d'un an, entre avril 2000 et avril 2001, l'action Cisco Systems a dévissé de 81 %, pour retrouver son niveau du début 1999. Ce n'est qu'au début de l'année 2003 que le cours a tendance à remonter, ce qui est considéré comme étant la fin de la crise boursière.

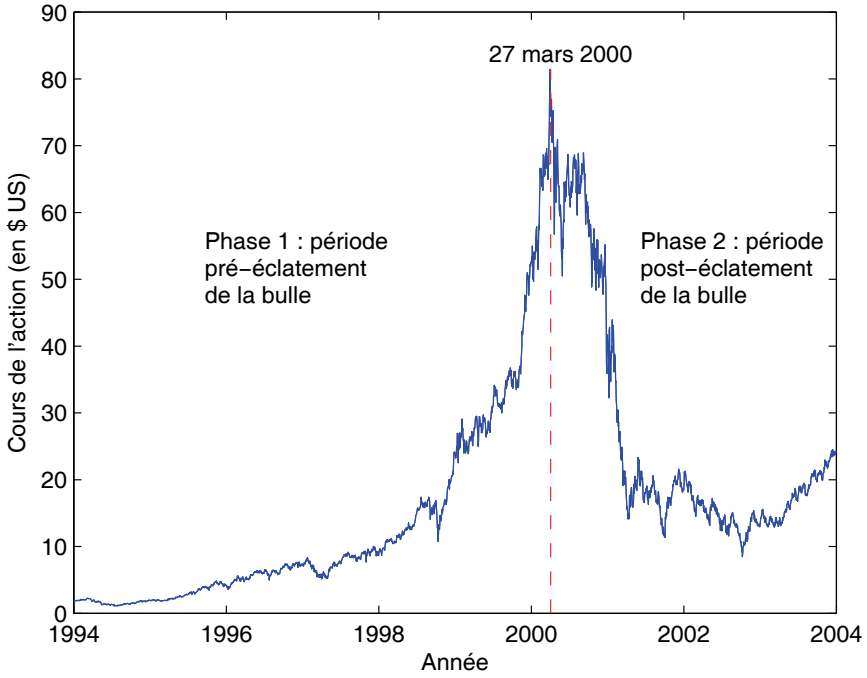

FIGURE 3.9 – Cours de l'action Cisco Systems entre le 01/01/1994 et le 31/12/2003.

Sur la figure 3.10, nous avons tracé, en mode linéaire (graphique du haut) et en mode log-log (graphique du bas), les fonctions d'autocorrélation empiriques des séries de la log-volatilité des rendements sur trois périodes : la période globale, la phase pré-bulle et la phase post-bulle. En observant le graphique en mode linéaire, il semblerait que la décroissance de la fonction d'autocorrélation de la période pré-bulle soit plus lente que celle de la période globale, qui elle-même est plus grand que celle de la période post-bulle. En terme de paramètre de mémoire longue, on s'attend donc à trouver une valeur de d plus grande pour la période pré-bulle, plus faible pour la période post-bulle, et intermédiaire pour la période globale ce

37. En réalité, il y a trois phases, la dernière commençant au début de l'année 2003 et synonyme de la fin de la crise boursière. Toutefois, pour des contraintes d'optimisation, nous négligeons cette troisième phase et incluons les données de l'année 2003 dans la phase 2.

38. Plus précisément : 34 % en 1995, 20 % en 1996, 31 % en 1997, 26 % en 1998 et 20 % en 1999 d'après les chiffres de Shiller (2000).

qui est logique puisque d_{global} apparaît comme la moyenne des coefficients d des deux sous-périodes. Puisque ce paramètre est liée à la pente de la fonction d'autocorrélation en mode log-log, nous allons confirmer ces impressions à l'aide du graphique du bas. En effet, la pente la plus raide correspond à la période post bulle. Par contre la différence de pente n'est pas si nette que cela pour la période globale et la période pré-bulle. Par conséquent, le paramètre de mémoire longue devrait-être assez proche.

Complétons ces impressions visuelles par les densités du paramètre d'hétérogénéité et les fonctions d'autocorrélation estimées sur les deux sous-périodes (voir figures 3.11 et 3.12). Nous avons tracé sur le graphique du haut la fonction de la densité de la variable d'hétérogénéité ainsi que son intervalle de confiance à 95%, et sur le graphique du bas les fonctions d'autocorrélation, estimée et empirique, à un horizon de 120 jours.

En ce qui concerne la période pré-bulle (voir figure 3.11), nous constatons (graphique du bas) que les fonctions d'autocorrélation, estimée et empirique, sont proches l'une de l'autre. De plus, l'intervalle de confiance à 95% de la fonction densité est très étroit. Nous en déduisons que l'optimisation a fonctionné, en d'autres termes le minimum global a bien été atteint.

L'allure de la densité du paramètre d'hétérogénéité (graphique du haut) est similaire à celle obtenue avec l'action Microsoft (voir figure 3.8) et montre une mémoire longue forte (divergence rapide au voisinage de 1. En somme, nous obtenons deux masses : une proche de 1 qui représente les agents auto-référents croyant uniquement en leur anticipation passée à moins qu'une information de très grande ampleur ne fasse irruption, et l'autre, s'étalant entre -1 et 0.2, qui englobe plusieurs catégories de comportements d'agents, mais qui ont toutefois un point commun : ils ne reproduisent pas leur anticipation passée. Certains n'y font pas référence et se focalisent sur les informations leur parvenant et sur la réalisation de la log-volatilité réalisée de la veille, alors que d'autres, moins rationnels, se basent sur l'opposé de leur anticipation passée combinée avec les informations et la réalisation passée.

Quant à la période post-bulle, nous commençons par constater que les fonctions d'autocorrélation empirique et estimée de la log-volatilité sont très proches l'une de l'autre et que l'intervalle de confiance à 95% de la densité de la variable d'hétérogénéité est bon (voir figure 3.12).

Ici, la situation est légèrement différente : premièrement, la valeur du paramètre de mémoire longue d est plus petite. Concrètement, cela signifie que la proportion d'agents reproduisant leur anticipation passée est moindre. Afin de comparer les comportements des investisseurs et d'observer d'éventuelles évolutions entre les deux phases, nous avons superposé les densités du paramètre d'hétérogénéité, sans leur intervalle de confiance pour une meilleure lisibilité.

La figure 3.13 nous aide à comparer le comportement des agents durant les deux phases. Nous observons qu'après l'éclatement de la bulle, la masse située entre -1 et 0.2 change. Tout d'abord, elle est comprimée entre -1 et 0, puis elle semble se diviser en deux masses. Par conséquent, d'une part la proportion d'agents prenant le contre-pied de leur anticipation passée augmente (ces agents peuvent reconnaître qu'ils ont fait une erreur, par exemple) et d'autre part la proportion d'agents qui utilisent principalement les informations et la valeur

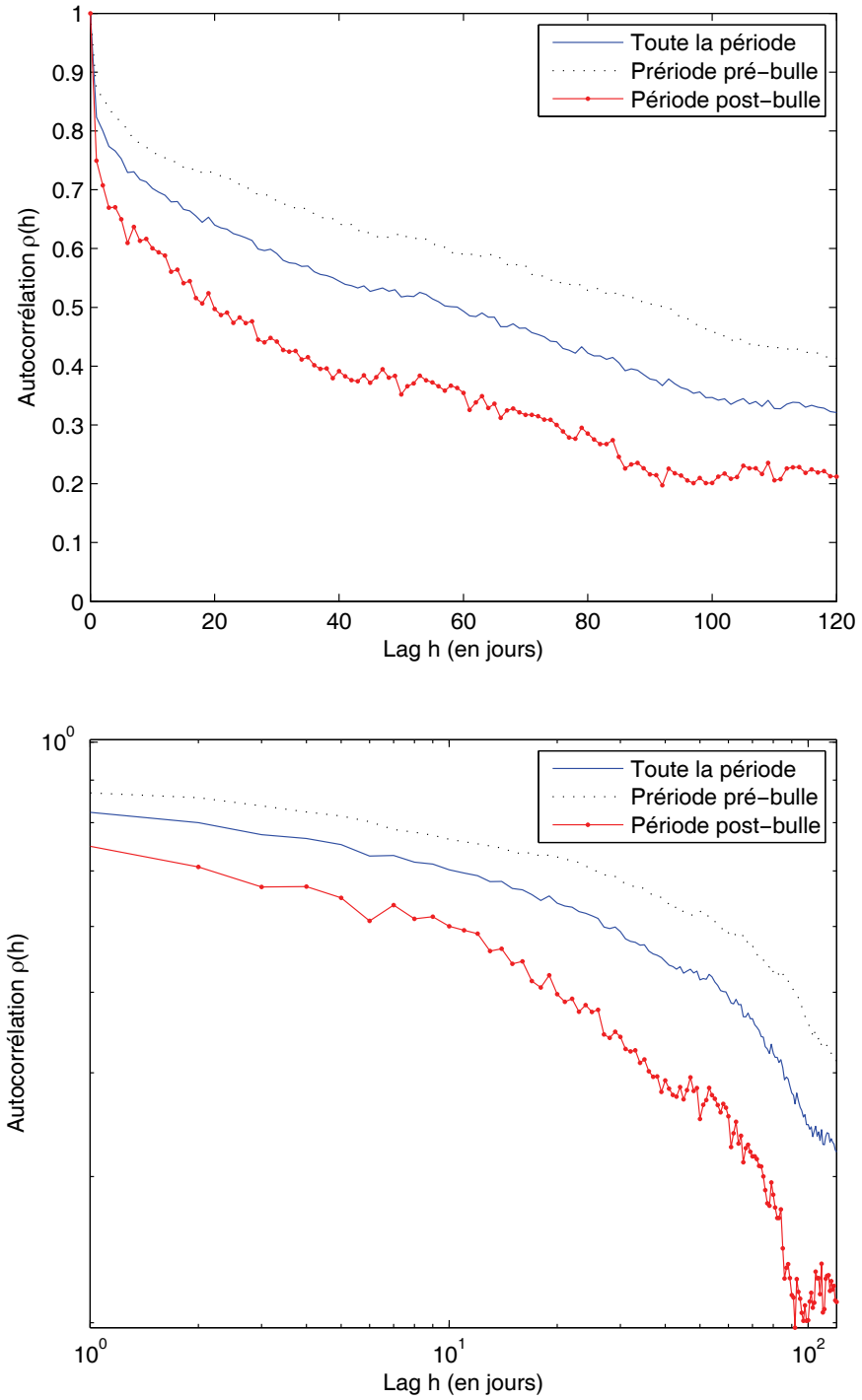

FIGURE 3.10 – Fonctions d'autocorrélation empiriques de la log-volatilité réalisée de Cisco Systems en mode linéaire (graphique du haut) et en mode log-log (graphique du bas), à horizon 120 jours, calculées sur les différentes périodes considérées : globale, pré-bulle et post-bulle.

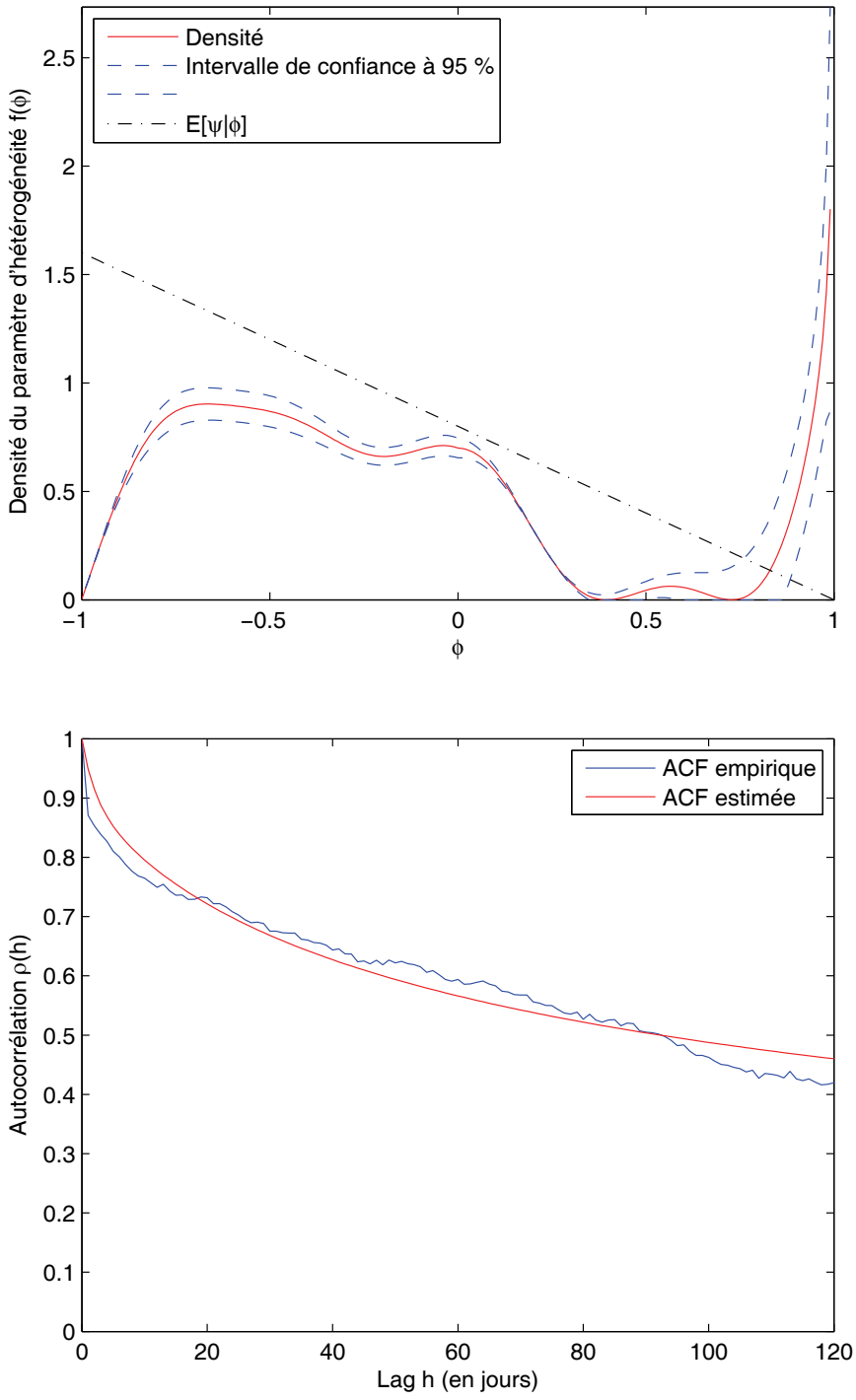

FIGURE 3.11 – Densité du coefficient d'hétérogénéité (en haut) et fonction d'autocorrélation de la log-volatilité réalisée de Cisco Systems, à horizon de 120 jours, durant la période pré-bulle.

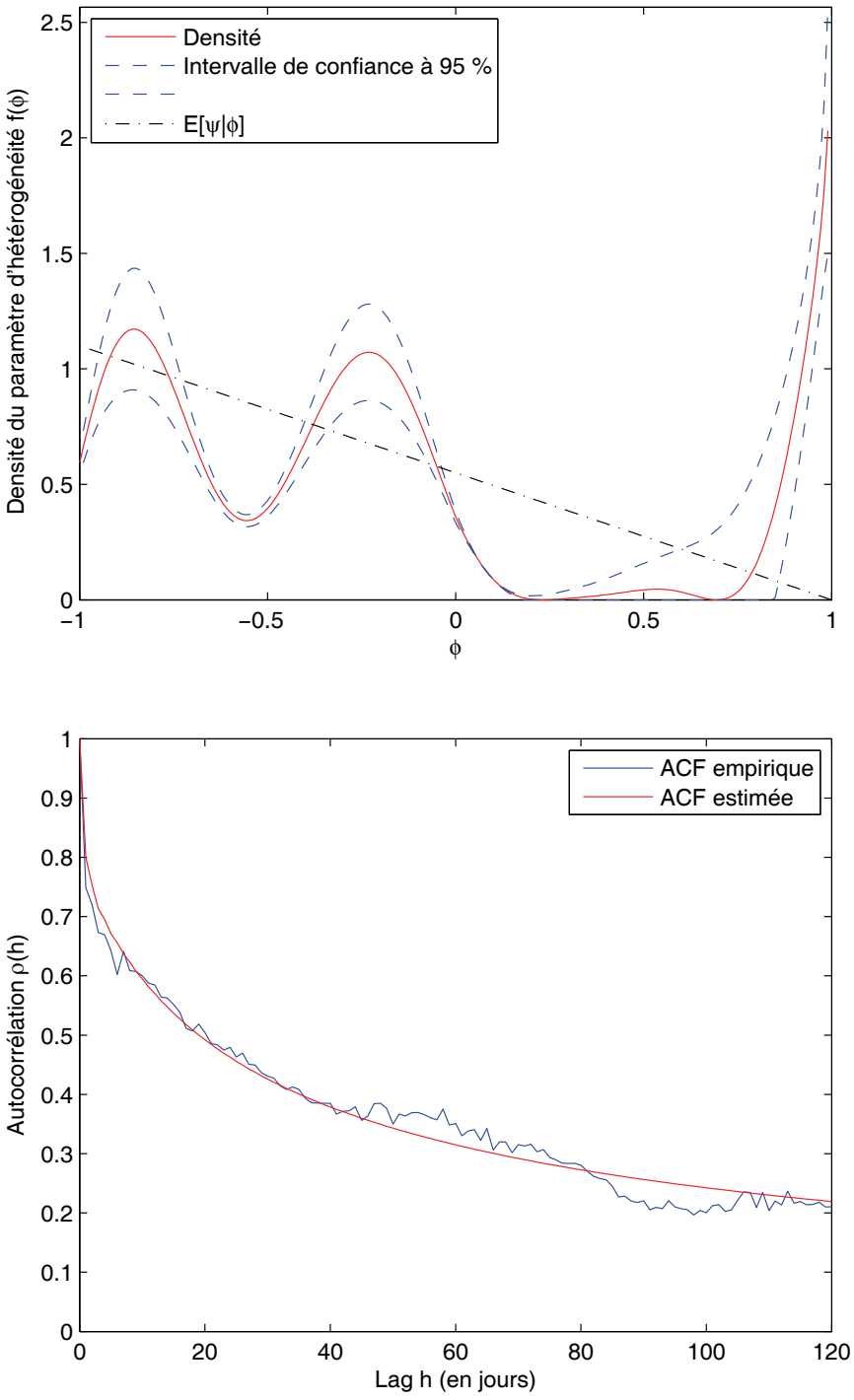

FIGURE 3.12 – Densité du coefficient d'hétérogénéité (en haut) et fonction d'autocorrélation de la log-volatilité réalisée de Cisco Systems, à horizon de 120 jours, durant la période post-bulle.

de la réalisation de la log-volatilité et le contraire de leur anticipation passée augmente aussi.

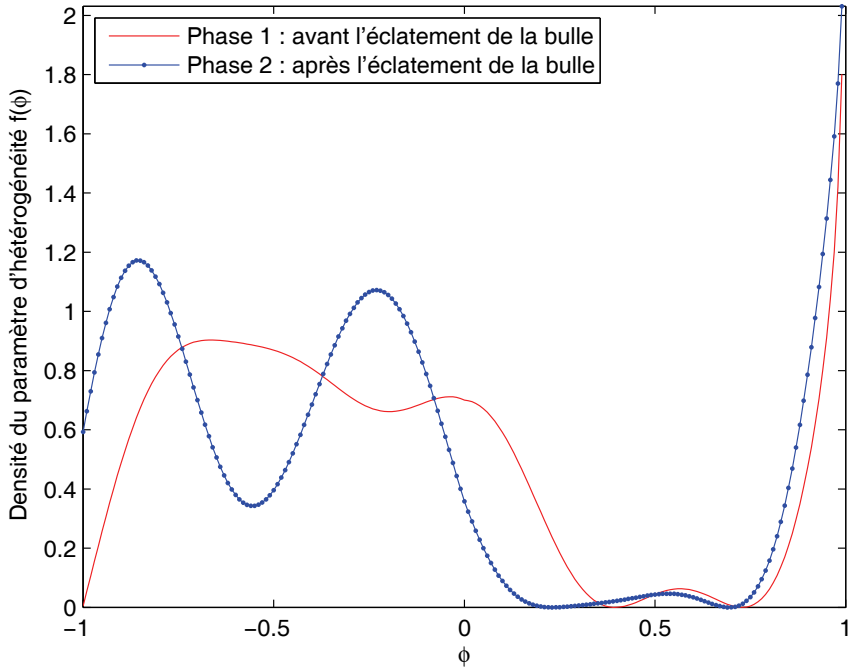

FIGURE 3.13 – Comparaison des densités du paramètre d'hétérogénéité φ de Cisco Systems estimées à partir des données des périodes pré et post-bulle.

3.3.2 Étude d'une action non technologique : Coca Cola

Nous reproduisons pour Coca-Cola les mêmes figures que pour Cisco Systems. En nous appuyant sue la figure 3.14, nous remarquons que, d'une part, la date de découpage n'est pas la même et, d'autre part, la bulle est moins prononcée. L'action Coca-Cola a été touché par la crise obligataire de fin 1998 où elle a perdu près d'un tiers de sa valeur en quelques semaines. La bulle est moins prononcée que dans le cas de CSCO.

La conséquence immédiate est montrée grâce à la figure 3.15. Les fonctions d'autocorrélation empiriques de la log-volatilité réalisée calculée sur les trois périodes considérées (totale, pré-bulle et post-bulle) sont beaucoup moins éloignées (voir graphique du haut). Nous nous attendons donc à obtenir des paramètres de mémoire longue assez proches, d'autant plus que le graphique du bas (en mode log-log) ne montrent pas de différence en terme de pente des fonctions d'autocorrélation.

Soyons, à présent, plus précis en ce qui concerne les densités et les fonctions d'autocorrélation de la log-volatilité déterminées à partir des sous-périodes (voir figures 3.16 et 3.17).

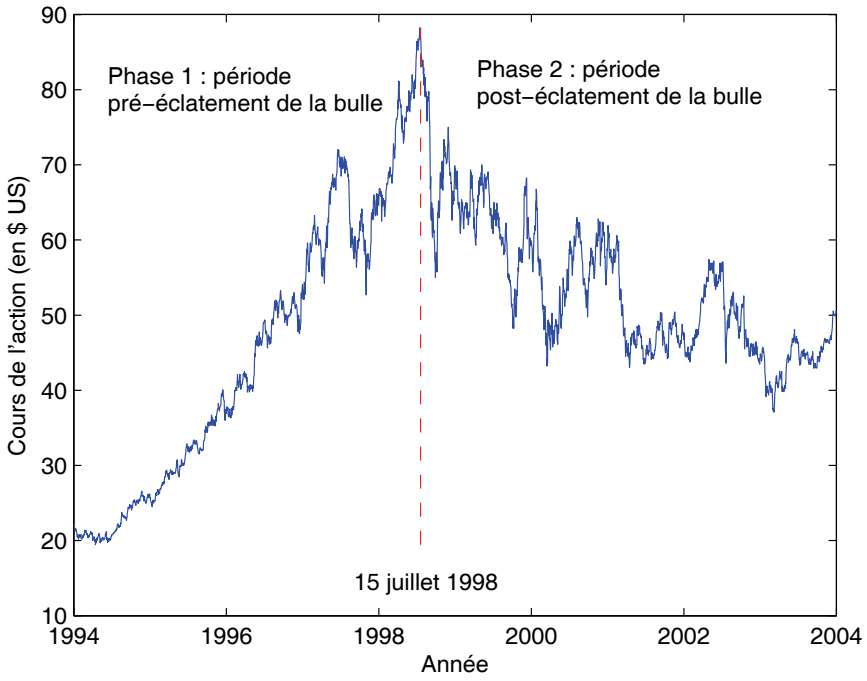

FIGURE 3.14 – Cours de l'action Coca Cola entre le 01/01/1994 et le 31/12/2003.

Pour les deux figures, remarquons que les fonctions d'autocorrélation estimées et empiriques (graphiques de droite) sont très proches, en particulier après l'éclatement de la bulle. Quant aux intervalles de confiance à 95% de la densité du paramètre d'hétérogénéité (graphiques de gauche), ils sont bons également.

Puisque l'allure des densités semble similaire (mais est-ce une surprise lorsque l'on se rappelle que les fonctions d'autocorrélation empirique étaient proches ?), superposons-les sur un même graphique pour mieux les comparer (voir figure 3.18). En observant la divergence au voisinage de 1, nous pouvons en déduire que la valeur du paramètre de mémoire longue est quasiment la même. Ensuite, si l'on regarde tout l'intervalle $[-1, 0.6]$, nous constatons une contraction globale et un recentrage vers 0. Après la bulle, les agents, dans l'ensemble, portent plus attention à l'information (exception faite du voisinage immédiat de 0). En particulier, la masse qui se trouvait entre 0.2 et 0.7 et qui représentait les agents qui croyaient à la fois en la continuité du marché et aux informations s'est recentrée en 0. Cela signifie que la proportion d'agents qui utilisent leur anticipation passée a diminué.

3.3.3 Observations générales

Regardons les valeurs obtenues pour le paramètre de mémoire longue d. Les valeurs du paramètre de mémoire longue obtenues avec notre modèle (voir tableau 3.7), confirment nos impressions visuelles. En effet, avant l'éclatement de la bulle, le paramètre de mémoire longue

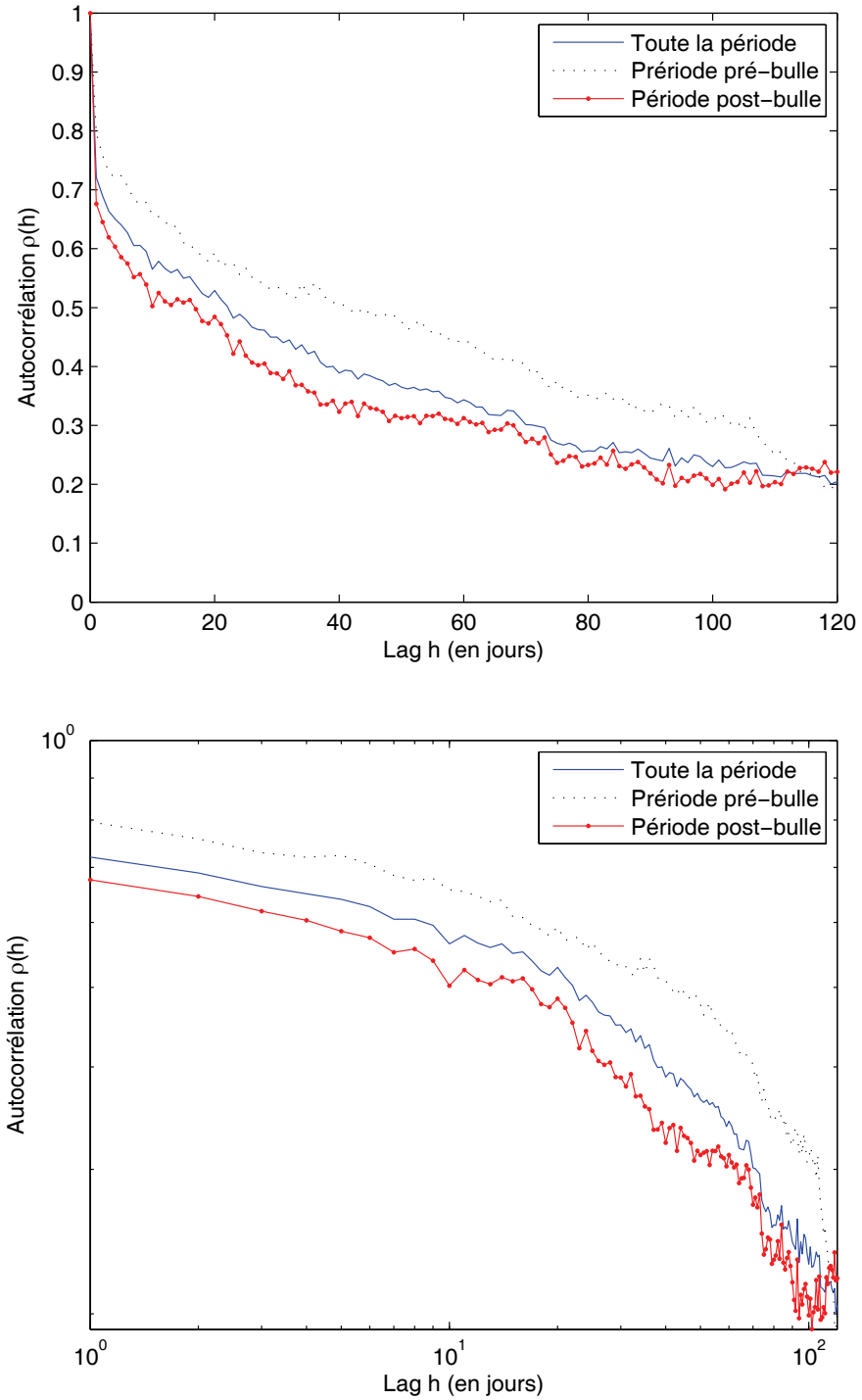

FIGURE 3.15 – Fonctions d'autocorrélation empiriques de la log-volatilité réalisée de Coca-Cola en mode linéaire (en haut) et en mode log-log (en bas), à horizon de 120 jours, calculées sur les différentes périodes considérées : globale, pré-bulle et post-bulle.

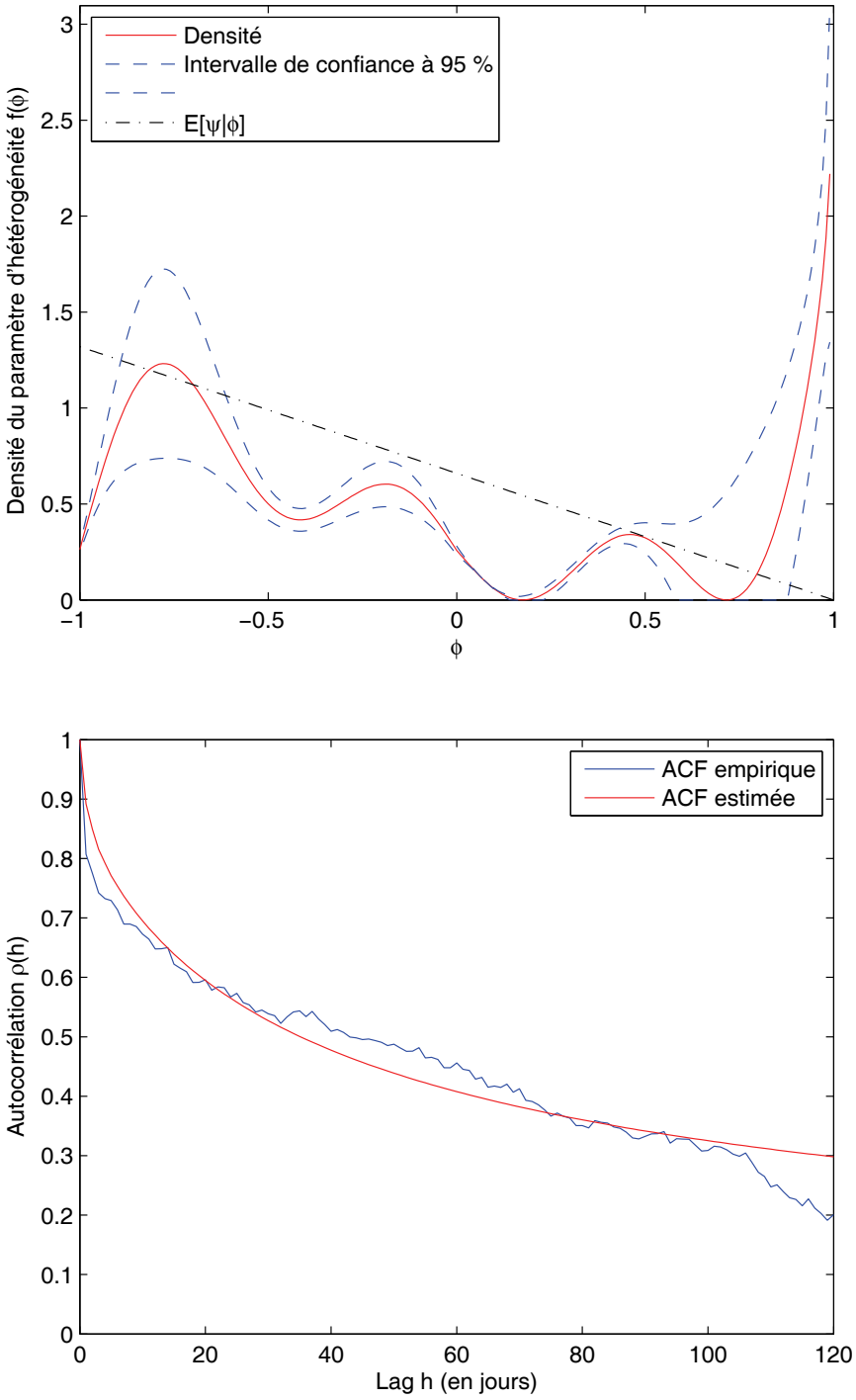

FIGURE 3.16 – Densité du coefficient d'hétérogénéité (en haut) et autocorrélation de la log-volatilité réalisée de Coca Cola (en bas), à horizon de 120 jours, durant la période pré-bulle.

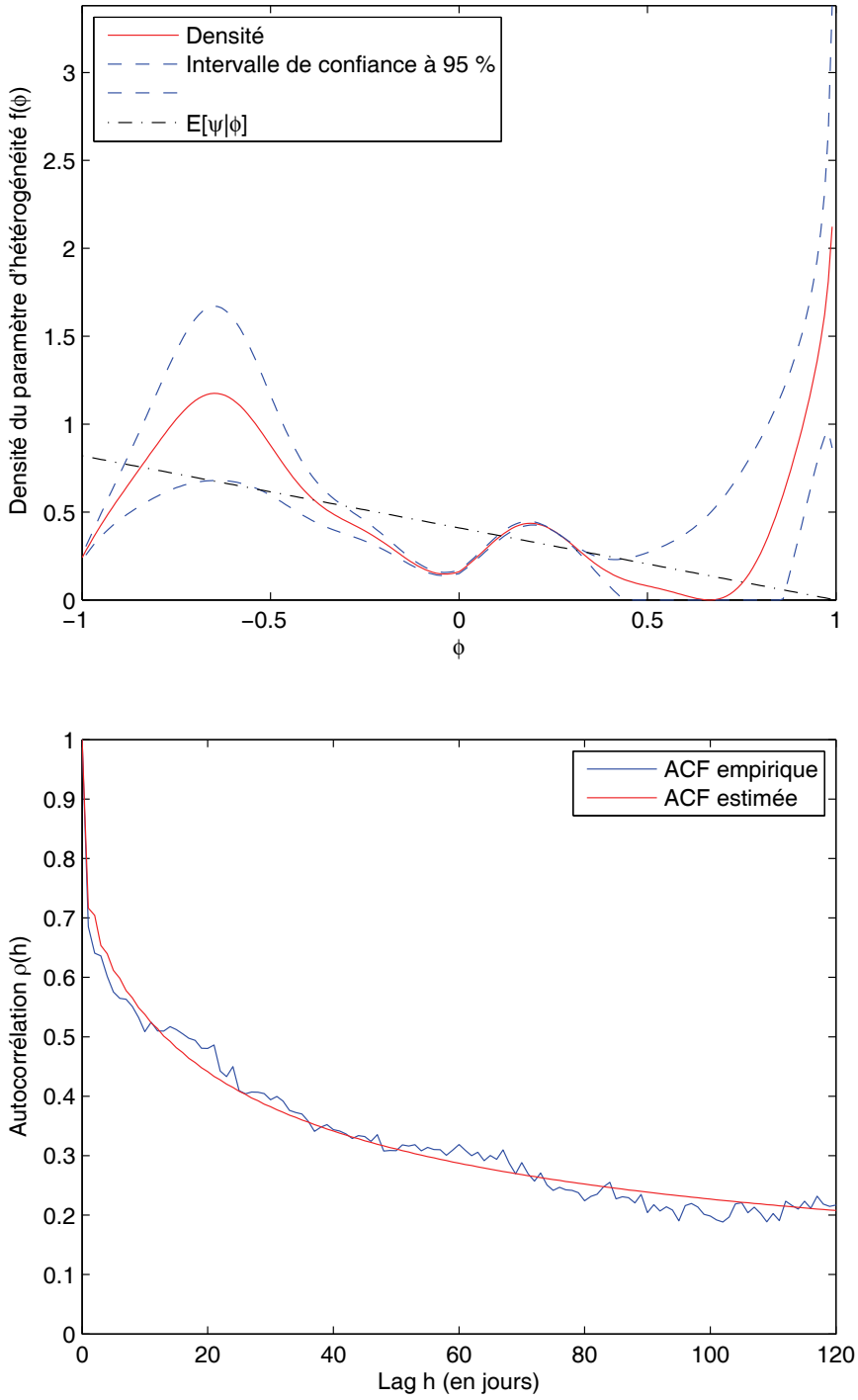

FIGURE 3.17 – Densité du coefficient d'hétérogénéité (en haut) et autocorrélation de la log-volatilité réalisée de Coca Cola (en bas), à horizon de 120 jours) durant la période post-bulle.

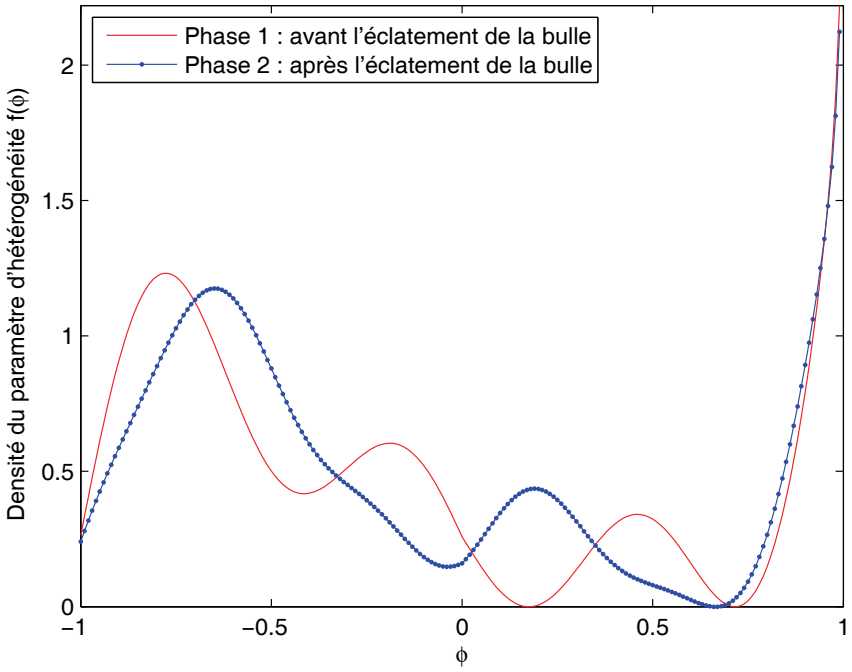

FIGURE 3.18 – Comparaison des densités du paramètre d'hétérogénéité φ de Coca Cola obtenues à partir des données des périodes pré et post-bulle.

est presqu'à chaque fois plus grand que celui obtenu en considérant la période globale : 4 fois sur 5 (dans le cas particulier, ces valeurs sont similaires). Au contraire, en observant avant l'éclatement de la bulle, d chute pour toutes les actions à l'exception de Coca-Cola.

Nous nous devons de mentionner que le très grand écart-type de d estimé pour Microsoft sur la seconde période nous permet de douter de la convergence de notre algorithme dans ce cas précis. Une explication possible est l'utilisation d'un nombre trop faible de données. En effet, pour la période post-bulle, nous ne disposons que de 900 jours pour 4 des 5 actions alors que 1600 données étaient utilisables pour les périodes pré-bulle. Dans le cas d'un nombre trop petit de données, le minimum global est plus difficile à atteindre. Par conséquent le résultat pourrait être faux.

Conclusion

Dans ce chapitre, nous avons commencé par définir la densité du paramètre d'hétérogénéité comme étant la somme pondérée d'une fonction de densité régulière quelconque qui admet un développement en série de Fourier et la densité d'une loi Beta$(2,1-d)$. Grâce à cette définition, le phénomène de mémoire longue peut être capturé par le modèle et une variété intéressante de comportements des acteurs du marché peut être interprétée. Nous

		Toute la période	Avant l'éclatement de la bulle	Après l'éclatement de la bulle	Jour du découpage
CSCO	d ($\bar{\sigma}$)	0.33 (0.06)	0.27 (0.04)	0.18 (0.04)	27/03/2000
	w	0.12	0.57	0.55	
	α	0.12	0.80	0.55	
MRK	d ($\bar{\sigma}$)	0.22 (0.05)	0.30 (0.04)	0.08 (0.04)	30/11/2000
	w	0.43	0.43	0.45	
	α	0.31	0.66	0.39	
KO	d ($\bar{\sigma}$)	0.16 (0.05)	0.19 (0.04)	0.20 (0.05)	15/07/1998
	w	0.46	0.41	0.47	
	α	0.62	0.66	0.41	
AIG	d ($\bar{\sigma}$)	0.10 (0.03)	0.26 (0.05)	-0.32 (0.04)	11/12/2000
	w	0.64	0.61	0.12	
	α	0.64	0.45	0.85	
MSFT	d ($\bar{\sigma}$)	0.21 (0.03)	0.34 (0.04)	*0.00 (2.10^5)	28/12/1999
	w	0.66	0.70	0.05	
	α	0.54	0.50	0.30	

TABLE 3.7 – Estimation du paramètre de mémoire longue d (et son écart-type), le paramètre α et le poids w en utilisant la méthode des autocorrélations pour quelques grandes capitalisations américaines, sur une période de 10 ans allant du 01/01/1994 au 31/12/2003. Remarque : l'étoile $*$ signifie que l'optimisation n'a pas fonctionné.

avons également discuté de l'estimateur des paramètres du modèle et établi ses propriétés asymptotiques. Puis, nous avons fourni et justifié de choix d'un algorithme que nous avons mis au point, d'une efficacité acceptable et d'une rapidité supérieure à l'algorithme génétique couramment utilisé pour traiter notre type de problème.

Ceci nous a permis de calibrer notre modèle sur les données de 24 actions américaines, de capitalisation variable, issues de différents secteurs d'activité. Nous avons mis en lumière typiquement deux allures de la densité du paramètre d'hétérogénéité qui ne semblent pas être reliées à la taille de l'entreprise ni au domaine d'activité. Ces deux densités ont un point commun : la quantité d'investisseurs accordant de l'importance à leur anticipation précédente au détriment des informations ou de la réalisation passée ($\varphi = 1$). Ce résultat n'est pas très surprenant car c'est ce type de comportement auto-référent et son degré d'hétérogénéité parmi les agents qui génèrent de la mémoire longue. La différence se situe pour des valeurs du paramètre d'hétérogénéité $\varphi \in [-1, 0.5]$. Dans un cas de figure, on trouve un grande quantité d'investisseurs qualifiés de rationnels, accordant de l'importance aux informations, privées ou publiques et à la réalisation passée de la log-volatilité réalisée, et une faible quantité d'investisseurs dont l'anticipation est basée sur l'opposé de leur estimation passée ainsi que de la réalisation passée sous réserve que α ne soit pas nul. Dans l'autre cas de figure, la proportion d'investisseurs rationnels diminue au profit de ces derniers et l'on voit apparaître

une autre catégorie d'investisseurs qui se basent à la fois sur l'opposé de leur anticipation passée, la réalisation passée et les informations ($\varphi \approx 0.5$).

Enfin, nous avons profité de la présence de la bulle internet (dont l'éclatement a eu lieu en mars-avril 2000) durant la période sur laquelle s'étend notre base de données pour étudier l'évolution du comportement des agents selon que le marché se trouvait dans une phase de croissance ou au contraire dans une phase de crise. Pour résumer, les résultats que nous obtenons sont ceux auxquels on pouvait légitimement s'attendre : pendant une période de croissance, qui voit le cours des actions augmenter de manière régulière, le nombre d'agents faisant confiance au marché et dont l'estimation du risque demeure la même jour après jour est plus grand que pendant une période de crise, qui voit le cours des actions chuter, durant laquelle les agents adoptent des comportements plus variés et moins auto-référents. En effet, bien que leur proportion ait diminué, certains continuent à fonder leurs anticipations sur leur précédente. D'autres ne croient plus du tout en leur anticipation passée et sont attentifs aux flux d'informations provenant de leurs réseaux (public et/ou privé). Une autre catégorie d'agents regroupe ceux qui admettent qu'ils ont fait une erreur la veille et prennent le contre-pied de leur anticipation passée. Enfin, certains font un mélange entre le contraire de leur anticipation passée, les informations récentes (du jour même) et la réalisation passée de la log-volatilité réalisée. Globalement, on observe une certaine tendance au recentrage de φ vers zéro (voir les figures 3.13 et 3.18). Cela signifie que de manière générale, après un krach, les agents ont une tendance à retourner collecter de l'information afin d'évaluer le niveau de risque des actions.

Annexes

3.A Coefficients du moment d'ordre k

$\mathrm{E}_1[\varphi^k]$ est donné par :

$$\mathrm{E}_1[\varphi^k] = \frac{1}{2}A_k + \sum_{n=1}^{\infty} a_n B_{n,k} + \sum_{n=1}^{\infty} b_n C_{n,k}, \tag{3.A.1}$$

où

$$A_k = \begin{cases} \dfrac{2}{k+1} & \text{si } k \text{ est pair,} \\ 0 & \text{si } k \text{ est impair,} \end{cases} \tag{3.A.2}$$

$$B_{n,k} = \begin{cases} (-1)^n \dfrac{2k}{(n\pi)^2} - \dfrac{k(k-1)}{(n\pi)^2} B_{n,k-2} & \text{si } k \text{ est pair,} \\ -\dfrac{k(k-1)}{(n\pi)^2} B_{n,k-2} & \text{si } k \text{ est impair,} \end{cases} \tag{3.A.3}$$

où $B_{n,0} = 0$ et $B_{n,1} = 0$.

$$C_{n,k} = \begin{cases} -\dfrac{k(k-1)}{(n\pi)^2} C_{n,k-2} & \text{si } k \text{ est pair,} \\ (-1)^{n-1} \dfrac{2}{n\pi} - \dfrac{k(k-1)}{(n\pi)^2} C_{n,k-2} & \text{si } k \text{ est impair,} \end{cases} \tag{3.A.4}$$

avec $C_{n,0} = 0$ et $C_{n,1} = (-1)^{n-1}\dfrac{2}{n\pi}$.

$\mathrm{E}_2[a^k]$ est simplement donné par :

$$\mathrm{E}_2[\varphi^k] = \frac{\Gamma(k+2)}{\Gamma(k+3-d)}\Gamma(3-d). \tag{3.A.5}$$

□

3.B Méthodes semi-paramétriques d'estimation du paramètre de mémoire longue d

Les méthodes semi-paramétriques sont des méthodes qui se concentrent exclusivement à l'estimation de d, par opposition aux méthodes paramétriques telles que l'estimation par maximum de vraisemblance qui estiment l'ensemble des paramètres en même temps ce qui entraîne d'une part un temps de calcul assez long et d'autre part des erreurs de spécification

de modèle. Il existe principalement deux méthodes dites semi-paramétriques. Une méthode couramment utilisée a été proposée par Geweke et Porter-Hudak (1983). Mais auparavant, Hurst (1951) avait proposé une méthode pour estimer la mémoire longue dans les séries temporelles qu'il avait appliquée à l'hydrologie. Les estimations obtenues par ces deux méthodes sont différentes, mais nous verrons qu'il existe une relation simple entre elles.

3.B.1 Méthode de Geweke et Porter-Hudak

Rappelons brièvement que la méthode de régression (ci-après GPH), décrite dans Brockwell et Davis (1990, pp. 162-168) par exemple, est basée sur le fait que la densité spectrale des séries temporelles à mémoire longue diverge au voisinage de 0

$$f(\lambda) \sim C \cdot \lambda^{-2d}. \tag{3.B.6}$$

En définissant les fréquences de Fourier $\lambda_j = 2\pi j/T \in]0, \pi[$ et les variables correspondantes $Z_j = \ln|1 - e^{-i\lambda_j}|^2$, le paramètre de mémoire longue d peut être vu comme le coefficient de régression linéaire du logarithme du périodogramme $I(\lambda_j)$ de l'échantillon sur ces variables Z_j

$$\log(I(\lambda_j)) = \alpha + dZ_j + \varepsilon_j, \tag{3.B.7}$$

avec le périodogramme défini par

$$I(\lambda) = \frac{1}{2\pi T}\left|\sum_{k=1}^{T} \exp(-i\lambda t)(X_t - \bar{X}_t)\right|^2, \tag{3.B.8}$$

où T est la taille du processus et $\bar{X}_t$ sa moyenne empirique.

Cela donne l'estimateur

$$\hat{d}_{GPH} = -\frac{\sum_{i=1}^{m}(Z_i - \bar{Z})(\omega_i - \bar{\omega})}{\sum_{i=1}^{T}(Z_i - \bar{Z})^2}. \tag{3.B.9}$$

Geweke et Porter-Hudak (1983) soutiennent qu'il existe une valeur m tel que $(\ln T)^2 \to 0$ quand $n \to \infty$ et que l'estimateur $\hat{d}_{GPH}$ converge asymptotiquement vers la loi normale :

$$\hat{d}_{GPH} \text{ est } \mathcal{AN}\left(d, \pi^2 \Big/ \left[6\sum_{i=1}^{T}(Z_i - \bar{Z})^2\right]\right), \tag{3.B.10}$$

sous réserve que la condition de stationnarité $d < 0.5$ soit vérifiée.

3.B.2 Méthode de Hurst

Nous allons à présent nous intéresser à l'estimation de l'intégrant fractionnaire d par la méthode de l'exposant de Hurst. Il existe une relation simple entre ces deux paramètres

caractérisant la mémoire longue. En notant

$$H = d + 0.5 \qquad d = [-0.5\ ;\ 0.5[\tag{3.B.11}$$

l'exposant de Hurst, il a été établi (voir Hosking (1984), Lo (1991), Peters (1994) entre autres) que la statistique R/S classique

$$\tilde{Q}_T = \frac{1}{S_T}\left[\max_{1\leq k\leq T}\sum_{j=1}^{k}(\omega_j - \bar{\omega}_T) - \min_{1\leq k\leq T}\sum_{j=1}^{k}(\omega_j - \bar{\omega}_T)\right], \tag{3.B.12}$$

où S_T est l'estimateur de la déviation standard habituel, et $\bar{\omega}_T$ la moyenne empirique des ω_j

$$S_T = \sqrt{\frac{1}{T}\sum_{j=1}^{T}(\omega_j - \bar{\omega}_T)^2}, \tag{3.B.13}$$

peut être reliée à H par l'équation de régression standard

$$\log(\tilde{Q}_T) = \log(c) + H\log(T). \tag{3.B.14}$$

Remarquons que le premier terme entre parenthèses dans (3.B.12) est le maximum sur k des sommes partielles des k premiers écarts de ω_j par rapport à la moyenne empirique, tandis que le second terme est le minimum sur k des mêmes sommes partielles.

Deuxième partie

Gestion de portefeuille

Introduction

Dans les années 1950, les premiers modèles de choix de portefeuille, telle l'approche moyenne-variance de Markowitz (1952), ont été élaborés dans un cadre discret mono-périodique. Puis, pour rendre compte du fait que l'investisseur peut rééquilibrer de temps à autre son portefeuille, les modèles ont été développés dans un cadre discret multi-périodique (Hakansson 1971). La même année est apparu un modèle en temps continu, proposé par Merton (1971) et utilisant une autre approche, celle de l'utilité espérée. Ce modèle est plus conforme à la réalité puisqu'un investisseur peut intervenir à tout moment sur le marché.

Par la suite, les deux approches ont évolué parallèlement : la première a été développée dans un cadre en temps continu au début des années 2000 par Zhou et Li (2000) et la deuxième a pris en compte les frais de transactions (Magill et Constantinides 1976). Or, tous ces modèles considèrent une volatilité constante, ce qui, compte tenu du caractère variable de celle-ci, conduit à une allocation non-optimale de la richesse du portefeuille et peut même se révéler particulièrement inefficace lorsque la volatilité connaît des pics comme pendant une crise financière. Par conséquent, la prise en compte, non seulement de la dynamique de la volatilité, mais aussi de ses autres caractéristiques comme le retour à la moyenne et la mémoire longue, nous semble un enjeu important en gestion de portefeuille. A ce titre, au début des années 2000 sont apparus les premiers modèles prenant en compte explicitement la volatilité stochastique (Benth, Karlsen, et Reikvam (2003), Chacko et Viceira (2005) ou encore Liu (2007)).

Notre but dans cette partie est de proposer un modèle de choix de portefeuille en temps continu avec une volatilité stochastique à mémoire longue et d'étudier l'impact de la mémoire longue sur l'allocation optimale lorsque l'investisseur cherche à maximiser l'utilité espérée de sa richesse terminale. À notre connaissance, jusqu'à présent, cette caractéristique très importante de la volatilité n'a pas été explicitement modélisée. À partir du cas très particulier du modèle de Merton (1971) dans lequel l'équation de diffusion du cours de l'actif risqué suit un mouvement brownien géométrique, nous remplaçons la volatilité constante de l'actif risqué par une volatilité stochastique à mémoire longue. Notre attention se porte sur le processus d'Ornstein-Uhlenbeck d'une part pour sa capacité à évoluer en processus fractionnaire (Fouque, Papanicolaou, et Sircar 2000) et d'autre part pour sa caractéristique de retour à la moyenne. Puisque contrairement à la volatilité, qui par définition est toujours positive, la log-volatilité peut prendre des valeurs négatives, nous supposons que sa dynamique suit un processus d'Ornstein-Uhlenbeck fractionnaire.

Le mouvement brownien fractionnaire, fondé comme son nom l'indique sur le mouvement brownien, est l'expression la plus simple d'un processus stochastique à mémoire longue. Il a été défini par Kolmogorov (1940), puis repris par Hurst (1951, 1957) en hydrologie, et plus tard en finance par Mandelbrot et Van Ness (1968) qui désignent par mouvement brownien fractionnaire d'exposant d ($-1/2 < d < 1/2$), une famille de fonctions aléatoires gaussiennes, résultat d'une moyenne mobile des incréments $dB(t)$ du mouvement brownien sur lequel il s'appuie et dont les incréments passés sont pondérés par le noyau $(t-s)^d$. L'utilisation d'un tel processus nécessite des précautions. En effet, un mouvement brownien fractionnaire n'est pas une semi-martingale. Par conséquent, la théorie classique du calcul stochastique ne peut pas s'appliquer et il a donc fallu l'adapter. A ce sujet, Lin (1995) et Dai et Heyde (1996) ont introduit les intégrales stochastiques et étendu le lemme d'Itō au mouvement brownien fractionnaire.

A notre connaissance, il n'existe pas de modèles de choix de portefeuille avec volatilité stochastique prenant en compte la mémoire longue de la volatilité. Barndorff-Nielsen et Shephard (2001) ont suggéré de modéliser la volatilité par une somme pondérée de processus d'Ornstein-Uhlenbeck non gaussiens. Ainsi, la densité de volatilité possède des queues épaisses et l'autocorrélation persiste à long terme. En outre, le niveau de volatilité peut connaître un bond brutal puis décroître de manière exponentielle. Cette technique a été utilisée par Benth, Karlsen, et Reikvam (2003) dans leur modèle de choix de portefeuille. Hélas, aucun paramètre permet de contrôler la caractéristique fondamentale des séries de volatilité qu'est la mémoire longue. Cette constatation fait partie des éléments qui nous ont encouragés à développer notre modèle.

Afin de comprendre le rôle du caractère stochastique de la volatilité d'une part, et de la mémoire longue d'autre part, le modèle a été développé en deux temps. Dans un premier temps, nous remplaçons la volatilité constante du modèle de Merton par une volatilité stochastique : la dynamique de la log-volatilité est traduite par un processus d'Ornstein-Uhlenbeck. Dans un deuxième temps, nous faisons évoluer le processus d'Ornstein-Uhlenbeck pour qu'il puisse prendre en compte la mémoire longue. Bien que le processus d'Ornstein-Uhlenbeck fractionnaire existe, nous avons opté pour un processus voisin en raison de la corrélation entre la volatilité et les rendements de l'actif risqué que nous voulions contrôler à tout moment. Nous avons donc remplacé le mouvement brownien du processus d'Ornstein-Uhlenbeck par une somme pondérée d'un mouvement brownien et d'un mouvement brownien fractionnaire indépendant. Ainsi, la mémoire longue est contrôlée par le paramètre de mémoire longue de ce dernier.

Nous dérivons les résultats analytiques du modèle à volatilité constante de Merton et de notre modèle à volatilité stochastique pour une généralisation du critère de gestion, réalisée au moyen d'un paramètre qui contrôle le poids accordé à l'utilité espérée de la consommation et à l'héritage. Pour certaines valeurs, l'investisseur, dont les préférences sont caractérisées par une fonction HARA, peut décider de laisser un capital à l'échéance, ou pas, ou de consommer durant la période de gestion, ou pas. Nous traitons ce problème par programmation dynamique et équation d'Hamilton-Jacobi-Bellman. Si pour le modèle de Merton, l'équation aux dérivées partielles obtenue peut se résoudre analytiquement, ce n'est pas le cas pour notre modèle à volatilité stochastique. Nous définissons donc un algorithme de résolution fondé sur l'approximation par les différences finies. Puis, nous étudions l'impact du caractère sto-

chastique de la volatilité en considérant le cas particulier d'un investisseur ayant pour but de maximiser l'utilité espérée de sa consommation avec héritage à l'échéance T. Puisque le modèle comporte un nombre conséquent de paramètres, pour faire l'économie de deux paramètres et ainsi simplifier cette étude, nous considérons que les préférences de l'investisseur sont dictées par une fonction d'utilité CRRA. Nous mettons en évidence les paramètres agissant sur l'ampleur de la variabilité de la volatilité et leurs rôles respectifs sur la consommation optimale et sur le poids optimal de l'actif risqué. En particulier, nous observons qu'au-delà un certain niveau d'incertitude, le plan de consommation et d'investissement change.

Ensuite, nous dérivons les résultats analytiques du modèle à volatilité stochastique à mémoire longue pour un investisseur, dont les préférences sont traduites par une fonction CRRA, ayant pour objectif de maximiser l'utilité espérée de sa richesse finale. Nous décidons de traiter un pur cas de choix de portefeuille afin de simplifier le problème. En effet, la modélisation de la mémoire longue via le mouvement brownien fractionnaire a transposé le système d'un cadre markovien à un cadre non-markovien. Par conséquent, la théorie classique par programmation dynamique et résolution de l'équation d'Hamilton-Jaccobi-Bellman est inadaptée. Nous procédons donc autrement et développons un algorithme récursif de résolution faisant appel à la technique de Monte Carlo. Puis, après s'être assuré de la convergence de nos deux modèles, le premiers à volatilité stochastique et le deuxième à volatilité stochastique à mémoire longue, dans le cas particulier où il n'y a pas de mémoire longue, nous mesurons l'impact de la mémoire longue de la volatilité sur l'allocation optimale. Nous montrons les conséquences en terme de performance de la mauvaise spécification du modèle ainsi que l'importance de prendre en compte la mémoire longue de la volatilité dans la modélisation.

Cette deuxième partie est organisée de la manière suivante. Dans le chapitre 4, nous effectuons une revue de la littérature sur l'évolution des modèles de choix de portefeuille en temps continu. Nous commençons par rappeler les différentes dimensions à prendre en compte dans la modélisation de choix de portefeuille. Puis, nous présentons le cas particulier du modèle de Merton en temps continu avec volatilité constante. Enfin, nous détaillons les différents modèles récents avec volatilité stochastique. Le chapitre 5 est consacré à la présentation de notre modèle de choix de portefeuille en temps continu avec volatilité stochastique, qui nous servira de support pour développer notre modèle à mémoire longue, et à l'étude de l'impact du caractère stochastique de la volatilité sur l'allocation optimale. Enfin, dans le chapitre 6, nous définissons le modèle avec volatilité stochastique à mémoire longue, et répondons aux questions de calibrage. Nous clôturons ce chapitre par l'étude de l'impact de la mémoire longue sur la politique optimale de l'investisseur.

Chapitre 4

Les modèles de choix de portefeuille : revue de littérature

En gestion de portefeuille, un investisseur dispose d'un capital (ou richesse initiale) qu'il peut répartir en général entre plusieurs actifs financiers risqués et un non risqué. Dès lors, il se demande quelle part de richesse il doit investir dans chacun de ces actifs. Cette problématique communément appelée choix de portefeuille par les économistes, simple en apparence, a donné naissance à une abondante littérature. En effet, les modèles se différencient les uns des autres par de nombreux critères :

- *leur approche* : il en existe trois. L'approche rendement-risque où, comme son nom l'indique, l'investisseur cherche le meilleur compromis possible entre le risque et le rendement (Markowitz 1952). L'approche par la théorie de l'utilité espérée consiste à traduire par une fonction d'utilité les préférences en terme de risque de l'investisseur et de maximiser le critère de gestion choisi (Merton 1971). Une troisième approche, reliée à la fonction d'utilité consiste à chercher la croissance optimale du portefeuille (Cover 1991).
- *leur cadre* : les modèles les plus simples sont développés dans un cadre mono-périodique (Markowitz 1952). L'investisseur forme son portefeuille au début de la période de gestion et ne peut par la suite plus intervenir. Puis, au début des années 1970, le cadre discret multi-périodique est apparu pour lui permettre de rééquilibrer par moment son portefeuille (Hakansson 1971). Au même moment, un autre modèle a ouvert la voie à la prise en compte d'un temps continu, plus fidèle à la réalité des marchés (Merton 1971).
- *leur prise en compte des imperfections de marché* : cette prise en compte relève plusieurs formes, à savoir les frais de transaction générés par chaque opération sur le marché (Magill et Constantinides 1976), les taxes (Leland 1999) et les liquidités (Kyle 1985).
- *leur prise en compte des facteurs de risque* : la modélisation la plus simple considère une volatilité constante. Or, la volatilité varie au fil du temps. C'est pourquoi une volatilité stochastique a été introduite dans les modèles plus évolués (Chacko et Viceira 2005). Récemment, un autre fait stylisé de la volatilité commence à être pris en compte : la mémoire longue (Benth, Karlsen, et Reikvam 2003). Dans une autre approche, Breeden (1979) relie la dynamique du prix des actifs risqués à n facteurs de risque.
- *les objectifs de l'investisseur* : dans l'approche par la théorie de l'utilité espérée, son

objectif est de maximiser un critère de gestion. Les critères de gestion diffèrent par l'horizon (fini ou infini) et par l'importance accordée à la satisfaction de la consommation par rapport à l'héritage.
- *les méthodes de résolution du problème.* Dans la modélisation intertemporelle, nous en rencontrons principalement deux : l'approche par programmation dynamique et équation de Hamilton-jacobi-Bellman (Merton 1971) et celle par martingale (Pliska 1986).

Compte tenu de la richesse de la littérature consacrée aux modèles de choix de portefeuille, il n'est pas possible de faire une revue de littérature exhaustive. Toutefois nous entamons ce chapitre en développant un peu plus les différentes dimensions évoquées ci-dessus (section 1). Puis, puisque dans cette partie, nous proposons un modèle de choix de portefeuille en temps continu avec volatilité stochastique un autre avec volatilité stochastique à mémoire longue, nous rappelons le premier modèle en temps continu : le modèle simplifié de Merton à volatilité constante dans le cas particulier où le prix de l'actif risqué suit un mouvement brownien géométrique (section 2). Enfin, dans la section 3 nous détaillons des modèles récents considérant une volatilité stochastique.

Quittard-Pinon (2003) fournit plus de précisions concernant le modèle de Markovitz, la théorie de l'utilité espérée et le modèle de Merton. Dans un souci de meilleure lisibilité, les notations des modèles que nous présentons ont été adaptées le plus possible au reste de la thèse.

4.1 Les critères prix en compte dans la modélisation

4.1.1 L'approche

L'approche rendement-risque consiste à trouver le meilleur compromis pour l'investisseur entre le risque du portefeuille et son rendement. Cette thématique a été étudiée par Roy (1952) au travers de son principe *safety first* stipulant que les investisseurs s'efforcent de minimiser la probabilité que le rendement de leur portefeuille tombe en dessous d'un niveau désigné subjectivement. La méthode moyenne-variance de Markowitz (1952) est la première dans ce domaine. Dans la modélisation de Markowitz, le rendement du portefeuille est caractérisé par la moyenne tandis que son risque l'est par la variance de des rendements. L'investisseur dispose de deux méthodes pour trouver son compromis optimal : d'une part il peut chercher à minimiser la variance de son portefeuille pour un niveau de rendement espéré fixé, et d'autre part pour un risque fixé il peut avoir pour but de maximiser le rendement espéré. Dans cette méthode, le risque est traduit par la variance des rendements. Mais ce n'est pas la seule possibilité. Par exemple Basak et Shapiro (2001) et Alexander et Baptista (2002) utilisent la *Value-at-Risk*. Quant à l'approche par la théorie de l'utilité espérée, elle considère que l'investisseur, comme tout le monde, a des besoins (se nourrir, s'habiller, se loger...) et des désirs (acquérir un bien immobilier, partir en vacances, se divertir...) qui entraînent naturellement des coûts et diffèrent d'un individu à l'autre. La façon dont il va utiliser son argent lui est propre et la modélisation doit en tenir compte. Une fonction d'utilité est employée pour caractériser sa consommation. Il en existe plusieurs types : les fonctions HARA caractéri-

sant l'aversion absolue au risque hyperbolique, CRRA (qui sont une sous-classe des fonctions HARA) traduisant l'aversion relative au risque constante ou encore la fonction d'utilité récursive introduite par Epstein et Zin (1989) dans un cadre discret puis adaptée au cadre continu par Duffie et Epstein (1992) pour ne citer que celles-ci. Leurs paramètres modélisent des comportements différents. Par exemple, les investisseurs pourront être considérés comme étant plus ou moins risquophobes. A noter que la méthode moyenne-variance conduit à des solutions explicites alors que l'approche par la théorie de l'utilité espérée, plus conforme à la théorie néoclassique, pas toujours (Quittard-Pinon 2003). La dernière approche évoquée précédemment est reliée à la fonction d'utilité : l'objectif consiste à chercher la croissance optimale du portefeuille (Cover 1991).

4.1.2 Le cadre

Le modèle de Markowitz (1952), bien qu'ayant marqué la gestion de portefeuille, souffre d'un défaut majeur. En effet, la plus grande critique à son égard concerne son cadre monopériodique, peu réaliste en pratique : les investisseurs peuvent intervenir à tout moment sur le marché et non pas seulement au moment de la formation de leur portefeuille. Cet inconvénient a été surmonté en transposant le modèle moyenne-variance dans un cadre multi-périodique. Ainsi, l'investisseur ne peut certes pas intervenir à tout moment, mais il a tout de même la possibilité de rééquilibrer son portefeuille épisodiquement (voir les modèles développés par Hakansson (1971), Samuelson (1986), Grauer et Hakansson (1993), Pliska (1997) ou encore Li et Ng (2000)). Dès lors, le fait de considérer une volatilité de l'actif risqué constante devient problématique puisqu'il est prouvé qu'elle varie au cours du temps. Par ailleurs, pour être totalement conforme à la réalité, la modélisation doit prendre en compte le fait que les acteurs du marché peuvent passer des ordres à tout moment durant la période d'ouverture. La même année que Hakansson (1971), Merton (1971) proposait un modèle qui allait révolutionner la gestion de portefeuille : le temps n'était plus discret mais bel et bien continu, autrement dit plus conforme à la réalité. A partir de cette date, les deux approches, celle par moyenne-variance et celle par la théorie de l'utilité espérée, ont été développées parallèlement. Au début des années 2000 sont apparus les premiers modèles moyenne-variance de choix de portefeuille en temps continu grâce aux travaux, entre autres, de Zhou et Li (2000), Portait, Bajeux-Besnainou, et Jordan (2001), Lim et Zhou (2002), Zhou et Yin (2003).

4.1.3 La prise en compte des imperfections de marché

La prise en compte des imperfections de marché diffère selon les modèles. Par exemple, dans le modèle de Merton, le gestionnaire intervient continuellement sur le marché pour maintenir la proportion d'actif risqué au niveau souhaité. Or, dans la réalité une telle pratique est impossible en raison des frais de transaction qui accompagnent chaque achat ou vente de titre : elle conduirait rapidement à la ruine du portefeuille. Par conséquent, pour remédier aux faiblesses du modèle de Merton, et ainsi le rendre plus réaliste, des chercheurs ont par la suite introduit les frais de transaction pour espacer dans le temps les interventions du gestionnaire. Ces modèles, développés depuis les années 1970, sont relativement peu nombreux en partie à cause des difficultés mathématiques générées par la prise en compte des frais de transactions.

Dans la littérature, nous en rencontrons deux types : les frais proportionnels au montant de la transaction (voir Magill et Constantinides (1976), Taskar, Klass, et Assaf (1988), Davis et Norman (1990), Akian, Sulem, et Taksar (2001) et Muthuraman et Kumar (2006)) et ceux proportionnels à la valeur du portefeuille (voir Morton et Pliska (1995) et Atkinson, Pliska, et Wilmott (1997)). D'une manière générale, la prise en compte des frais de transaction génère la formation d'une zone dans le plan défini par l'actif non risqué et l'actif risqué à l'intérieur de laquelle aucune transaction n'a lieu. En d'autres termes, dans la zone de non-transaction, le bénéfice attendu par l'investisseur du rééquilibrage de son portefeuille après que sa composition se soit écartée de l'allocation optimale est inférieur aux coûts occasionnés. Ainsi, l'investisseur n'interviendra que de temps en temps sur le marché. Jouini, Koehl, et Touzi (1999) considèrent une taxe sur les bénéfices. Quant à Leland (1999), il considère à la fois des frais de transaction proportionnels au montant de la transaction et une composante de taxe sur les plus-values du capital. Kyle (1985) et Gabaix, Gopikrishnan, Plerou, et Stanley (2006) suggèrent une façon alternative de prendre en compte les imperfections de marché en considérant la liquidité. Le principe est que les mouvements du marché sont la conséquence de l'activité de très grands investisseurs institutionnels dont les ordres passés provoquent un pic dans les rendements et le volume, même en l'absence d'arrivée d'information concernant la valeur fondamentale de l'entreprise.

4.1.4 La prise en compte des facteurs de risque

Les modèles diffèrent également par la prise en compte des facteurs de risque. Précédemment, nous avons brièvement évoqué un inconvénient du modèle de Markowitz (1952) : à l'image du modèle de Merton (1971) et de tous les modèles développés jusqu'au début des années 2000, il fait l'hypothèse d'une volatilité constante. Or, c'est un fait bien établi : la volatilité des rendements d'actions fluctue au cours du temps. Ainsi, dans tous les modèles à volatilité constante, le risque du portefeuille n'est pas correctement estimé. Par conséquent, les positions prises par l'investisseur peuvent s'avérer pénalisantes surtout si la volatilité des actions change brutalement et de manière importante comme lorsque que le marché doit faire face à une crise [39]. La prise en compte de la dynamique de la volatilité dans la gestion de portefeuille représente donc un enjeu majeur. Celle-ci aura pour effet d'inciter le gérant à agir plus fréquemment sur le marché. En somme, les deux facteurs à prendre en compte en gestion de portefeuille, les frais de transaction et la dynamique de la volatilité vont avoir une influence contraire sur le gestionnaire : le premier va espacer dans le temps ses interventions, tandis que le second va au contraire l'inciter à rééquilibrer son portefeuille. Hélas, compte tenu de la difficulté mathématique de résolution de chacun de ces facteurs séparément, il n'existe pas à notre connaissance de modèle les traitant de front, à l'exception de celui de Atkinson et Mokkhavaisa (2003). Les années 2000 ont vu apparaître des modèles à volatilité stochastique de Chacko et Viceira (2005), Liu (2007) que nous détaillerons dans une prochaine section. Benth, Karlsen, et Reikvam (2003) sont allés un peu plus loin en utilisant la modélisation de la volatilité de Barndorff-Nielsen et Shephard (2001) capable de traduire d'autres caractéristiques telles que la persistance à long terme. Mais il existe d'autres façons de prendre en

39. Par exemple, Schwert (1989) a constaté que pendant la Grande Dépression de 1929-1939 aux États-Unis, la volatilité était en moyenne deux à trois fois plus élevée qu'en temps normal, mais ce cas reste exceptionnel : pour les autres crises, de moindre envergure certes, la volatilité retombe à la normale au bout de quelques semaines voire de quelques jours.

compte les facteurs de risque : par exemple, Breeden (1979) relie la dynamique des prix à n facteurs de risque, Brennan et Xia (2002) considèrent un taux d'inflation stochastique et et Poncet et Lioui (2001) un taux d'intérêt stochastique.

4.1.5 L'objectif de l'investisseur

Dans l'approche par théorie de l'utilité espérée, nous trouvons plusieurs critères pour traduire l'objectif de l'investisseur. Le plus général d'entre eux consiste à maximiser l'utilité espérée de la consommation avec héritage sur un horizon de gestion fini (voir, entre autres, Liu (2007)). Il s'agit par exemple d'un particulier qui, tout en consommant, désire se constituer une retraite complémentaire. Des variantes existent : par exemple l'investisseur peut décider de ne pas laisser de capital à la fin de la période de gestion (voir Merton (1971), Magill et Constantinides (1976)). Dans ce cas, la consommation augmente fortement à l'horizon de gestion pour ne laisser aucun capital, ce qui n'est pas très réaliste. Ou bien encore, dans un cas plus théorique, l'investisseur peut avoir pour but de maximiser l'utilité espérée de sa richesse finale, sans consommer (Benth, Karlsen, et Reikvam 2003). Dans d'autres critères, le temps n'est plus fini, mais infini. Alors, l'investisseur peut avoir pour objectif de maximiser l'utilité espérée de sa consommation sans héritage sur un long horizon de gestion (voir Davis et Norman (1990), Chacko et Viceira (2005), Muthuraman et Kumar (2006)). Enfin, un autre objectif est de maximiser le taux de croissance asymptotique de son portefeuille (Taskar, Klass, et Assaf 1988, Akian, Sulem, et Taksar 2001), critère dont Cover (1991) a justifié l'intérêt.

4.1.6 La méthode de résolution

Enfin, en ce qui concerne les méthodes de résolution, il en existe principalement deux : l'approche classique par programmation dynamique et l'utilisation de l'équation d'Hamilton-Jacobi-Bellman développée par Merton (1971), et reprise dans de très nombreux articles, tels que Benth, Karlsen, et Reikvam (2003), Chacko et Viceira (2005), Liu (2007) et l'approche plus récente par martingale introduite par Pliska (1986), Karatzas, Lehoczky, et Shreve (1987), Cox et Huang (1989, 1991) et He et Pearson (1991). Dans cette deuxième approche, les auteurs ont relié l'utilité marginale de la richesse terminale du portefeuille optimal à la densité de la mesure de martingale en utilisant des techniques de dualité convexe. Puisque par la suite nous n'utiliserons pas cette méthode, nous n'entrerons pas dans les détails. Quant à la méthode classique, nous la détaillerons plus amplement lorsque nous y ferons appel dans le chapitre suivant. A noter toutefois que des variantes existent. Par exemple, Atkinson et Mokkhavaisa (2003) dans leur modèle à volatilité stochastique et frais de transaction, une fois obtenue l'équation aux dérivées partielles par l'approche classique, utilisent la technique par perturbation pour mesurer les changements provoqués par les frais de transaction, supposés suffisamment petits, sur la fonction d'utilité indirecte du problème avec volatilité stochastique sans frais de transaction.

4.2 Un cas particulier du modèle de Merton à volatilité constante

Merton (1971) propose un modèle général de choix de portefeuille en temps continu où l'investisseur dispose d'un actif non risqué et de n actifs risqués dont la volatilité des rendements est constante. Dans cette section, nous nous intéressons à un cas très particulier du modèle de Merton que nous utiliserons comme modèle étalon par la suite (voir section 5.1 dans le chapitre 5) : celui à un actif risqué dont le cours est caractérisé par un mouvement brownien géométrique. Puisque dans le prochain chapitre ce modèle sera détaillé, nous nous contenterons ici de rappeler uniquement ses principaux éléments.

Dans son modèle, Merton, émet les hypothèses (i) il n'y a pas de frais de transaction, (ii) l'agent est un *price-taker*, (iii) les ventes à découvert sont autorisées, (iv) l'achat ou la vente d'actions peut se faire à tout moment. Il considère la situation d'un investisseur cherchant à déterminer la répartition optimale entre sa consommation et son investissement sur un intervalle de temps $[0, T]$. Il dispose de deux instruments financiers[40] : un placement dans un actif sans risque, au taux instantané r, dont le cours est noté $S_0(t)$ et un placement dans un actif risqué dont le cours $S_1(t)$ est décrit par un mouvement brownien géométrique. Les processus d'évolution des prix de ces deux actifs sont

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (4.1) \\ dS_1(t) = \mu S_1(t)dt + S_1(t)\sigma dB(t), & (4.2) \end{cases}$$

où μ, la dérive, représente le rendement moyen de l'actif risqué et σ, le coefficient de diffusion, sa volatilité. De plus, $B(t)$ est un processus de Wiener standard (Karatzas et Shreve 1991).

Merton suppose que la consommation instantanée de l'investisseur, notée $c(t)dt$, est financée par sa stratégie de placement, supposée auto-financée. La dynamique de sa richesse, notée $W(t)$, est alors donnée par

$$dW(t) = \left[(\mu - r)w(t)W(t) + rW(t) - c(t)\right] dt + w(t)W(t)\sigma dB(t). \tag{4.3}$$

L'investisseur cherche à maximiser l'utilité espérée de sa consommation à un horizon de gestion T, ce qui se traduit par

$$\max_{c(t),w(t)} \mathrm{E}\left[\int_0^T \mathrm{e}^{-\delta t} u(c(t))dt\right], \tag{4.4}$$

sous les contraintes de gestion $c(t) \geq 0$, $W(0) = W_0 > 0$ et où $u(c)$ est une fonction d'utilité supposée croissante et concave et le paramètre δ fournit une mesure de l'impatience à consommer de l'investisseur (Ingersoll 1987). Merton (1971) utilise l'approche classique par programmation dynamique et le théorème d'Hamilton-Jacobi-Bellman pour résoudre ce problème.

Dans le cas d'une fonction d'utilité de type HARA :

$$u(c) = \frac{1-\gamma}{\gamma}\left(\frac{\beta c}{1-\gamma} + \eta\right)^{\gamma}, \qquad \gamma < 1 \text{ et } \gamma \neq 0, \tag{4.5}$$

40. C'est le cas le plus simple. Merton (1971) développe aussi le cas général à n actifs risqués.

la consommation optimale et le poids optimal de l'actif risqué sont donnés par

$$c^*(t) = \frac{\delta - \gamma\nu}{1-\gamma} \frac{W(t) + \frac{(1-\gamma)\eta}{\beta}\frac{1-\mathrm{e}^{-r(T-t)}}{r}}{1-\mathrm{e}^{-\frac{\delta-\gamma\nu}{1-\gamma}(T-t)}} - \frac{(1-\gamma)\eta}{\beta}, \tag{4.6}$$

$$w^*(t) = \frac{\mu - r}{(1-\gamma)\sigma^2} + \frac{\mu - r}{(1-\gamma)\sigma^2}\frac{(1-\gamma)\eta}{\beta}\frac{1-\mathrm{e}^{-r(T-t)}}{r}\frac{1}{W(t)}, \tag{4.7}$$

où $\nu = r + \frac{(\mu-r)^2}{2(1-\gamma)\sigma^2}$. L'expression du poids optimal de l'actif risqué est composé de deux termes : le premier est la demande du portefeuille efficient, tandis que le second caractérise la demande de couverture inter-temporelle de l'investisseur. Lorsque le paramètre η vaut zéro, cette couverture est toujours nulle.

Pour une fonction d'utilité de type CRRA ($\eta = 0$ et $\beta = (1-\gamma)^{-\frac{1-\gamma}{\gamma}}$) les expressions de c^* et de w^* se simplifient

$$c^*(t) = \frac{\delta - \gamma\nu}{1-\gamma} \frac{W(t)}{1-\mathrm{e}^{-\frac{\delta-\gamma\nu}{1-\gamma}(T-t)}}, \tag{4.8}$$

$$w^* = \frac{\mu - r}{(1-\gamma)\sigma^2}. \tag{4.9}$$

En conclusion, que ce soit dans le cas d'une utilité CRRA où l'expression du poids optimal de l'actif risqué est égal à une constante, ou dans le cas d'une utilité HARA, l'investisseur devra perpétuellement rééquilibrer son portefeuille pour garder cette proportion constante. Si cette attitude est réalisable en l'absence de frais de transaction, appliquée au marché où tout échange de titre génère des coûts, elle va se révéler désastreuse : le portefeuille sera rapidement pénalisé par les coûts trop nombreux et ce comportement aboutira inéluctablement à sa ruine.

Un autre inconvénient majeur de ce modèle est qu'il considère une volatilité constante de l'actif risqué. Naturellement, cette hypothèse simplificatrice n'est pas conforme à la réalité puisque la volatilité évolue au cours du temps (pour plus de détails, se reporter à la revue de littérature sur la volatilité, section 1.2.1). En particulier, en période de crise financière, la volatilité atteint des niveaux beaucoup plus élevés. Cela entraîne donc une mauvaise estimation du risque du portefeuille et par conséquent une allocation non-optimale du portefeuille.

4.3 Prise en compte de la volatilité stochastique dans les modèles en temps continu à partir années 2000

Il faut attendre le début des années 2000, pour que des modèles d'allocation d'actifs intègrent une volatilité stochastique. La volatilité des rendements d'action n'est pas complètement observable en direct sur le marché. Toutefois, l'usage des données haute-fréquences permet, presque en temps direct, d'avoir accès à la volatilité. En outre, les agents ont une perception quotidienne du niveau de la volatilité. Il est alors naturel qu'ils cherchent à introduire cette part variable d'information dans leurs méthodes de gestion de portefeuille.

4.3.1 Le modèle de Chacko et Viceira

Chacko et Viceira (2005) font l'hypothèse d'un marché incomplet, dans lequel l'investisseur, qui a le choix entre deux actifs, l'un non risqué et l'autre risqué ayant un rendement moyen constant, gère son portefeuille en temps continu sur un long horizon de gestion. Leur modèle est défini ainsi :

$$\begin{cases} dS_0(t) = rS_0(t)dt & (4.10) \\ dS_1(t) = \mu S_1(t)dt + \sqrt{v_t}S_1(t)dB(t), & (4.11) \\ dv(t) = \kappa(\alpha - v(t))v(t)dt - \sigma v(t)^{3/2}dZ(t), & (4.12) \end{cases}$$

où $S_0(t)$ est le cours de l'actif sans risque de rendement r, $S_1(t)$ est celui de l'actif risqué et $v(t)$ est la variance instantanée des rendements de l'actif risqué. Les variations de la volatilité sont corrélées avec les rendements de l'actif risqué par la relation $\mathrm{Cov}(dB(t), dZ(t)) = \rho dt$.

Les préférences de l'investisseur sont décrites par la fonction d'utilité récursive de Duffie et Epstein (1992) définie par

$$J(t) = \mathrm{E}_t \left[\int_t^{+\infty} f(c(s), J(s))ds \right], \tag{4.13}$$

où $f(c, J)$ est de la forme :

$$f(c,J) = \begin{cases} \frac{\beta\gamma}{1-\frac{1}{\psi}} J \left[\left(\frac{c}{(\gamma J)^{\frac{1}{\gamma}}} \right)^{1-\frac{1}{\psi}} - 1 \right], & \text{si } \psi \in]0, +\infty[\setminus \{1\}, \\ \beta\gamma J \log\left(\frac{c}{(\gamma J)^{\frac{1}{\gamma}}} \right), & \text{si } \psi = 1, \end{cases} \tag{4.14}$$

avec $\psi > 0$. β est le taux de préférence temporelle et $1 - \gamma$ le coefficient d'aversion relative au risque.

L'investisseur a pour objectif de maximiser l'utilité espérée de sa consommation à horizon infini, autrement dit l'équation (4.13).

Lorsque $\psi = 1$, Chacko et Viceira (2005) ont démontré qu'il existe une solution explicite au problème de maximisation. La fonction d'utilité indirecte est donnée par

$$J(W(t), v(t)) = \exp\left\{ \frac{A}{v(t)} + B \right\} \frac{W(t)^\gamma}{\gamma}, \tag{4.15}$$

ce qui conduit aux expressions de la consommation optimale et du poids optimal de l'actif risqué

$$c(t) = \beta W(t), \tag{4.16}$$

$$w(t) = \frac{\mu - r}{(1-\gamma)v(t)} + \frac{\rho\sigma}{(1-\gamma)v(t)} A, \tag{4.17}$$

où A et B sont solutions du système d'équations

$$aA^2 + bA + c = 0, \tag{4.18}$$

$$\gamma\left(\beta\log(\beta) + r - \beta\right) - \beta B + \kappa\alpha A = 0, \tag{4.19}$$

avec $a = \frac{\sigma^2}{2\gamma(1-\gamma)}\left[1-\gamma(1-\rho^2)\right]$, $b = \frac{\rho\sigma(\mu-r)}{1-\gamma} - \frac{\beta+\kappa}{\gamma}$ et $c = \frac{(\mu-r)^2}{2(1-\gamma)}$. Le poids optimal de l'actif risqué a deux composantes : la première est la demande du portefeuille moyenne-variance et la deuxième est la demande de couverture intertemporelle de Merton. Les deux composantes sont des fonctions linéaires de $1/v(t)$, ce qui implique que leur rapport est indépendant du niveau de volatilité. Par ailleurs la couverture est toujours nulle lorsque la volatilité de l'actif risqué est nulle ($\sigma = 0$) ou lorsque qu'elle varie mais que l'investisseur ne peut pas utiliser l'actif risqué pour se couvrir contre la volatilité ($\rho = 0$).

Ce modèle permet de conclure que lorsque l'aversion au risque relative est supérieure ou égale à 1 ($\gamma \leq 0$) , la demande du portefeuille optimal pour des actions inclut une composante de couverture intertemporelle dont le signe est le même que celui de la corrélation instantanée ρ entre la volatilité et le rendement. En outre, à partir de données d'actions américaines, Chacko et Viceira (2005) montrent que leur volatilité n'est pas assez variable et persistante pour générer des grandes demandes de couverture intertemporelle.

4.3.2 Le modèle de Liu ou le modèle d'Heston adapté aux problèmes de choix de portefeuille

Liu (2007) utilise le modèle de Heston (1993) [41] pour décrire la volatilité stochastique des rendements de l'actif risqué. Sous les mêmes hypothèses que Merton (1971), Liu (2007) résout explicitement des problèmes de choix de portefeuille dynamique jusqu'à obtenir la solution d'une équation différentielle ordinaire, quand l'investisseur a une utilité CRRA. Puis, il applique sa solution à des problèmes qui jusqu'alors n'ont pas été traités dans la littérature et notamment le problème de choix de portefeuille à 2 actifs (l'un non risqué et l'autre risqué) avec une volatilité de type Heston. De plus, le modèle permet de fixer arbitrairement une corrélation entre le brownien du rendement et celui de la volatilité.

Dans le modèle d'Heston, la dynamique du prix de l'actif risqué et celle de la variance s'écrivent

$$\begin{cases} dS_1(t) = \mu S_1(t)dt + \sqrt{v(t)}S_1(t)dB(t), & (4.20) \\ dv(t) = \kappa(\alpha - v(t))dt + \sigma\sqrt{v(t)}dZ(t), & (4.21) \end{cases}$$

où $\mathrm{Cov}(dB(t), dZ(t)) = \rho dt$. La dynamique de l'actif non risqué a pour expression $dS_0(t) = rS_0(t)dt$.

L'investisseur, dont les préférences sont représentées par une fonction CRRA, cherche à maximiser l'utilité espérée de sa consommation avec héritage à l'échéance T

$$\max_{c(t),w(t)} \mathrm{E}\left[\int_0^T \theta \mathrm{e}^{-\delta t}\frac{c(t)^\gamma}{\gamma}dt + (1-\theta)\mathrm{e}^{-\delta T}\frac{W(T)^\gamma}{\gamma}\right]. \quad (4.22)$$

L'aversion au risque de l'investisseur est donné par $1-\gamma$.

41. Heston avait dérivé une solution à forme fermée pour le prix d'une option d'achat européenne à partir d'un actif avec volatilité stochastique. Il avait ensuite introduit des taux d'intérêt stochastiques et appliqué son modèle aux options sur obligation et aux options sur devises étrangères.

En utilisant l'approche par programmation dynamique, Liu détermine explicitement l'expression du poids optimal de l'actif risqué lorsqu'il n'y a pas de consommation ($\theta = 0$) :

$$w^*(t) = \frac{\mu - r}{(1-\gamma)v(t)} + \rho\sigma\frac{(\mu-r)^2}{(1-\gamma)v(t)}f(v,t), \tag{4.23}$$

avec

$$f(v,t) = \begin{cases} \dfrac{1}{\left(\frac{1-\gamma}{\gamma}\kappa v(t) - (\mu-r)\sigma\rho + \sqrt{A}\right) + \frac{\sqrt{A}e^{\sqrt{A}t/2}}{\sinh(\sqrt{A}t/2)}} & \text{si } A \geq 0, \\ \dfrac{1}{\left(\frac{1-\gamma}{\gamma}\kappa v(t) - (\mu-r)\sigma\rho\right) + \sqrt{A}\coth(\sqrt{A}t/2)} & \text{si } A \leq 0, \end{cases} \tag{4.24}$$

où

$$A = \left(\frac{1-\gamma}{\gamma}\kappa v(t) - (\mu-r)\sigma\rho\right)^2 - \left(1 + \frac{1-\gamma}{\gamma}\frac{1}{\rho^2}\right)((\mu-r)\sigma\rho)^2. \tag{4.25}$$

Comme pour le précédent modèle, le poids optimal de l'actif risqué possède deux composantes : la première est la demande du portefeuille efficient et la deuxième est la couverture intertemporelle. Ici aussi, la couverture est nulle lorsque $\sigma = 0$ ou lorsque $\rho = 0$. Mais le rapport entre la demande du portefeuille efficient et la demande de couverture n'est pas indépendant du niveau de volatilité puisque elle intervient dans la fonction $f(v,t)$.

Le poids de l'actif risqué dans le portefeuille est toujours positif et décroît de façon monotone vers 0 quand l'aversion au risque est très forte, si $\left(2\kappa\rho + \sigma\frac{\mu-r}{v(t)}\right)\sigma\frac{\mu-r}{v(t)} \leq 0$. De plus, il est négatif pour une faible aversion au risque et atteint une valeur maximale puis décroît vers 0 quand l'aversion au risque tend vers $+\infty$, si $\left(2\kappa\rho + \sigma\frac{\mu-r}{v(t)}\right)\sigma\frac{\mu-r}{v(t)} < 0$.

4.3.3 Le modèle de Benth, Karlsen et Reikvam

Benth, Karlsen, et Reikvam (2003) modélisent la volatilité de l'actif risqué comme l'ont suggéré Barndorff-Nielsen et Shephard (2001). Ces derniers proposent d'utiliser une somme pondérée de processus d'Ornstein-Uhlenbeck non gaussiens. De cette manière, le modèle possède plusieurs propriétés des séries temporelles de la volatilité telles que des queues épaisses et une dépendance à long terme. Par ailleurs, le niveau de volatilité peut occasionnellement connaître un bond brutal suivi d'une décroissance exponentielle. Benth, Karlsen, et Reikvam (2003) étudient le problème d'optimisation de portefeuille de Merton par l'approche par programme dynamique.

L'investisseur a le choix entre une obligation sans risque et une action dont les dynamiques des prix sont fournies par

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (4.26) \\ dS_1(t) = \left[\mu + \left(\frac{1}{2} + \beta\right)v(t)\right]S_1(t)dt + \sqrt{v(t)}S_1(t)dB(t), & (4.27) \\ v(t) = \left(\sum_{j=1}^{m}\omega_j Y_j(t)\right)^2, & (4.28) \end{cases}$$

où $v(t)$ est la variance des rendements. La dynamique des processus d'Ornstein-Uhlenbeck donnée par $dY_j(t) = -\lambda_j Y_j(t)dt + dZ_j(\lambda_j t)$, où $j = 1, \ldots, m$ et λ_j est un nombre positif. Les $\omega_j > 0$ dont la somme vaut 1, sont les poids accordés au processus j.

En l'absence de consommation, l'hypothèse du portefeuille auto-financé conduit à la dynamique de la richesse

$$dW(t) = \left\{ \left[\mu + \left(\frac{1}{2} + \beta \right) v(t) \right] w(t) - r \right\} W(t)dt + \sqrt{v(t)} w(t) W(t) dB(t). \tag{4.29}$$

Si l'objectif de l'investisseur est de maximiser l'utilité de sa richesse terminale :

$$\max_{w} \mathrm{E}_t \left[u(W(T)) \right], \tag{4.30}$$

le poids optimal de l'actif risqué, lorsque les préférences de l'investisseur sont caractérisées par une utilité CRRA, s'écrit :

– cas 1 : si $\frac{\frac{1}{2}+\beta}{1-\gamma} \geq 1$,

$$w^*(t) = 1 \quad \text{pour } v \in [0, +\infty[, \tag{4.31}$$

– cas 2 : si $\frac{\frac{1}{2}+\beta}{1-\gamma} \in]0, 1[$,

$$w^*(t) = \begin{cases} 1, & \text{si } v \in [0, v_0[, \\ \bar{w}(t) & \text{si } v \in [v_0, \infty[, \end{cases} \tag{4.32}$$

– cas 3 : si $\frac{\frac{1}{2}+\beta}{1-\gamma} < 0$,

$$w^*(t) = \begin{cases} 1, & \text{si } v \in [0, v_0[, \\ \bar{w}(t) & \text{si } v \in [v_0, v_1], \\ 0 & \text{si } v \in]v_1, \infty[, \end{cases} \tag{4.33}$$

où $v_0 = \frac{\mu - r}{(1-\gamma) - (\frac{1}{2}+\beta)}$, $v_1 = -\frac{\mu - r}{\frac{1}{2}+\beta}$ et la fonction $\bar{w}(t)$ est définie par

$$\bar{w}(t) = \frac{\mu - r}{(1-\gamma)v} + \frac{\frac{1}{2}+\beta}{1-\gamma}. \tag{4.34}$$

L'expression de $\bar{w}(t)$ est constituée de deux termes : le premier correspond à la demande de portefeuille efficient et la deuxième est la couverture intertemporelle. Nous remarquons qu'avec la définition de la volatilité de Benth, Karlsen, et Reikvam (2003), un aspect de $\bar{w}(t)$ est surprenant : la couverture intertemporelle est constante. Elle ne dépend ni du temps, ni de la volatilité. De plus, si la volatilité n'a pas d'influence sur le rendement espéré ($\beta = -1/2$) de l'actif risqué, celui-ci devient constant et le poids optimal de l'actif risqué vaut uniquement la demande de portefeuille efficient.

Conclusion

Ces modèles récents en temps continu présentent l'avantage de prendre en compte les caractéristiques de la volatilité des rendements d'action, ce qui améliore sensiblement la perception du risque du portefeuille et ainsi conduit à une meilleure allocation optimale de la richesse de l'investisseur que le permettaient les modèles statiques.

Outre le caractère stochastique, Chacko et Viceira (2005) et Liu (2007) s'attachent à modéliser le retour à la moyenne de la volatilité. Hélas, ces modèles font l'impasse sur une

caractéristique fondamentale de la volatilité : la mémoire longue (voir par exemple Campbell, Lo, et MacKinlay (1997)). Même le modèle de Benth, Karlsen, et Reikvam (2003) qui tente de prendre en compte la dépendance à longue portée de la volatilité, n'est pas véritablement un modèle à mémoire longue. Or, les modèles de volatilité prenant en compte ce phénomène donnent de meilleurs résultats en terme de prédiction que les autres : par exemple, le modèle ARFIMA est meilleur que le modèle ARMA et le modèle FIGARCH que le modèle GARCH (voir section 1.3.3). Cela est tout à fait logique puisque la valeur actuelle de la volatilité aura un impact sur le niveau de volatilité dans 6 mois et même au-delà. Par conséquent, la prise en compte de la mémoire longue dans les modèles de choix de portefeuille devrait permettre une meilleure gestion du portefeuille par "anticipation".

Chapitre 5

Modèle de choix de portefeuille en temps continu avec volatilité stochastique

Notre but dans cette deuxième partie de thèse est de développer un modèle de choix de portefeuille en temps continu avec une volatilité stochastique à mémoire longue. Pour y parvenir, nous nous basons sur le cas particulier du modèle de Merton (1971) en temps continu et à volatilité constante où la dynamique du prix de l'actif risqué suit un mouvement brownien géométrique, et procédons en deux étapes. Ce chapitre constitue la première d'entre elles et consiste à remplacer la volatilité constante par une volatilité stochastique. Dans ce but, nous modélisons la log-volatilité de l'actif risqué par un processus d'Ornstein-Uhlenbeck qui présente l'avantage de pouvoir capturer de manière simple le phénomène de retour à la moyenne de la volatilité. En plus de l'intérêt théorique, le développement d'un tel modèle nous permet d'étudier l'impact du caractère stochastique de la volatilité de l'actif risqué sur l'allocation optimale.

Pour que ce chapitre soit le plus complet possible, nous nous intéressons à 3 critères de gestion à horizon fini : la maximisation de l'espérance de l'utilité de la consommation avec héritage, puis sans héritage et la seule maximisation de l'héritage. Ici, la notion d'héritage n'est pas comprise au sens successoral, mais au sens de ce qu'il reste à la fin de la période de gestion. Nous considérons la fonction d'utilité caractérisant l'aversion absolue au risque hyperbolique (ci-après notée HARA) pour les deux premiers critères précédemment cités. Cette classe de fonctions d'utilité possède la particularité de pouvoir retrouver, pour certaines valeurs d'autres fonctions d'utilité, en particulier celle caractérisant l'aversion relative au risque constante (notée ci-après CRRA). Celle-ci présente l'intérêt de nous permettre d'étudier plus facilement l'impact de la volatilité stochastique sur la gestion de portefeuille puisqu'elle entraîne la suppression de deux paramètres de notre problème.

Pour résoudre les problèmes de choix de portefeuille que nous nous posons, nous utilisons l'approche par programmation dynamique. Une telle approche suppose que la volatilité soit connue au temps t. Or, nous l'avons vu dans la première partie de la thèse, la volatilité

contrairement au cours de l'action, n'est pas observable directement sur le marché. Elle est calculée grâce à des données historiques au moyen par exemple de l'estimateur de la volatilité réalisée. Toutefois, tous les acteurs du marché se font une idée du niveau de volatilité actuelle et peuvent être tentés de l'introduire dans leur stratégie de gestion. Ainsi, l'approche par programmation dynamique se révèle pertinente.

Ce chapitre est composé de trois sections, les deux premières axées uniquement sur la dérivation des résultats analytiques et la troisième dédiée aux illustrations numériques. En particulier, la section 1 est consacrée au modèle à volatilité constante de Merton. Dans la section 2, un modèle avec une volatilité stochastique est développé à partir de ce dernier. Enfin, dans une troisième section, nous procédons aux simulations et analysons les résultats obtenus.

5.1 Modèle à volatilité constante

Dans son article de 1971, Merton propose un modèle de choix de portefeuille dans lequel le cours de l'actif risqué est représenté par une équation de diffusion générale. Nous reprenons ici un cas particulier de son modèle pouvant se traiter explicitement. Il s'agit du cas où la dynamique de l'actif risqué est caractérisé par un mouvement brownien géométrique et où l'investisseur cherche à maximiser l'utilité espérée de la consommation seule et considère une fonction d'utilité de type HARA. Autrement dit, à l'horizon de gestion T, aucun héritage n'est versé, ce qui n'est pas très courant dans la réalité. En effet, nous avons plutôt tendance à rencontrer des personnes cherchant à accumuler un patrimoine pour financer leur retraite tout en consommant. De plus, un inconvénient apparaît à l'approche de l'horizon de gestion : puisque dans ce critère, il n'y pas d'héritage au temps T, la consommation augmente fortement juste avant l'échéance pour ne laisser aucun capital.

Nous décidons donc de généraliser le modèle de Merton pour qu'il puisse traiter d'autres problèmes de maximisation : celui de la maximisation de l'utilité espérée de la consommation avec héritage à l'horizon de gestion T et celui de la maximisation de l'utilité espérée de sa richesse finale en T. Pour le premier critère, il peut s'agir, par exemple, d'un investisseur qui finance sa consommation tant qu'il est en vie et qui, à son décès en T, lègue un capital à ses héritiers. Quant au deuxième critère, c'est un pur cas de choix de portefeuille où, à partir d'un capital initial, l'investisseur cherche à le faire fructifier au mieux à un horizon de placement donné.

5.1.1 Définition du modèle

Le modèle à volatilité constante de Merton (1971) dans le cas particulier qui nous intéresse évoqué ci-dessus s'écrit

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (5.1) \\ dS_1(t) = \mu S_1(t)dt + \sigma_s S_1(t)dB(t), & (5.2) \end{cases}$$

où $S_0(t)$ représente la valeur d'une unité de l'actif sans risque au temps t et $S_1(t)$ celle d'une unité de l'actif risqué au temps t. Le taux sans risque est caractérisé par r, le rendement moyen de l'actif risqué par μ et sa volatilité par σ_s, une constante.

Soit $W(t)$ la richesse totale au temps t et $w(t)$ la fraction de richesse investie dans l'actif risqué. La richesse au temps $t + dt$ est égale à la richesse au temps t augmentée de l'accroissement entre t et $t + dt$ des richesses investies dans les deux actifs, risqué et non risqué, à laquelle on aura retranché la consommation $c(t)dt$ du gestionnaire sur la période considérée. En résumé, l'équation de la richesse sous forme différentielle est donnée par

$$dW(t) = [(\mu - r)w(t)W(t) + rW(t) - c(t)]\, dt + w(t)W(t)\sigma_s dB(t). \tag{5.3}$$

Les processus $c(t)$ et $w(t)$, $t \geq 0$, sont adaptés [42] à la filtration naturelle, notée $\mathcal{F}_t$, associée à $B(t)$.

Déterminer l'allocation du portefeuille optimal revient à résoudre un problème de maximisation. Nous allons traiter le problème général, comme le fait Liu (2007), en maximisant l'utilité espérée de la consommation avec héritage :

$$\max_{c(t),w(t)} \mathrm{E}\left[\int_0^T \theta \mathrm{e}^{-\lambda t} u(c(t))dt + (1-\theta)H(W(T),T)\right], \tag{5.4}$$

où $H(W(T),T)$ est la fonction héritage et $u(x)$ la fonction d'utilité strictement concave pour modéliser un investisseur risquophobe. Le paramètre λ caractérise l'impatience à consommer [43] de l'investisseur : plus λ est grand, plus la préférence pour le présent de l'agent économique est prononcé. Ce paramètre peut aussi être interprété comme le taux d'escompte. Quant au paramètre θ, il a été introduit pour traiter une multitude de critères de gestion, notamment les trois suivants :

- si $\theta = 1/2$: nous nous intéressons au problème classique de maximisation de l'utilité espérée de la consommation avec héritage à l'échéance T,
- si $\theta = 1$: nous maximisons l'utilité espérée de la consommation sans héritage à l'échéance T,
- si $\theta = 0$: nous nous penchons sur le cas où il n'y a pas de consommation (et donc $\lambda = 0$), mais uniquement un héritage à la date T.

Puisque nous allons faire appel à une fonction d'utilité correspondant à l'aversion absolue au risque hyperbolique (HARA)

$$u(c) = \frac{1-\gamma}{\gamma}\left(\frac{\beta c}{1-\gamma} + \eta\right)^{\gamma}, \tag{5.5}$$

où $\gamma < 1$, $\gamma \neq 0$ et $\beta > 0$, nous posons, pour obtenir une solution explicite du problème :

$$H(W(T),T) = \mathrm{e}^{-\lambda T}\frac{W(T)^{\gamma}}{\gamma}. \tag{5.6}$$

42. Cela signifie que $c(t)$ et $w(t)$ sont mesurables au temps t, ce qui veut dire, concrètement, que la consommation optimale et le poids optimal de l'actif risqué sont observables au temps t.

43. Plus λ est grand, plus le terme $\mathrm{e}^{-\lambda t}$ diminuera rapidement lorsque t augmentera. Donc moins d'importance sera accordée à l'utilité de la consommation au temps t' par rapport à celle au temps $t < t'$, ce qui caractérise une impatience à consommer. Le lecteur est invité à se référer à Ingersoll (1987) pour plus de détails.

L'objectif de l'investisseur se traduit donc ainsi :

$$\max_{c(t),w(t)} \mathrm{E}\left[\int_0^T \theta \mathrm{e}^{-\lambda t} u(c(t))dt + (1-\theta)\mathrm{e}^{-\lambda T}\frac{W(T)^\gamma}{\gamma}\right]. \tag{5.7}$$

Remarque 1. La fonction d'utilité CRRA $u(c) = \frac{c^\gamma}{\gamma}$, où $\gamma < 1, \gamma \neq 0$, s'obtient en fixant $\eta = 0$ et $\beta = (1-\gamma)^{-\frac{1-\gamma}{\gamma}}$.

Pour déduire les équations d'optimalité, l'équation (5.7) doit être réécrite sous forme de programme dynamique pour pouvoir appliquer le principe d'optimalité de Hamilton-Jacobi-Bellman. Étant donné que d'une part ce résultat est classique et que d'autre part nous proposons une extension de ce modèle dans la section suivante, nous n'entrerons pas ici dans les détails, mais donnerons uniquement les expressions du poids optimal de l'actif risqué et le cas échéant la consommation optimale de l'investisseur.

5.1.2 Étude des critères de gestion avec consommation de l'investisseur

Les critères de gestion avec consommation de l'investisseur correspondent au cas où $\theta = 0$. Lorsque l'investisseur a une utilité HARA et cherche à maximiser l'utilité espérée de sa consommation avec ou sans héritage à l'horizon de gestion T, les expressions de la consommation optimale $c^*(t)$ et du poids optimal de l'actif risqué $w^*(t)$ sont :

$$\begin{cases} c^*(W,t) = \dfrac{W}{F(t)} + K_1\left(\dfrac{1-\mathrm{e}^{-r(T-t)}}{r}\dfrac{1}{F(t)} - 1\right), & (5.8) \\ w^*(W,t) = \dfrac{\mu - r}{(1-\gamma)\sigma_s^2} + K_1\dfrac{\mu-r}{(1-\gamma)\sigma_s^2}\dfrac{1-\mathrm{e}^{-r(T-t)}}{r}\dfrac{1}{W}, & (5.9) \end{cases}$$

où

$$F(t) = \frac{1}{K_2} + \left[\frac{\beta^{-\frac{\gamma}{1-\gamma}}}{1-\gamma}\left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\gamma}} - \frac{1}{K_2}\right]\mathrm{e}^{-K_2(T-t)}, \tag{5.10}$$

avec $K_1 := \frac{(1-\gamma)\eta}{\beta}$, $K_2 := \frac{\lambda - \gamma\nu}{1-\gamma}$ où $\nu := r + \frac{(\mu-r)^2}{2(1-\gamma)\sigma_s^2}$.

Les formulations (5.8) et (5.9) sont valables quel que soit le critère de gestion utilisé. En effet, θ n'intervient qu'au travers de la fonction $F(t)$, présente dans l'expression de la consommation optimale : nous observons dans l'équation (5.10) que $F(t)$ décroît en fonction de θ, par déduction la consommation optimale croit avec θ, ce à quoi il fallait s'attendre étant donné l'objectif de l'investisseur (voir équation 5.7). En effet, plus θ est grand, plus l'investisseur accorde d'importance à sa consommation agrégée au dépend de l'héritage.

Quant au poids optimal de l'actif risqué dans le portefeuille, nous constatons qu'il est composé de deux termes : un premier constant, la demande de portefeuille efficient, et un second, variable dans le temps. Si la richesse du portefeuille augmente, alors le deuxième terme diminue en valeur absolue et le poids optimal se rapproche de celui pour une fonction CRRA (voir ci-après). Et inversement. Le signe de ce second terme dépend du signe de η (car $\beta > 0$ et $\gamma < 1$) : si η est négatif, alors il est négatif, et inversement si η est positif.

Remarque 2. Pour une fonction CRRA, les formulations de l'équation aux dérivées partielles, de la consommation optimale, du poids optimal de l'actif risqué et de la fonction d'utilité indirecte restent valable. Il suffit d'attribuer à η et β les valeurs adéquates précédemment citées (voir remarque 1). Dans ce cas, $K_1 = 0$ et les équations (5.8) et (5.9) se simplifient :

$$\begin{cases} c^*(W,t) = \dfrac{W}{F(t)}, & (5.11) \\ w^* = \dfrac{\mu - r}{(1-\gamma)\sigma_s^2}. & (5.12) \end{cases}$$

Le poids optimal de l'actif risqué est constant. Par ailleurs, nous constatons que deux facteurs agissent conjointement dans la proportion d'actif risqué dans le portefeuille optimal. Premièrement le risque : plus il est grand, plus le poids optimal de l'actif risqué est petit et inversement. Deuxièmement, l'aversion au risque de l'investisseur, $1-\gamma$: plus $1-\gamma$ est grand (investisseur risquophobe) plus le poids de l'actif risqué est petit et inversement, ce qui est logique.

5.1.3 Étude du critère de gestion sans consommation de l'investisseur

Pour étudier le critère de gestion sans consommation (ce qui se traduit par $\theta = 0$), nous cherchons uniquement à maximiser l'héritage à la date T. L'expression du poids optimal de l'actif risqué est

$$w^* = \frac{\mu - r}{(1-\gamma)\sigma_s^2}. \tag{5.13}$$

Dans ce cas, nous retrouvons l'écriture du poids optimal de l'actif risqué pour un investisseur ayant une utilité CRRA et pour objectif de maximiser l'utilité espérée de sa consommation avec ou sans héritage.

5.2 Modèle à volatilité stochastique

Nous reprenons les équations (5.1) et (5.2), en y apportant une modification : nous remplaçons la volatilité constante σ_s de l'actif risqué par une volatilité stochastique $V(t)$, contrôlée par une équation différentielle stochastique. Une méthode classique consiste à modéliser la volatilité par un processus en racine carrée du type CIR (Cox, Ingersoll, et Ross 1985). Nous choisissons une autre méthode qui consiste à utiliser un processus d'Ornstein-Uhlenbeck, non pas pour décrire la volatilité (ce n'est pas possible car cela voudrait dire que la volatilité pourrait prendre des valeurs négatives), mais la log-volatilité qui, elle, peut prendre des valeurs négatives. Notre objectif est d'utiliser un processus de retour à la moyenne puisqu'il s'agit d'une des caractéristiques de la volatilité des rendements (voir Campbell, Lo, et MacKinlay (1997), Fouque, Papanicolaou, et Sircar (2000)).

5.2.1 Définition du modèle

Notre modèle avec volatilité stochastique s'écrit

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (5.14)\\ dS_1(t) = \mu S_1(t)dt + V(t)S_1(t)dB(t), & (5.15)\\ d(\ln V(t)) = \kappa(\alpha - \ln V(t))dt + \sigma dZ(t), & (5.16) \end{cases}$$

avec $\mathrm{E}[dB(t)dZ(t)] = \rho dt$ et $\mathrm{E}[dB^2(t)] = \mathrm{E}[dZ^2(t)] = dt$. Nous reprenons les mêmes notations que dans la section précédente à la différence que la volatilité constante σ_s est remplacée par une volatilité stochastique $V(t)$. Nous soulignons le fait que nous avons délibérément maintenu μ constant afin que l'étude permette véritablement de se concentrer sur les effets de variation de la volatilité. Enfin, κ caractérise le temps de retour à la moyenne du processus d'Ornstein-Uhlenbeck de la log-volatilité, α la moyenne à long terme de la log-volatilité et σ le coefficient de diffusion [44] de la log-volatilité.

Pour plus de simplicité et de compréhension, nous donnons l'expression de l'équation de diffusion de la volatilité. En appliquant le théorème d'Itō (voir théorème 1 en annexe 5.A) à la fonction $f(X(t)) = \mathrm{e}^{X(t)}$ où $X(t) = \ln V(t)$, nous obtenons l'équation différentielle stochastique de la volatilité et le système devient :

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (5.17)\\ dS_1(t) = \mu S_1(t)dt + V(t)S_1(t)dB(t), & (5.18)\\ dV(t) = \left(\dfrac{\sigma^2}{2} + \kappa(\alpha - \ln V(t))\right)V(t)dt + \sigma V(t)dZ(t), & (5.19) \end{cases}$$

Remarque 3. En posant $\sigma = 0$ et $\kappa = 0$, alors $dV(t) = 0$. Cela implique que l'on retombe sur le cas traité précédemment où la volatilité était constante.

En reprenant les mêmes notations que dans la section précédente, l'équation de la richesse en notation différentielle s'écrit

$$dW(t) = \left[(\mu - r)w(t)W(t) + rW(t) - c(t)\right]dt + w(t)W(t)V(t)dB(t), \tag{5.20}$$

et les équations permettant de décrire l'état du système à un instant t se résument à

$$\begin{cases} dW(t) = \left[(\mu - r)w(t)W(t) + rW(t) - c(t)\right]dt + w(t)W(t)V(t)dB(t), & (5.21)\\ dV(t) = \left(\dfrac{\sigma^2}{2} + \kappa(\alpha - \ln V(t))\right)V(t)dt + \sigma V(t)dZ(t). & (5.22) \end{cases}$$

L'investisseur poursuit le même objectif que dans la section précédente :

$$\max_{c(t),w(t)} \mathrm{E}\left[\int_0^T \theta \mathrm{e}^{-\lambda t}u(c(t))dt + (1-\theta)\mathrm{e}^{-\lambda T}\frac{W(T)^\gamma}{\gamma}\right]. \tag{5.23}$$

44. Nous attirons l'attention du lecteur sur le fait que σ n'est pas la volatilité de la log-volatilité, mais seulement une composante comme nous l'expliquons au moyen de l'équation (5.44) dans la sous-section 5.3.2.

Pour déduire les équations d'optimalité, l'équation (5.23) doit être réécrite sous forme de programme dynamique pour pouvoir appliquer le principe d'optimalité de Hamilton-Jacobi-Bellman. On définit pour cela la fonction d'utilité indirecte

$$J(W,V,t) = \max_{c(t),w(t)} \mathrm{E}_t \left[\int_t^T \theta \mathrm{e}^{-\lambda s} u(c(s)) ds + (1-\theta)\mathrm{e}^{-\lambda t} \frac{W(T)^\gamma}{\gamma} \right], \tag{5.24}$$

où E_t représente l'espérance conditionnelle étant donnée l'information jusqu'au temps t. Nous notons $V(t) = V$ et $W(t) = W$.

D'après le théorème d'Hamilton-Jacobi-Bellman, la fonction d'utilité indirecte $J(W,V,t)$ vérifie l'équation

$$0 = \max_{c(t),w(t)} \left\{ \theta \mathrm{e}^{-\lambda t} u(c(t)) + \mathcal{L}J(W,V,t) + J_t \right\}, \tag{5.25}$$

où $\mathcal{L}J(W,V,t)$ est l'opérateur différentiel générateur de J ayant pour expression

$$\begin{aligned} \mathcal{L}J(W,V,t) = {} & \left[(\mu - r)w(t)W + rW - c(t)\right] J_W + V \left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V) \right] J_V \\ & + \frac{\sigma^2}{2} V^2 J_{VV} + \frac{1}{2} w(t)^2 W^2 V^2 J_{WW} + \rho\sigma w(t) W V^2 J_{VW}. \end{aligned} \tag{5.26}$$

Dorénavant, nous sommes prêts à étudier les trois principaux critères de gestion de l'investisseur.

5.2.2 Étude des critères de gestion avec consommation de l'investisseur

Les critères de gestion avec consommation de l'investisseur correspondent au problème de maximisation de l'utilité espérée de la consommation avec ou sans héritage (5.23) pour $\theta \neq 0$. La généralisation au cadre stochastique des expressions de la consommation optimale $c^*(W,V,t)$ et du poids de l'actif risqué dans la richesse totale $w^*(W,V,t)$, que nous démontrons dans l'annexe 5.B, est donnée par

$$\begin{cases} c^*(W,V,t) = \dfrac{W}{F(V,t)} + K_1 \left(\dfrac{1-\mathrm{e}^{-r(T-t)}}{r} \dfrac{1}{F(V,t)} - 1 \right), & (5.27) \\ w^*(W,V,t) = \dfrac{\mu - r}{(1-\gamma)V^2} + K_1 \dfrac{\mu - r}{(1-\gamma)V^2} \dfrac{1-\mathrm{e}^{-r(T-t)}}{r} \dfrac{1}{W} \\ \qquad\qquad + \rho\sigma \dfrac{F_V(V,t)}{F(V,t)} \left(1 + K_1 \dfrac{1-\mathrm{e}^{-r(T-t)}}{r} \dfrac{1}{W} \right), & (5.28) \end{cases}$$

où $K_1 := \frac{(1-\gamma)\eta}{\beta}$ et $F(V,t)$ est solution de l'équation aux dérivées partielles suivante :

$$\begin{aligned} 0 = {} & 1 - K_2(V)F + F_t + \left\{ V \left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V) \right] + \frac{\gamma}{1-\gamma}(\mu - r)\rho\sigma \right\} F_V \\ & + \frac{\sigma^2}{2} V^2 F_{VV} - \gamma(1-\rho^2) \frac{\sigma^2}{2} V^2 \frac{F_V^2}{F}, \end{aligned} \tag{5.29}$$

avec la condition limite $F(V,T) = \frac{\beta^{-\frac{\gamma}{1-\gamma}}}{1-\gamma}\left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\gamma}}$ pour tout V, où $\theta \neq 0$ et $K_2(V) := \frac{\lambda-\gamma\nu(V)}{1-\gamma}$ où $\nu(V) := r + \frac{(\mu-r)^2}{2(1-\gamma)V^2}$.

L'expression de la consommation optimale $c^*(t)$ est quasiment la même que pour le cas à volatilité constante : seule la fonction $F(t)$ est devenue $F(V,t)$. Quant à l'expression du poids optimal de l'actif risqué $w^*(V,t)$, un terme a été rajouté par rapport à l'équation (5.9) : celui où $\rho\sigma\frac{F_V(V,t)}{F(V,t)}$ est en facteur. Les deux premiers termes représentent la généralisation triviale du terme classique de la demande en actif risqué, et le troisième terme représente la demande de couverture intertemporelle qui, en regardant de plus près l'équation (5.28), est toujours nulle dans deux cas. Premièrement, lorsque le coefficient de diffusion de la log-volatilité est nulle. Cela signifie que la volatilité est déterministe, donc non risquée et par conséquent il n'y a besoin de couverture. Deuxièmement, lorsque l'investisseur ne peut avoir recours à l'actif risqué pour se couvrir des changements de volatilité ($\rho = 0$).

Remarque 4. Pour une utilité CRRA de l'investisseur, la consommation optimale et le poids optimal de l'actif risqué se simplifie

$$\begin{cases} c^*(W,V,t) = \dfrac{W}{F(V,t)}, & (5.30) \\ w^*(V,t) = \dfrac{\mu-r}{(1-\gamma)V^2} + \rho\sigma\dfrac{F_V(V,t)}{F(V,t)}. & (5.31) \end{cases}$$

Nous constatons que la richesse W n'intervient plus dans l'expression de $w^*(V,t)$ qui possède deux composantes. La première est la demande d'actif risqué du modèle de Merton, pour une utilité CRRA de l'investisseur. La deuxième est la demande de couverture intertemporelle, qui comme pour une utilité HARA est nulle lorsque $\sigma = 0$ et $\rho = 0$.

Toutes les équations aux dérivées partielles ne se résolvent pas analytiquement. Hélas, c'est le cas ici pour l'équation (5.29). Il faudra donc employer une méthode numérique, qui sera détaillée ultérieurement dans la section 5.3.2, pour obtenir une approximation de $F(V,t)$.

5.2.3 Étude du critère de gestion sans consommation de l'investisseur

Lorsque nous fixons $\theta = 0$, nous retrouvons un problème uniquement de choix de portefeuille : l'investisseur a pour objectif de maximiser l'utilité de sa richesse finale en T. L'expression du poids optimal de l'actif risqué $w^*(V,t)$ devient

$$w^*(V,t) = \frac{\mu-r}{(1-\gamma)V^2} + \rho\sigma\frac{F_V(V,t)}{F(V,t)}. \tag{5.32}$$

où $F(V,t)$ est solution de l'équation aux dérivées partielles

$$0 = \frac{\gamma\nu(V)}{1-\gamma}F + F_t + \left\{V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] + \frac{\gamma}{1-\gamma}(\mu-r)\rho\sigma\right\}F_V + \frac{\sigma^2}{2}V^2F_{VV} - \gamma\frac{\sigma^2}{2}V^2(1-\rho^2)\frac{F_V^2}{F}, \tag{5.33}$$

avec la condition limite $F(V,T) = |\gamma|^{-\frac{1}{1-\gamma}}$ pour tout V, et où $\nu(V) := r + \frac{(\mu-r)^2}{2(1-\gamma)V^2}$.

Nous retombons sur l'expression du poids optimal de l'actif risqué dans le portefeuille (5.31) d'un individu dont les préférences sont caractérisées par une fonction CRRA, et qui cherche à maximiser l'utilité espérée de sa consommation avec héritage. Naturellement, les poids ne seront pas les mêmes puisque les équations aux dérivées partielles diffèrent légèrement, ce qui conduit à l'obtention de deux fonctions F distinctes.

5.3 Étude de l'impact de la volatilité stochastique en gestion de portefeuille

Dans cette section, nous commençons par définir le problème étudié, puis nous détaillons la méthode que nous avons développée pour résoudre numériquement l'équation aux dérivées partielles simplifiée du problème et la manière par laquelle nous avons étudié l'impact de la volatilité stochastique sur l'allocation optimale. Ensuite, nous fournissons le paramétrage du modèle. Enfin, nous présentons les résultats de cette étude.

5.3.1 Définition du problème étudié

Le problème est défini par le choix d'une fonction d'utilité et d'un critère de gestion. Puisque le modèle possède un grand nombre de paramètres, nous avons décidé de nous intéresser à la fonction d'utilité CRRA. Celle-ci nous permettra de supprimer β et η de l'équation du problème. En posant $\beta = (1-\gamma)^{-\frac{1-\gamma}{\gamma}}$ et $\eta = 0$, la fonction d'utilité s'écrit

$$u(c) = \frac{c^\gamma}{\gamma}, \qquad \gamma < 1 \text{ et } \gamma \neq 0. \tag{5.34}$$

L'aversion relative au risque est donnée par $1-\gamma$. Si γ se rapproche de 1, l'investisseur tend à être indifférent au risque, si γ prend de faibles valeurs négatives (de l'ordre de -2), il est moyennement risquophobe et si γ prends de grandes valeurs négatives (de l'ordre de -20), alors il est très risquophobe. L'utilisation de cette fonction d'utilité permet aussi de simplifier les expressions de la fonction d'utilité indirecte, de l'équation aux dérivées partielles simplifiée du problème, de la consommation optimale et du poids optimal de l'actif risqué. Nous allons résumer ces modifications dans le cas d'une volatilité constante, puis dans celui d'une volatilité stochastique.

Par ailleurs, nous décidons de traiter le problème de maximisation de l'utilité espérée de la consommation avec héritage à l'échéance T en posant $\theta = 1/2$. Nous rappelons les expressions de la consommation optimale et du poids optimal de l'actif risqué dans le portefeuille dans les deux cas de figures :

Cas 1 : volatilité constante

La gestion optimale est définie par

$$\begin{cases} c^*(W,t) = \dfrac{W}{F(t)}, & (5.35) \\ w^* = \dfrac{\mu - r}{(1-\gamma)\sigma_s^2}, & (5.36) \end{cases}$$

avec $F(t)$ ayant pour expression

$$F(t) = \frac{1}{K_2}\left[1 + (K_2 - 1)\mathrm{e}^{-K_2(T-t)}\right], \tag{5.37}$$

et où $K_2 := \frac{\lambda - \gamma\nu}{1-\gamma}$ et $\nu := r + \frac{(\mu-r)^2}{2(1-\gamma)\sigma_s^2}$.

Cas 2 : volatilité stochastique

La gestion optimale est définie par

$$\begin{cases} c^*(W,V,t) = \dfrac{W}{F(V,t)}, & (5.38) \\ w^*(V,t) = \dfrac{\mu - r}{(1-\gamma)V^2} + \rho\sigma\dfrac{F_V(V,t)}{F(V,t)}, & (5.39) \end{cases}$$

où $F(V,t)$ est la solution de l'équation aux dérivées partielles non linéaire

$$\begin{aligned} 0 = {} & 1 - K_2(V)F(V,t) + F_t(V,t) \\ & + \left\{ V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] + \frac{\gamma}{1-\gamma}(\mu - r)\rho\sigma \right\} F_V(V,t) \\ & + \frac{\sigma^2}{2}V^2 F_{VV}(V,t) - \gamma\frac{\sigma^2}{2}V^2(1-\rho^2)\frac{F_V(V,t)^2}{F(V,t)}, \end{aligned} \tag{5.40}$$

sous la condition limite $F(V,T) = 1$, pour tout V et où $K_2(V) := \frac{\lambda - \gamma\nu(V)}{1-\gamma}$ avec $\nu(V) := r + \frac{(\mu-r)^2}{2(1-\gamma)V^2}$.

5.3.2 Méthodologie

Pour obtenir les valeurs de la consommation optimale et du poids optimal de l'actif risqué dans le cadre du modèle à volatilité stochastique, il nous faut résoudre l'équation aux dérivées partielles non linéaire (5.40). Nous proposons ici une méthode itérative de résolution numérique. Ensuite, puisque notre objectif initial est de mesurer l'impact de la volatilité stochastique sur l'allocation optimale, nous détaillons notre manière de procéder.

Résolution numérique de l'équation aux dérivées partielles

Nous commençons par réécrire l'équation aux dérivées partielles (5.40) :

$$F_t(V,t) = A + BF(V,t) + CF_V(V,t) + DF_{VV}(V,t) + E\frac{F_V(V,t)^2}{F(V,t)}, \tag{5.41}$$

pour tout V et pour tout $t \in [0,T]$, avec les coefficients A, B, C, D et E ayant pour expression

$$\begin{aligned}
A &= -1,\\
B &= \frac{\lambda - \gamma\nu(V)}{1-\gamma},\\
C &= -\left[V\left(\frac{\sigma^2}{2} + \kappa(\alpha - \log V\right) + \frac{\gamma(\mu - r)}{1-\gamma}\rho\sigma\right],\\
D &= -\frac{\sigma^2}{2}V^2,\\
E &= \gamma(1-\rho^2)\frac{\sigma^2}{2}V^2,
\end{aligned}$$

où $F(V,T) = 1$ et $\nu(V) := r + \frac{(\mu-r)^2}{2(1-\gamma)V^2}$.

Dans de nombreux problèmes où nous sommes amenés à résoudre une équation aux dérivées partielles d'une fonction à deux variables (espace et temps), comme l'équation de la chaleur en physique, nous avons l'habitude de discrétiser ces deux variables en même temps. Cela est possible lorsqu'en plus de la condition limite, nous connaissons les conditions aux bords. Dans le cas qui nous concerne, nous ne connaissons ni $F(V_{min},t)$ ni $F(V_{max},t)$. Par conséquent, il n'est pas possible d'appliquer ce schéma de résolution. Nous proposons donc un autre schéma dont l'algorithme est le suivant :

1. Discrétiser le temps : $t = [0, \tau, 2\tau, \cdots, T-\tau, T]$, où $\tau = T/N$,
2. Partir de la condition limite $F(V,T)$ connue (voir équation (5.41)),
   (a) En déduire $F_V(V,T) = 0$ et $F_{VV}(V,T) = 0$,
   (b) Obtenir $F_t(V,T) = B - 1$ grâce à l'équation (5.41),
   (c) Enfin, obtenir $F(V,T-\tau) \approx F(V,T) - \tau F_t(V,T)$,
3. Ensuite, pour $k = 1 \cdots N-1$,
   (a) Partir de $F(V,T-k\tau)$ qui vient d'être calculée,
   (b) En déduire $F_V(V,T-k\tau)$ et $F_{VV}(V,T-k\tau)$ par la méthode des différences finies,
   (c) Calculer $F_t(V,T-k\tau)$ grâce à l'équation (5.41),
   (d) Enfin, obtenir $F(V,T-(k+1)\tau) \approx F(V,T-k\tau) - \tau F_t(V,T-k\tau)$,

Cet algorithme nous fournit donc pour tout t dans $[0,\tau,2\tau,\cdots,T]$ $F(V,t)$, naturellement, mais aussi $F_V(V,t)$ nécessaire à l'expression de $w^*(V,t)$ dans l'équation (5.39). Ensuite, il ne nous reste plus qu'à discrétiser l'intervalle de volatilité sur lequel nous voulons calculer $c^*(V,t)$ et $w^*(V,t)$ et de répéter cet algorithme pour chaque point de la discrétisation.

A noter que l'algorithme nécessite un réglage fin de la discrétisation du temps pour converger. Les conditions de convergence n'ont pas été étudiées, mais la compatibilité des résultats obtenus par ce modèle et celui à volatilité constante a été vérifiée.

Mesure de l'impact de la volatilité stochastique

Lorsque la volatilité est stochastique, elle va s'écarter de la valeur de la volatilité constante σ_s, qui correspond à sa valeur moyenne de long terme, en prenant des valeurs plus grandes ou plus petites. C'est ce que l'on appelle par la suite la *stochasticité* de la volatilité. Dans cette sous-section, nous voulons étudier l'impact d'une part de la stochasticité et d'autre part du degré de stochasticité de la volatilité sur la consommation optimale et le poids optimal de l'actif risqué.

En fixant $\kappa = 0$, et $\sigma = 0$, alors d'après (5.16) $d\ln V(t) = 0$ et par conséquent $\ln V(t) = \ln V(0) = \alpha$: la volatilité du modèle à volatilité stochastique devient constante. Concrètement, $\sigma = 0$ signifie que la volatilité est déterministe puisqu'elle vérifierait $dV(t) = \kappa\left(\alpha - \ln V(t)\right) dt$ et $\kappa = 0$ que le retour à la moyenne est infini, autrement dit qu'il n'y a pas de retour à la moyenne. Le degré de stochasticité va être donné par ces deux paramètres. Regardons de plus près leur rôle respectif.

En considérant l'équation différentielle stochastique de la log-volatilité (5.16) et en appliquant le lemme d'Itō (voir théorème 1 en annexe 5.A) à la fonction $f(X(t), t) = X(t)\mathrm{e}^{\kappa t}$, nous obtenons une nouvelle formulation de la log-volatilité :

$$\ln V(t) = \alpha + \mathrm{e}^{-\kappa t}(\ln V(0) - \alpha) + \sigma \int_0^t \mathrm{e}^{-\kappa(t-s)} dZ(s). \tag{5.42}$$

La moyenne et la variance de la log-volatilité s'obtiennent trivialement :

$$\mathrm{E}\left[\ln V_t\right] = \alpha + \mathrm{e}^{-\kappa t}(\ln V(0) - \alpha), \tag{5.43}$$

$$\mathrm{Var}\left[\ln V_t\right] = \frac{\sigma^2(1 - \mathrm{e}^{-2\kappa t})}{2\kappa}. \tag{5.44}$$

Or, s'il est intéressant de connaître ces deux données, il est encore plus significatif de connaître la moyenne et la variance de la volatilité. Nous avons admis, dans la première partie de la thèse, l'hypothèse que la densité de la log-volatilité est une densité de loi normale. Ainsi, la volatilité suit une loi log-normale dont les paramètres sont donnés par (5.43) et (5.44), et nous en déduisons que

$$\mathrm{E}\left[V_t\right] = \mathrm{e}^{\alpha + \mathrm{e}^{-\kappa t}(\ln V(0) - \alpha) + \frac{\sigma^2(1 - \mathrm{e}^{-2\kappa t})}{4\kappa}}, \tag{5.45}$$

$$\mathrm{Var}\left[V_t\right] = \left(\mathrm{e}^{\frac{\sigma^2(1 - \mathrm{e}^{-2\kappa t})}{2\kappa}} - 1\right) \mathrm{e}^{2(\alpha + \mathrm{e}^{-\kappa t}(\ln V(0) - \alpha)) + \frac{\sigma^2(1 - \mathrm{e}^{-2\kappa t})}{2\kappa}}. \tag{5.46}$$

Nous constatons que le niveau moyen de volatilité dépend du temps, de σ et de κ. De même, la volatilité de la volatilité qui représente une part d'incertitude quant au niveau futur de la volatilité dépend du temps et de ces deux paramètres. En particulier, lorsque σ augmente, l'incertitude augmente et lorsque κ augmente, elle diminue.

Le rôle joué par σ est plutôt évident au regard de l'équation (5.16). En effet, c'est le coefficient placé devant l'incrément du mouvement brownien fractionnaire. Par conséquent, plus il prendra de grandes valeurs, plus l'incertitude sera grande quant à la valeur de la log-volatilité et inversement. C'est ce que confirme l'analyse de l'expression (5.46) : la variance de la volatilité est une fonction croissante de σ.

En ce qui concerne le rôle du paramètre de retour à la moyenne, il est plus ambigu. Lorsque κ augmente, le temps de retour à la moyenne diminue. Cela peut être interprété de deux manières différentes. D'une part, nous pouvons dire qu'il sera plus difficile au processus de la log-volatilité de s'écarter de sa valeur moyenne et donc de prendre de grandes valeurs, ce qui limitera l'amplitude de l'écart à la valeur moyenne de long terme et donc l'incertitude. Mais d'autre part, nous pouvons aussi estimer que le fait d'avoir des retours à la moyenne plus fréquents génère une source supplémentaire d'incertitude puisque, en ramenant plus souvent la volatilité à son niveau moyen, cela va produire des variations de fortes amplitudes. D'un point de vue mathématique, en observant l'expression de la variance de la volatilité (5.46), nous en déduisons qu'une augmentation de κ diminue la variance de la volatilité et donc l'incertitude. Quoi qu'il en soit, nous constatons également qu'il existe une interaction entre ces deux paramètres.

Nous fixons donc ces paramètres à des niveaux raisonnables que nous discutons dans la prochaine sous-section et nous faisons varier tour à tour chacun de ces paramètres afin de rendre la volatilité plus stochastique. Pour mesurer son impact au travers des paramètres σ et κ, nous nous intéressons à la variation de la consommation optimale[45] $\frac{c^*(W,V,t)-c^*(W,\sigma_s,t)}{c^*(W,\sigma_s,t)}$ en fonction de la variation de la volatilité annuelle $\frac{V-\sigma_s}{\sigma_s}$, et celle du poids optimal de l'actif risqué[46] $\frac{w^*(V,t)-w^*(\sigma_s)}{w^*(\sigma_s)}$, toujours en fonction de la variation de la volatilité. La moyenne de long terme de la volatilité ($t \to \infty$) s'écrit $\mathrm{e}^{\alpha+\frac{\sigma^2}{4\kappa}}$. Nous désirons qu'elle soit égale à σ_s, d'où la relation :

$$\alpha = \ln \sigma_s - \frac{\sigma^2}{4\kappa}. \tag{5.47}$$

5.3.3 Choix des paramètres

L'investisseur a le choix entre un actif non risqué, un compte bancaire rémunéré au taux annuel fixe $r = 5\%$, et un actif risqué, une action, par exemple, dont le rendement annuel espéré est $\mu = 15\%$. Nous fixons $\sigma = 1$ et $\kappa = 2$ (autrement dit un retour à la moyenne de 6 mois) et la moyenne à long terme de la volatilité annuelle à $\sigma_s = 30\%$, ce qui entraîne d'après la relation (5.47) : $\alpha = -1.08$ par an. Afin de prendre en compte l'effet de levier, nous affectons une valeur négative à la corrélation entre le mouvement brownien de l'actif risqué et celui de sa volatilité : $\rho = -0.25$. Le but de l'investisseur, moyennement risquophobe ($\gamma = -2$ ce qui donne un coefficient d'aversion relative égal à 3) et légèrement impatient à consommer ($\lambda = 3\%$), est de maximiser l'utilité espérée de sa consommation avec héritage à un horizon de gestion $T = 1$ an. Le tableau 5.1 récapitule tous les paramètres du modèle avec volatilité stochastique. Par ailleurs, nous avons constaté à partir des séries de volatilité réalisée d'actions qu'il est réaliste de choisir $\kappa = 2$ et $\sigma = 1$.

45. $c^*(W, \sigma_s, t)$ est utilisé pour insister sur le fait qu'il s'agit de la consommation optimale dans le cadre du modèle à volatilité constante, notée précédemment $c^*(W, t)$.

46. Ici aussi, $w^*(\sigma_s)$ est employé pour signifier qu'il s'agit du poids optimal de l'actif risqué obtenu par le modèle à volatilité constante.

Actif non risqué	Taux sans risque	$r = 5\%$
Actif risqué	Rendement moyen	$\mu = 15\%$
Log-volatilité	Retour à la moyenne	$1/\kappa = 1/2$
	Moyenne de long terme	$\alpha = -1.08$
	Volatilité	$\sigma = 1$
	Corrélation	$\rho = -0.25$
Critère	Utilité + héritage	$\theta = 1/2$
	Impatience	$\lambda = 3\%$
Fonction d'utilité	CRRA	$\gamma = -2$
		$\beta = \left(\frac{1}{1-\gamma}\right)^{\frac{1-\gamma}{\gamma}}$
		$\eta = 0$

TABLE 5.1 – Paramètres en base annuelle du modèle de choix de portefeuille avec volatilité stochastique.

5.3.4 Résultats

Notre objectif, ici, est de présenter l'impact de la volatilité stochastique sur la consommation optimale, puis sur le poids optimal de l'action dans le portefeuille. Mais auparavant, à titre de référence, nous montrons dans le cas constant la surface de la consommation optimale en fonction du temps et de la volatilité annuelle ainsi que la courbe du poids optimal de l'action dans le portefeuille en fonction de la volatilité.

La figure 5.1 nous montre, la surface du taux de consommation optimale $\frac{c^*(t)}{W}$ en fonction de sa volatilité annuelle dans le cas constant, comprise entre 4% et 100%, et du temps t pour différentes aversion au risque de l'investisseur : sur le graphique du haut, l'investisseur est moyennement risquophobe ($\gamma = -2$), sur celui du milieu il est très risquophobe ($\gamma = -19$) et enfin sur celui du bas, il est presque indifférent au risque ($\gamma = 0.5$). Pour chaque graphique, nous avons assemblé les courbes de la consommation optimale en fonction du temps que nous avons obtenues pour chaque point de la discrétisation de la volatilité. Pour pouvoir analyser le comportement global de l'investisseur, nous avons superposé sur la figure 5.2 les courbes du poids optimal de l'actif risqué en fonction de la volatilité annuelle, dans le cas constant, comprise entre 4 et 100% (le poids est indépendant du temps dans ce cas) pour les différentes valeurs de γ précédemment citées.

En observant la figure 5.1, nous remarquons immédiatement que pour des grandes valeurs de volatilité, les trois surfaces du taux de consommation optimale sont similaires. Par exemple, la consommation, 1 mois après le début de la période de gestion et pour une volatilité annuelle de 60%, vaut 53.86% si l'investisseur est moyennement averse au risque ($\gamma = -2$), 53.98% s'il est très averse au risque $\gamma = -19$) et, enfin, 51.50% s'il est presque neutre face au risque $\gamma = 0.5$). En revanche, le changement de comportement se situe au niveau du poids accordé à l'actif risqué (voir figure 5.2) : l'investisseur moyennement risquophobe investira 9.29% de sa richesse dans l'actif risqué, tandis que l'investisseur très risquophobe seulement 1.39% et l'investisseur presque neutre un peu plus de la moitié : 55.74%.

Pour des petites valeurs de volatilité, la situation est différente : la consommation est

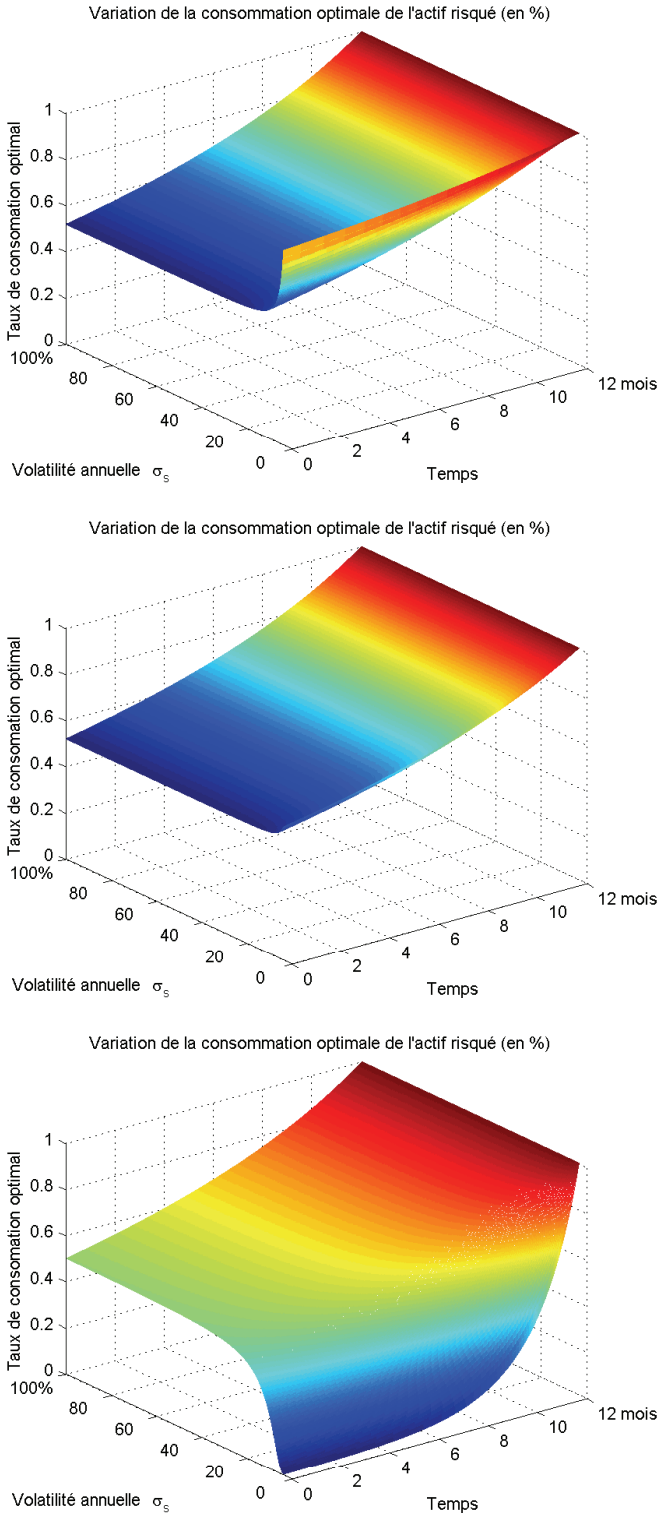

FIGURE 5.1 – Consommation optimale en fonction du temps t et de la volatilité σ_s obtenue dans le cadre du modèle à volatilité constante et à horizon de gestion $T = 1$ an pour différents coefficients d'aversion au risque : 3 (en haut), 20 (au milieu) et 0.5 (en bas).

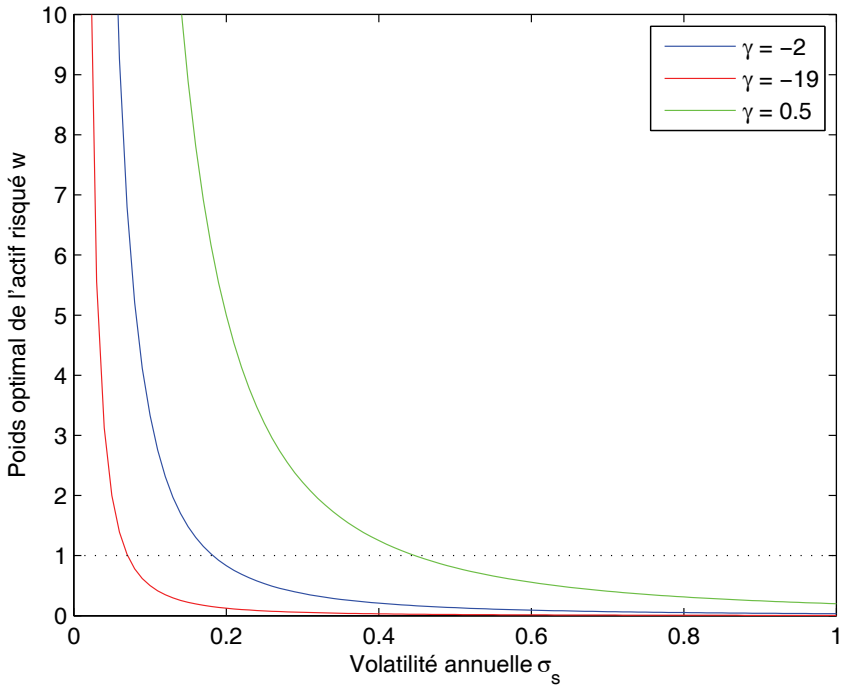

FIGURE 5.2 – Poids de l'actif risqué en fonction de la volatilité annuelle obtenu dans le cadre du modèle à volatilité constante pour différentes aversion au risque de l'investisseur.

significativement affectée par l'aversion au risque de l'investisseur. En effet, pour une aversion au risque moyenne, lorsque que la volatilité tend vers 0, la consommation augmente. Lorsque le risque de l'actif risqué devient très faible, l'investisseur emprunte (pour une volatilité annuelle inférieure à 18.26%, le poids est supérieur à 1), place tout son capital dans l'actif risqué dont le rendement moyen est de 15%, et en profite pour consommer davantage. Au contraire, lorsque qu'il est presque neutre face au risque, sa consommation diminue. Dans ce cas, il emprunte une somme encore plus importante (le poids de l'actif risqué est supérieur à 1 lorsque la volatilité annuelle passe en dessous de 44.72%) qu'il place entièrement dans l'actif risqué et diminue sa consommation pour en tirer le plus grand bénéfice possible.

En somme, l'aversion au risque a essentiellement un impact sur le taux de consommation optimale pour des niveaux faibles de la volatilité : un investisseur moyennement risquophobe consommera davantage et empruntera pour placer la totalité de son capital dans l'actif risqué, tandis qu'un investisseur plutôt neutre face au risque préférera emprunter davantage, placer la totalité de son capital dans l'actif risqué et diminuer sa consommation pour tirer un maximum de profit de la situation. Au-delà d'un certain niveau de volatilité, quel que soit l'aversion au risque (dont le coefficient vaut $1-\gamma$), le taux de consommation, pour un temps donné, atteint un seuil qui augmente au fur et à mesure que l'on se rapproche de l'horizon de gestion. Lorsque σ_s tend vers l'infini, K_2 tend rapidement vers $\frac{\lambda-\gamma r}{1-\gamma}$ et $\frac{c^*(W,t)}{W}$ tend vers $\frac{K_2}{1+(K_2-1)e^{-K_2(T-t)}}$ qui est très peu sensible à la valeur de γ (lorsque $\gamma \to -\infty$, $K_2 \to r$). Par ailleurs, en $t = T$, le taux vaut 1, ce qui signifie que l'investisseur consomme la totalité de sa richesse. Ce résultat peut paraître surprenant puisque le critère de gestion employé tenait

compte du versement d'un capital à l'échéance. En réalité, il n'y a pas d'erreur. L'explication de ce phénomène réside dans la simplification, obtenue au moyen d'une fonction utilité CRRA, de l'expression de la consommation optimale (5.8) : le terme indépendant de W s'annule et $\frac{c^*(W,T)}{W} = F(T) = 1$. L'aversion au risque, dont le coefficient vaut $1 - \gamma$, a naturellement un impact sur la richesse investie dans l'actif risqué. En effet, au regard de l'expression (5.36), lorsque le coefficient d'aversion au risque augmente, le poids diminue.

Impact de la volatilité stochastique au travers du paramètre σ de l'équation de la log-volatilité sur l'allocation optimale

Nous fixons le retour à la moyenne du processus de log-volatilité à 6 mois ($\kappa = 2$) et nous étudions d'une part la variation de la consommation optimale obtenue dans le cadre du modèle à volatilité stochastique par rapport à celle obtenue dans le cadre du modèle à volatilité constante en fonction de la variation de la volatilité annuelle (voir figure 5.3) et d'autre part, selon le même procédé, la variation du poids optimal de l'actif risqué en fonction de la variation de la volatilité annuelle (voir figure 5.4), pour des valeurs du coefficient de diffusion de la log-volatilité comprises entre 0.5 et 2. Les deux figures sont constituées de trois graphiques montrant ces courbes à différentes échéances de l'horizon de gestion : 10 mois, 6 mois et 2 mois (autrement dit $t = 2, 6$ et 10 mois).

La figure 5.3 nous enseigne plusieurs choses. Pour commencer, lorsque la variation de la volatilité annuelle vaut 0, c'est-à-dire lorsque la volatilité de l'actif risqué est égale à sa valeur moyenne de long terme ($V = \sigma_s$), plus σ est grand, autrement dit plus la stochasticité de la volatilité est importante, moins l'investisseur consomme et d'après la figure 5.4, moins il répartit sa richesse dans l'actif risqué. En somme, lorsque l'incertitude augmente, pour un risque égal à sa moyenne de long terme, l'investisseur moyennement risquophobe privilégie davantage le placement sans risque et, puisque le rendement de son portefeuille est affecté, il réduit son train de vie.

A présent, de manière plus générale, nous observons que l'allure des courbes de la variation de la consommation dépend du degré de stochasticité de la volatilité : lorsque $\sigma \in \{0.5, 1\}$, une variation positive (resp. négative) de la volatilité entraîne une variation positive (resp. négative) de la consommation, puis dès qu'elle dépasse un certain seuil, le contraire se produit (c'est le cas pour $\sigma = 2$). Quant au poids optimal de l'actif risqué, l'allure est la même quel que soit le degré de stochasticité : plus V est grand (resp. petit) par rapport à σ_s, plus la proportion d'actif risqué diminue (resp. augmente). A noter que plus le degré de stochasticité est important, plus le poids accordé à l'actif risqué est faible. En résumé, face à une incertitude relativement faible, l'augmentation (resp. la diminution) du risque entraîne une augmentation (resp. diminution) de la consommation couplée à une diminution (resp. augmentation) du poids de l'actif risqué, mais si l'incertitude devient suffisamment importante, alors l'augmentation du risque entraîne une diminution de la consommation et une diminution de la part relative de la richesse investie dans l'actif risqué. Donc l'incertitude engendrée par σ semble être interprété par l'investisseur comme un risque additionnel : lorsque l'incertitude est faible, face à l'augmentation du risque, l'investisseur privilégie sa consommation et le placement sans risque.

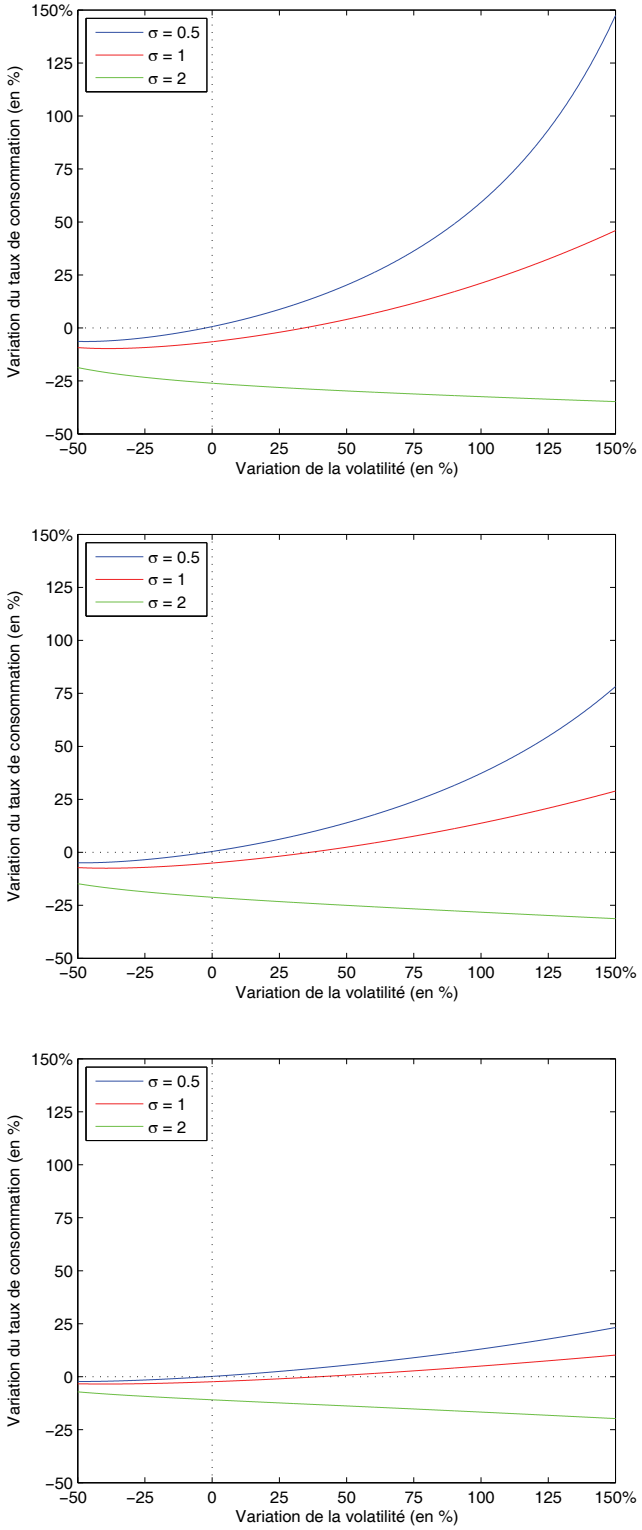

FIGURE 5.3 – Variation de la consommation optimale en fonction de la variation de la volatilité pour un retour à la moyenne du processus de log-volatilité fixé à 6 mois ($\kappa = 2$) et différentes valeurs de σ et du temps t : 2 mois (en haut), 6 mois (au milieu) et 10 mois (en bas).

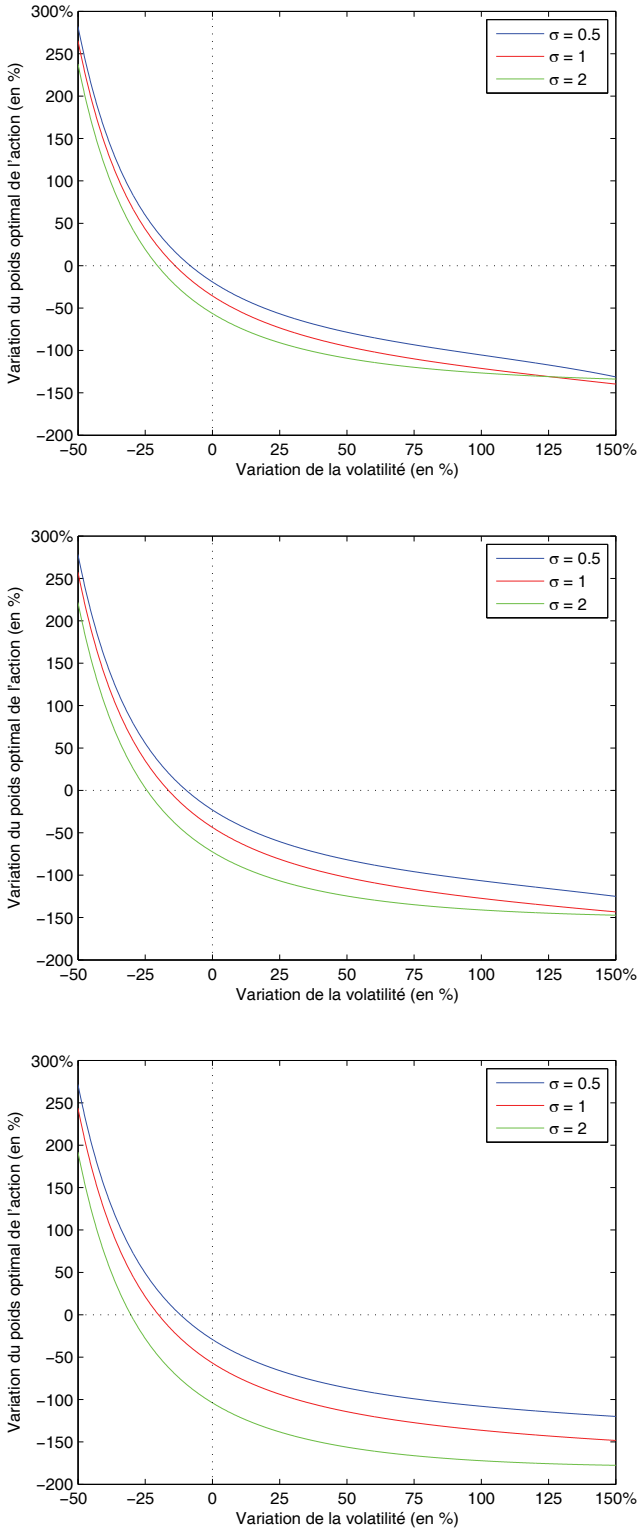

FIGURE 5.4 – Variation du poids optimal de l'actif risqué en fonction de la variation de la volatilité annuelle pour un retour à la moyenne du processus de log-volatilité fixé à 6 mois ($\kappa = 2$) et différentes valeurs de σ et du temps t : 2 mois (en haut), 6 mois (au milieu) et 10 mois (en bas).

Enfin, globalement, l'éloignement de l'échéance de gestion a un impact à la fois sur la variation de la consommation et sur la variation du poids optimal, mais de façon contraire. En effet, lorsque que l'échéance de gestion se rapproche, l'impact de la variation de la volatilité sur la variation de consommation optimale diminue, alors que son impact sur la variation du poids optimal de l'actif risqué augmente.

Impact de la volatilité stochastique au travers du paramètre de retour à la moyenne sur l'allocation optimale

Sur les figures 5.5 et 5.6, nous traçons les mêmes graphiques que sur les figures 5.3 et 5.4 mais inversons les rôles de κ et σ. σ est fixée à 1 et le retour à la moyenne prend différentes valeurs : 3 mois, 6 mois et 1 ans ($\kappa = 4, 2$ et 1).

En ce qui concerne la consommation optimale, les premières constatations semblent être légèrement différentes de celles émises précédemment : quel que soit le temps restant jusqu'à l'horizon de gestion, une diminution (resp. augmentation) de la volatilité annuelle entraîne une diminution (resp. augmentation) de la consommation. En outre, nous remarquons que plus κ est faible, moins la variation de la volatilité a d'influence sur la variation de la consommation. Nous avons établi qu'une diminution de ce paramètre entraînait une augmentation du degré de stochasticité. En particulier, lorsque $\kappa = 1$, c'est-à-dire lorsque la stochasticité est la plus importante, la courbe est quasiment horizontale ce qui laisse à penser que le seuil de stochasticité au-delà duquel le comportement change que nous avons décrit dans l'analyse précédente est proche. Par ailleurs, pour une même variation de la volatilité, l'importance de la variation de la consommation diminue lorsque l'horizon de gestion se rapproche. Ces résultats sont donc cohérents avec les résultats précédents.

Quant à la figure 5.4, les renseignements qu'elle nous fournis sont un peu différents de ceux de la figure 5.4. Premièrement, nous observons qu'à l'approche de l'horizon de gestion, l'impact de la variation de la volatilité annuelle sur la variation de la consommation augmente très légèrement pour $\kappa \in 1, 2$, mais diminue pour $\kappa = 4$.

Autre fait encore plus marquant : la stochasticité de la volatilité générée par κ a une influence négligeable sur le poids de l'actif risqué lorsque la volatilité prend des valeurs plus petites que sa moyenne à long terme. Au contraire, plus la variation est importante, plus l'impact de κ est important. Lorsque $V = \sigma_s$, le poids accordé à l'actif risqué dans le cadre du modèle à volatilité stochastique est inférieur à celui accordé dans le cadre du modèle à volatilité constante. Cela résulte du fait que dans le premier modèle, $\sigma = 1$. Par conséquent, la stochasticité de la volatilité n'est pas négligeable et puisque l'investisseur est moyennement risquophobe, il opte pour plus de sécurité en diminuant sa fraction de richesse investie dans l'actif risqué. Lorsque la variation de la volatilité diminue (reps. augmente), la variation du poids augmente (resp. diminue).

En résumé, l'investisseur, lorsqu'il est confronté à un risque plus élevé, préfère consommer moins et, par l'effet des vases communicants, privilégier son héritage placé en grande majorité dans l'actif non risqué. En outre, lorsque le risque augmente, plus la vitesse de retour à la moyenne est lente, autrement dit plus l'incertitude est grande, moins l'augmentation de

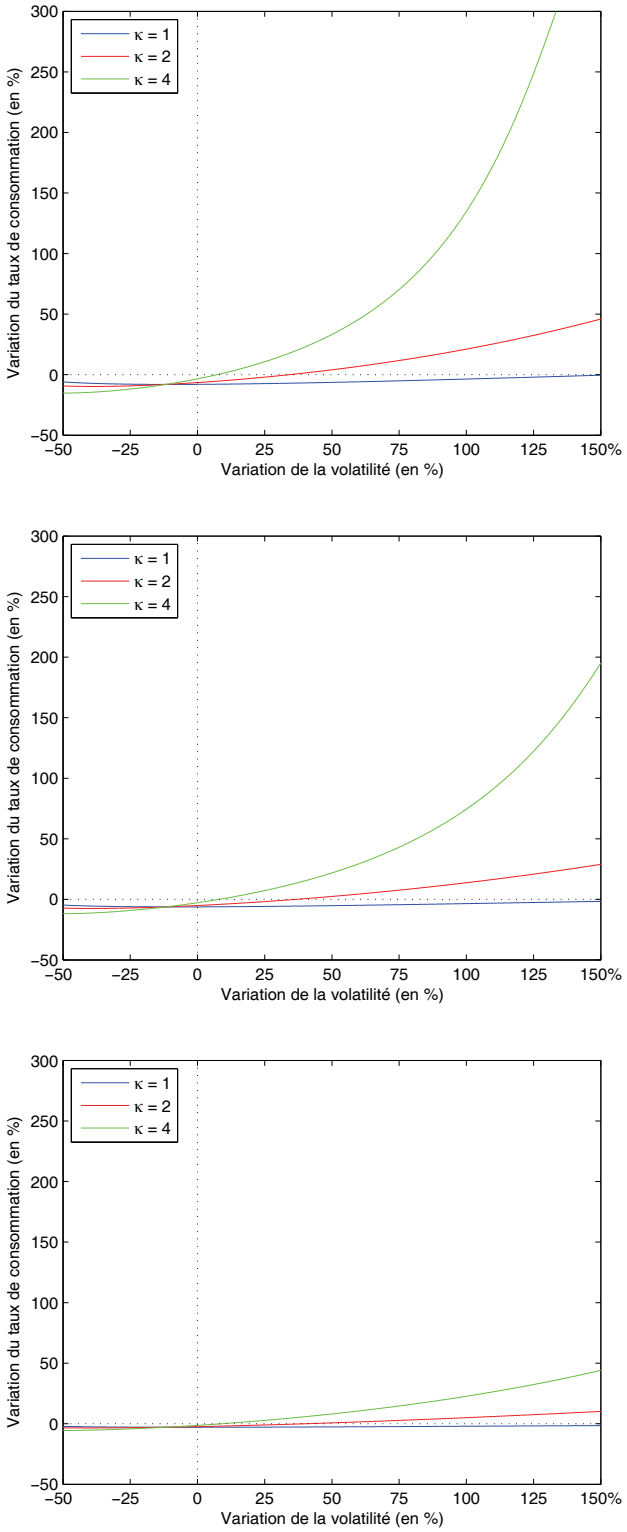

FIGURE 5.5 – Variation de la consommation optimale fonction de la variation de la volatilité pour une volatilité de la log-volatilité fixée ($\sigma = 1$) et différentes valeurs du retour à la moyenne du processus de log-volatilité κ et du temps t : 2 mois (en haut), 6 mois (au milieu) et 10 mois (en bas).

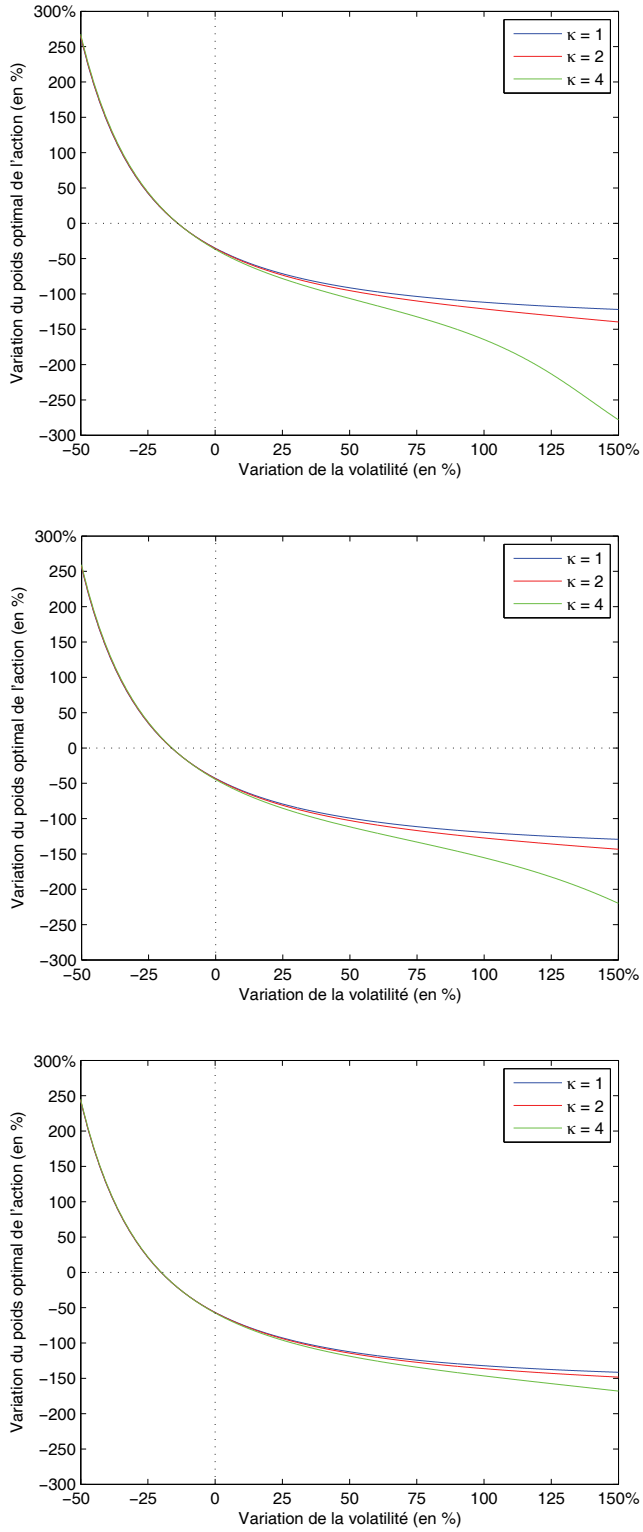

FIGURE 5.6 – Variation du poids optimal de l'actif risqué en fonction de la variation de la volatilité pour une volatilité de la log-volatilité fixée ($\sigma = 1$) et différentes valeurs du retour à la moyenne du processus de log-volatilité κ et du temps t : 2 mois (en haut), 6 mois (au milieu) et 10 mois (en bas).

la consommation de l'investisseur est importante et moins il diminue la proportion d'actif risqué. Ce dernier point peut paraître surprenant puisqu'il est en contradiction avec l'analyse du degré de stochasticité généré par σ. Il existe donc une interaction entre les paramètres σ et κ concernant le degré de stochasticité. Si le rôle de σ est clairement établi, celui de κ est plus ambigü. La contradiction en est probablement la conséquence.

A ce stade de notre analyse, nous insistons sur l'importance du rôle du paramètre ρ qui contrôle la corrélation entre le rendement de l'actif risqué et sa volatilité dans l'allocation optimale : $\mathrm{Corr}(dB(t), dZ(t)) = \rho$, où $dB(t)$ est le mouvement brownien de la dynamique des rendements et $dZ(t)$ celui de la volatilité. Or, cela signifie que si la volatilité est prévisible au temps t, le rendement de l'actif risqué l'est également, du moins en partie selon la valeur de ρ. Cette modélisation, pourtant utilisée dans de nombreux modèles tels que ceux de Chacko et Viceira (2005) et Liu (2007) ou encore le modèle de Heston (1993) dans un autre domaine, conduit à une forme d'arbitrage. Dans nos simulations, nous avons employé $\rho = -0.25$, c'est à dire qu'une augmentation de la volatilité est liée à un rendement négatif. Si à présent nous posons $\rho = 0.25$, alors une augmentation du risque devrait conduire à un rendement positif. Par conséquent, l'augmentation du risque devrait ne pas entraîner une aussi grande diminution du poids dans l'actif risqué et dans certains cas peut-être même une augmentation en dépit du caractère moyennement risquophobe de l'investisseur. C'est ce qu'illustre la figure 5.7 qui correspond au premier graphique de la figure 5.6 avec $\rho = 0.25$ au lieu de $\rho = -0.25$. Dans ce cas, la couverture intertemporelle est positive et s'accroît significativement lorsque le temps de retour à la moyenne diminue (voir figure 5.8, graphique du haut). Inversement, lorsque ρ est négatif, la couverture intertemporelle est négative. Nous observons même, pour le jeu de paramètres choisi, une symétrie par rapport à l'axe des abscisses entre les deux couvertures.

Etude du taux de croissance

Pour achever l'étude du modèle à volatilité stochastique, nous analysons le taux de croissance du portefeuille

$$g(t) = \mathrm{E}\left[\frac{dW(t)}{W(t)}\right]\frac{1}{dt}. \tag{5.48}$$

En remplaçant $dW(t)$ et $w(t)$ par leur expression respectives dans (5.20) et (5.39), nous obtenons

$$g(t) = r + \frac{(\mu - r)^2}{(1-\gamma)}\mathrm{E}\left[\frac{1}{V(t)^2}\right] + \mathrm{E}\left[\rho\sigma(\mu - r)\frac{F_V(V(t),t)}{F(V(t),t)} - \frac{1}{F(V(t),t)}\right]. \tag{5.49}$$

Or, $1/V^2$ peut s'écrire $\mathrm{e}^{-2\ln V}$. Puisque $\ln V$ est une loi normale dont les paramètres sont donnés par les équations (5.43) et (5.44), alors $\mathrm{e}^{-2\ln V}$ est une loi log-normale. Afin que la moyenne ne dépende pas du temps, nous posons $\ln V(0) = \alpha$. Nous en déduisons la forme définitive du taux de croissance du portefeuille :

$$g(t) = r + \frac{(\mu - r)^2}{(1-\gamma)}\mathrm{e}^{-2\alpha + \frac{\sigma^2}{\kappa}} + \rho\sigma(\mu - r)\mathrm{E}\left[\frac{F_V(V(t),t)}{F(V(t),t)}\right] - \mathrm{E}\left[\frac{1}{F(V(t),t)}\right]. \tag{5.50}$$

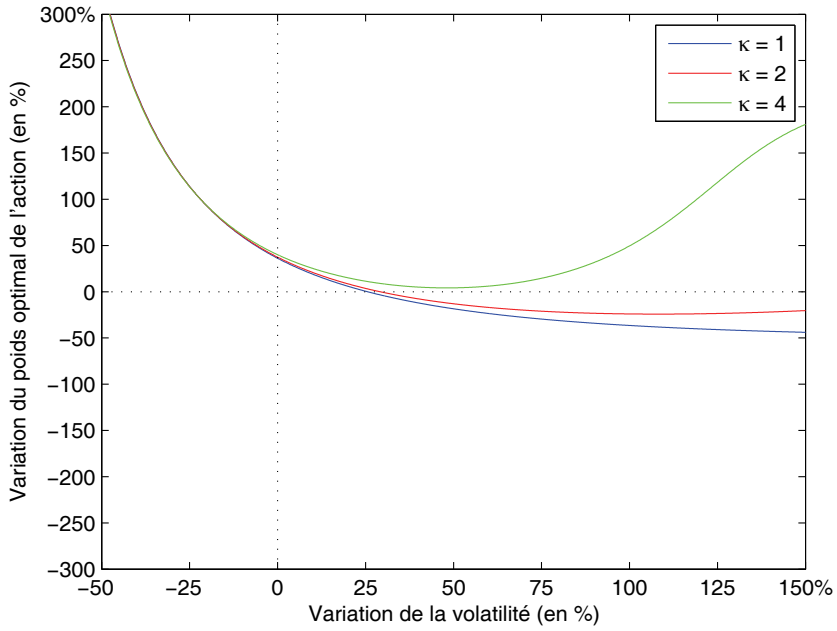

FIGURE 5.7 – Variation du poids optimal de l'actif risqué en fonction de la variation de la volatilité pour $\rho = 0.25$, $\sigma = 1$ et différentes valeurs du retour à la moyenne du processus de log-volatilité κ au temps $t = 2$ mois.

Nous remarquons que $g(T)$ se calcule explicitement étant donné que $F(V,T) = 1$ pour tout V et $F_V(V,T) = 0$ pour tout V :

$$g(T) = r + \frac{(\mu - r)^2}{(1-\gamma)} \mathrm{e}^{-2\alpha + \frac{\sigma^2}{\kappa}} - 1. \tag{5.51}$$

Nous traçons sur la figure 5.9 le taux de richesse du portefeuille en fonction du temps pour différentes valeurs du rendement de l'actif risqué.

En fonction du rendement de l'actif risqué, le taux de croissance connaît différentes tendances : lorsque $\mu = 15\%$ et 30%, il décroît sur toute la période de gestion, avec une accélération de la décroissance pour $\mu = 15\%$ dans quatre derniers mois. En revanche, lorsque $\mu = 45\%$, il connaît deux phases : une décroissance qui s'étend sur les dix premiers mois, puis une augmentation sur les deux derniers mois.

Par ailleurs, nous observons que le taux de richesse est constamment négatif pour les deux plus petits rendements, autrement dit la consommation est supérieure à l'accroissement de richesse du portefeuille, mais cet écart diminue d'autant plus que le rendement de l'actif risqué est élevé. En somme, dès le début de la période de gestion, l'investisseur consomme sans prêter particulièrement attention à son capital. Cette observation est consistante avec le fait que pour une fonction d'utilité CRRA, le gestionnaire ne laisse pas de capital à l'échéance. Lorsque le rendement est suffisamment grand, par exemple $\mu = 45\%$, la consommation devient inférieure à l'accroissement de richesse sur les trois premiers mois et l'investisseur épargne.

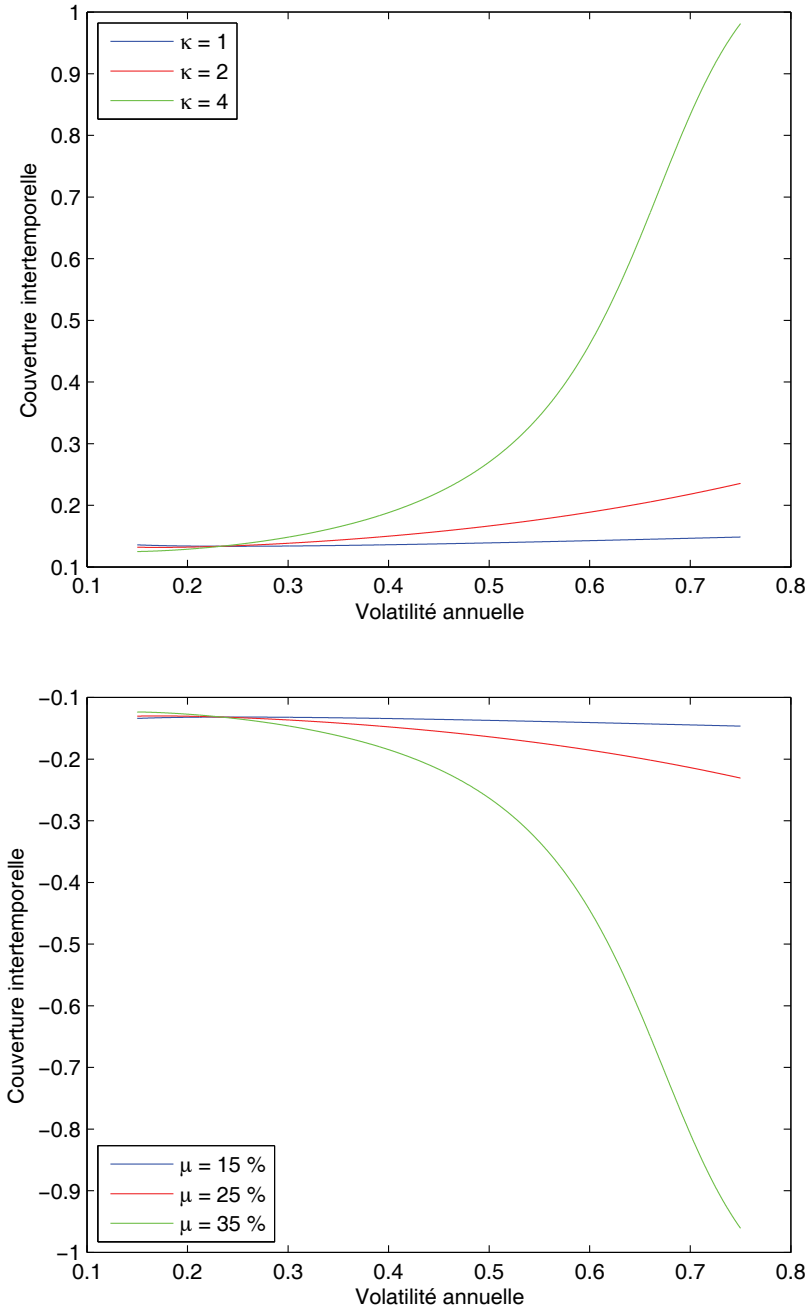

FIGURE 5.8 – Couverture intertemporelle pour $\sigma = 1$ et différentes valeurs du retour à la moyenne du processus de log-volatilité κ au temps $t = 2$ mois. Sur le graphique du haut $\rho = 0.25$ et sur celui du bas $\rho = -0.25$.

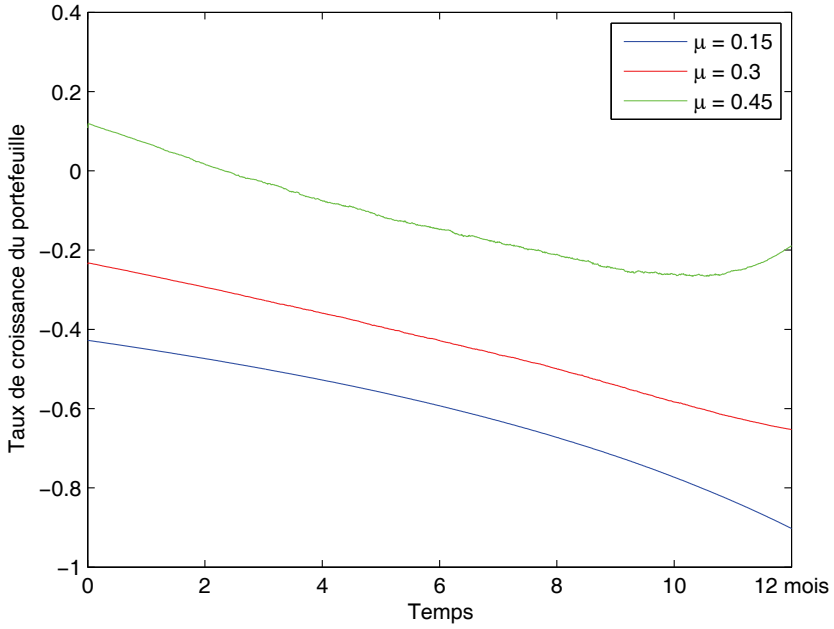

FIGURE 5.9 – Taux de croissance du portefeuille en fonction du temps pour différentes valeurs du rendement de l'actif risqué : 15, 30 et 45%.

Ensuite, l'inverse se produit jusqu'à la fin de la période de gestion.

Conclusion

Dans ce chapitre, nous avons proposé un modèle de choix de portefeuille en temps continu avec une volatilité stochastique. Nous utilisons un processus de retour à la moyenne de type Ornstein-Uhlenbeck pour modéliser non pas la volatilité, mais la log-volatilité en raison des valeurs négatives que peut prendre un tel processus. L'intérêt de l'utilisation de ce processus est double. Premièrement, le retour à la moyenne est une des caractéristiques bien connues des séries de volatilité des rendements. Deuxièmement, ce processus peut, au moyen d'une modification, devenir un processus à mémoire longue désigné sous le nom de processus d'Ornstein-Uhlenbeck fractionnaire et ainsi donner naissance à un modèle de choix de portefeuille en temps continu avec volatilité stochastique à mémoire longue.

A partir de notre modèle avec volatilité stochastique, nous avons dérivé les résultats analytiques de trois critères de gestion fréquemment rencontrés dans la littérature : la maximisation de l'utilité espérée de la consommation avec héritage à l'échéance T, le même critère mais sans héritage et la maximisation de l'utilité espérée de la richesse finale, ainsi qu'une classe générale de fonction d'utilité : la classe HARA. En utilisant l'approche par programmation dynamique et l'équation d'Hamilton-Jacobi-Bellman, nous avons obtenu une équation aux dérivées partielles que nous avons ensuite simplifiée au maximum, à défaut de pouvoir la résoudre analytiquement. Enfin, nous avons proposé une méthode d'approximation numérique

basée sur la méthode des différences finies ainsi qu'un algorithme de résolution du problème étudié.

De cette façon, nous avons été capable d'étudier l'impact du caractère stochastique de la volatilité sur l'allocation optimale. Pour illustrer cela, nous avons considéré un investisseur moyennement averse au risque souhaitant faire fructifier son patrimoine pendant un an tout en consommant et laisser un capital à l'échéance. L'amplitude de l'écart relatif entre la volatilité stochastique et la volatilité constante, que nous avons appelé degré de stochasticité, est contrôlée par les paramètres σ (le coefficient de diffusion de la log-volatilité) et κ (le retour à la moyenne du processus de la log-volatilité). Nous avons montré que

- Lorsque la volatilité stochastique s'écarte de la moyenne de long terme de la volatilité en prenant des valeurs positives, l'importance du degré de stochasticité à la fois sur la consommation optimale et sur le poids optimal de l'actif risqué, qu'elle soit contrôlée par σ ou κ, est très nette. En revanche, lorsqu'elle prend des valeurs négatives, l'impact est moindre.
- L'incertitude générée par σ sur la valeur de la volatilité agit comme une source de risque supplémentaire : en-dessous d'un certain seuil,
  - lorsque la volatilité prend des valeurs supérieures à sa moyenne de long terme, l'augmentation de variation de la consommation est d'autant plus faible et celle du poids optimal de l'actif risqué d'autant plus forte que le degré de stochasticité est important.
  - lorsque la volatilité prend des valeurs inférieures à sa moyenne de long terme, la diminution du risque de l'action est en partie compensée par le risque additionnel lié à l'incertitude. C'est pourquoi la consommation diminue moins vite lorsque la volatilité de l'action diminue. En outre, l'augmentation de la variation du poids de l'actif risqué est moins forte.

  Puis au-dessus de ce seuil, la variation du taux de consommation est une fonction décroissante de la variation de la volatilité.
- Au contraire du rôle de σ, celui de κ est moins évident. Il semblerait que lorsque κ augmente, l'incertitude décroît. En outre, contrairement à l'incertitude engendrée par σ, celle formée par κ ne se traduirait pas en terme de risque additionnel. En effet,
  - bien que l'incertitude générée par κ a la même influence sur la variation de la consommation que celle induite par σ,
  - c'est le contraire qui se produit pour la variation du poids optimal de l'actif risqué.

De plus, d'une manière générale, l'impact de la stochasticité de la volatilité au travers des paramètres σ et κ sur la variation de consommation s'atténue au fur et à mesure que l'on se rapproche de l'horizon de gestion, alors que l'impact sur la variation du poids optimal de l'actif risqué augmente. En se rapprochant de l'échéance, l'investisseur moyennement risquophobe préfère investir de moins en moins dans l'actif sans risque pour atteindre son objectif de gestion et puisque le rendement attendu est moindre, il consent à réduire sa consommation.

Toutefois, nous nous devons d'attirer l'attention sur un point important lié à la modélisation de l'effet de levier ayant une grande influence sur les résultats. Dans de nombreux modèles comme dans le nôtre, l'effet de levier est pris en compte au moyen de la corrélation ρ entre ll'accroissement du mouvement brownien des rendements $dB(t)$ et celui de la volatilité $dZ(t)$. Donc connaître la volatilité au temps t permet de déduire, en partie, le rendement au temps t.

Annexes

5.A Rappels sur le mouvement brownien

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité avec la mesure de probabilité $\mathbb{P}$. Par la suite, nous exprimons l'information disponible au temps $t \in [0, T]$ par la filtration $\mathcal{F}_t$ sur l'espace de probabilité, définie comme suit.

Définition 12. *Une filtration sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ est une famille croissante $(\mathcal{F}_t)_{0 \leq t \leq T}$ de sous tribus de $\mathcal{A}$*

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_T = \mathcal{A}, \quad \mathcal{F}_s \subseteq \mathcal{F}_t \text{ pour } s \leq t. \tag{5.A.1}$$

Le mouvement brownien est le nom donné par Robert Brown en 1828 pour décrire le mouvement irrégulier des particules de pollen en suspension dans l'eau. Ce mouvement aléatoire a, depuis, été utilisé dans de nombreux domaines, et en particulier en finance où Bachelier (1900) l'a employé pour modéliser le cours des actions. Mais il faut attendre une vingtaine d'années pour assister aux premières études "mathématiques" [47] de Wiener (1923). Puis, Lévy (1948) approfondit les travaux de Wiener. Le mouvement brownien est défini de la manière suivante :

Définition 13. *Un mouvement brownien est un processus continu, adapté à la filtration $\mathcal{F}_t$, noté $B = \{B(t), \mathcal{F}_t; 0 \leq t < \infty\}$ et défini sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, ayant les propriétés suivantes :*

1. $B(0) = 0 \qquad \mathbb{P} - p.s.,$
2. *Pour tout $0 \leq s < t$, les accroissements $B(t) - B(s)$ sont indépendants de $\mathcal{F}_s$,*
3. *Pour tout $0 \leq s < t$, les accroissements $B(t) - B(s)$ sont gaussiens de moyenne nulle et de variance $t - s$.*

On dit que $B(t)$ est un processus à accroissements indépendants et stationnaires.

Pour la dérivation des résultats analytiques des modèles de choix de portefeuille, nous faisons appel au lemme d'Itō.

Définition 14. *On appelle processus d'Itō, un processus $X = \{X(t), 0 \leq t \leq T\}$ à valeurs dans $\mathbb{R}$ tel que*

$$dX(t) = \mu(t)dt + \sigma(t)dB(t), \tag{5.A.2}$$

avec

- *$\mu = \{\mu(t), 0 \leq t \leq T\}$ et $\sigma = \{\sigma(t), 0 \leq t \leq T\}$ des processus adaptés à la filtration $\mathcal{F}_t$,*
- *$P\left[\int_0^T |\mu(s)|ds < \infty\right] = 1$,*

47. Preuve de l'existence du mouvement brownien

- $P\left[\int_0^T |\sigma(s)|^2 ds < \infty\right] = 1.$

Théorème 1 (Lemme d'Itō)**.** Soit X un processus d'Itō et $f : \mathbb{R} \times [0, \infty[\longrightarrow \mathbb{R}$ une fonction dont les dérivées partielles premières et secondes existent et sont continues. Alors, pour tout $0 \leq t \leq T$,

$$f(X(t),t) = f(X(0),0) + \int_0^t \frac{\partial f}{\partial s}(X(s),s)ds + \int_0^t \frac{\partial f}{\partial x}(X(s),s)dX(s) + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(X(s),s)\sigma^2(s)ds. \quad (5.A.3)$$

5.B Démonstration de la section 5.2.2

Nous rappelons le problème de maximisation à résoudre :

$$\begin{cases} 0 = \max\limits_{c(t),w(t)} \left\{ \theta \mathrm{e}^{-\lambda t} u(c(t)) J_t + \left[(\mu - r)w(t)W + rW - c(t)\right] J_W + \frac{\sigma^2}{2} V^2 J_{VV} \right. \\ \qquad \left. + V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] J_V + \frac{1}{2} w(t)^2 W^2 V^2 J_{WW} + \rho\sigma w(t) W V^2 J_{VW} \right\}, \quad (5.B.1) \\ J(W,V,T) = (1-\theta)\mathrm{e}^{-\lambda T}\frac{W^\gamma}{\gamma}, \qquad \forall\, W, V. \quad (5.B.2) \end{cases}$$

Les conditions du premier ordre nous fournissent la consommation optimale et le poids optimal de l'actif risqué :

$$\begin{cases} c^* = u'^{-1}\left(\frac{\mathrm{e}^{\lambda t}}{\theta} J_W\right), & (5.B.3) \\ w^* = -\frac{(\mu - r)}{W(t)V(t)^2}\frac{J_W}{J_{WW}} - \frac{\rho\sigma}{W(t)}\frac{J_{VW}}{J_{WW}}. & (5.B.4) \end{cases}$$

L'expression du poids optimal de l'actif risqué se démarque du modèle à volatilité constante dans la mesure où premièrement la volatilité constante σ_s a été remplacée par $V(t)$ et, deuxièmement, la dérivée partielle de la fonction valeur J par rapport à V et à W a donné naissance à un second terme faisant apparaître la corrélation ρ entre le brownien du rendement et celui de la volatilité.

Les conditions du second ordre qui nous impose la concavité de la fonction J est admise. Par conséquent, en remplaçant $w(t)$ dans l'équation (5.B.1) par son expression dans l'équation (5.B.4), on obtient l'équation aux dérivées partielles générale du problème :

$$0 = \theta \mathrm{e}^{-\lambda t} u(c^*) + J_t + (rW - c^*)J_W + V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] J_V + \frac{\sigma^2}{2} V^2 J_{VV} - \frac{(\mu - r)^2}{2V^2}\frac{J_W^2}{J_{WW}} - \frac{\sigma^2}{2}\rho^2 V^2 \frac{J_{VW}^2}{J_{WW}} - (\mu - r)\rho\sigma\frac{J_W J_{VW}}{J_{WW}}, \quad (5.B.5)$$

avec la condition limite $J(W,V,T) = (1-\theta)\mathrm{e}^{-\lambda T}\frac{W^\gamma}{\gamma}$ pour tout W et V. Pour retomber sur l'équation aux dérivées partielles obtenue à partir du modèle à volatilité constante, il faut remplacer V par la volatilité constante σ_s et supprimer tous les termes J_V, J_{VV} et J_{VW}. On remarque donc que dans le cas d'un modèle à volatilité stochastique l'équation aux dérivées partielles du problème s'est considérablement compliquée, et contient désormais la dérive de la volatilité, son coefficient de diffusion, et pour les termes croisés, les coefficients de corrélation ρ entre le rendement de l'actif risqué et la volatilité ainsi que le coefficient de diffusion de la log-volatilité σ.

Comme dans le cas de la volatilité constante, nous nous intéressons maintenant à la fonction d'utilité HARA :

$$u(c) = \frac{1-\gamma}{\gamma}\left(\frac{\beta c}{1-\gamma} + \eta\right)^\gamma \tag{5.B.6}$$

A partir de l'équation (5.B.3), nous en déduisons

$$c^* = \frac{1-\gamma}{\beta}\left[\left(\frac{\mathrm{e}^{\lambda t}}{\theta}\frac{J_W}{\beta}\right)^{\frac{1}{\gamma-1}} - \eta\right], \tag{5.B.7}$$

et

$$u(c^*) = \frac{1-\gamma}{\gamma}\left(\frac{e^{\lambda t}}{\theta}\frac{J_W}{\beta}\right)^{\frac{\gamma}{\gamma-1}}. \tag{5.B.8}$$

En reportant les expressions de c^* et $u(c^*)$ respectivement de (5.B.7) et (5.B.8) dans l'équation aux dérivées partielles générale (5.B.5) et en factorisant les termes en $(J_W)^{\frac{\gamma}{1-\gamma}}$, il vient

$$\begin{aligned} 0 = \frac{(1-\gamma)^2}{\gamma}\beta^{\frac{\gamma}{1-\gamma}}\theta\mathrm{e}^{-\lambda t}\left(\frac{\mathrm{e}^{\lambda t}}{\theta}J_W\right)^{\frac{\gamma}{\gamma-1}} + J_t + \left[rW + (1-\gamma)\frac{\eta}{\beta}\right]J_W \\ + V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right]J_V + \frac{\sigma^2}{2}V^2 J_{VV} \\ - \frac{(\mu-r)^2}{2V^2}\frac{J_W^2}{J_{WW}} - \frac{\sigma^2}{2}\rho^2 V^2\frac{J_{VW}^2}{J_{WW}} - (\mu-r)\rho\sigma\frac{J_W J_{VW}}{J_{WW}}. \end{aligned} \tag{5.B.9}$$

Cette équation aux dérivées partielles est compliquée. Nous allons, en procédant étape par étape, simplifier son écriture.

Commençons par poser

$$J(W,V,t) = \theta\mathrm{e}^{-\lambda t}H(W,V,t), \tag{5.B.10}$$

afin de faire disparaître les termes en $\theta\mathrm{e}^{-\lambda t}$ de l'équation aux dérivées partielles. Nous soulignons le fait qu'ici H n'est pas la fonction héritage. Il vient

$$\begin{aligned} 0 = \frac{(1-\gamma)^2}{\gamma}\beta^{\frac{\gamma}{1-\gamma}}(H_W)^{\frac{\gamma}{\gamma-1}} + H_t - \lambda H + \left(rW + (1-\gamma)\frac{\eta}{\beta}\right)H_W \\ + V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right]H_V + \frac{\sigma^2}{2}V^2 H_{VV} \\ - \frac{(\mu-r)^2}{2V^2}\frac{H_W^2}{H_{WW}} - \frac{\sigma^2}{2}\rho^2 V^2\frac{H_{VW}^2}{H_{WW}} - (\mu-r)\rho\sigma\frac{H_W H_{VW}}{H_{WW}}. \end{aligned} \tag{5.B.11}$$

Si $\gamma > 0$ alors, d'après l'équation (5.24), $J(W,V,t) > 0$. Inversement, lorsque $\gamma < 0$, alors $J(W,V,t) < 0$. Par conséquent, en s'inspirant de la solution théorique du modèle à volatilité constante pour séparer la variable W de V, posons

$$H(W,V,t) = \left[W + \frac{(1-\gamma)\eta}{\beta r}\left(1 - \mathrm{e}^{-r(T-t)}\right)\right]^{\gamma} \mathrm{sgn}(\gamma) G(V,t), \tag{5.B.12}$$

où $G(V,t)$ est une fonction positive. Cela entraîne, pour tout $\gamma < 1, \gamma \neq 0$, la simplification

$$\begin{aligned} 0 = (1-\gamma) K G^{\frac{\gamma}{\gamma-1}} - (\lambda - \gamma\nu(V)) G + G_t + \frac{\sigma^2}{2} V^2 G_{VV} \\ + \left\{ V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] + \frac{\gamma}{1-\gamma}(\mu - r)\rho\sigma \right\} G_V + \frac{\gamma}{1-\gamma}\frac{\sigma^2}{2}\rho^2 V^2 \frac{G_V^2}{G}, \end{aligned} \tag{5.B.13}$$

avec $K := (1-\gamma)\beta^{\frac{\gamma}{1-\gamma}}\left(\frac{1}{|\gamma|}\right)^{\frac{1}{1-\gamma}}$ et $\nu(V) := r + \frac{(\mu-r)^2}{2(1-\gamma)V^2}$. Enfin, posons

$$G(V,t) = (KF(V,t))^{1-\gamma}, \tag{5.B.14}$$

pour faire disparaître la puissance du terme $G^{\frac{\gamma}{\gamma-1}}$:

$$\begin{aligned} 0 = 1 - K_2(V) F + F_t + \left\{ V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] + \frac{\gamma}{1-\gamma}(\mu - r)\rho\sigma \right\} F_V \\ + \frac{\sigma^2}{2} V^2 F_{VV} - \gamma\frac{\sigma^2}{2}(1-\rho^2) V^2 \frac{F_V^2}{F}, \end{aligned} \tag{5.B.15}$$

avec $K_2(V) := \frac{\lambda - \gamma\nu(V)}{1-\gamma}$. La condition limite s'écrit maintenant

$$F(V,T) = \frac{\beta^{-\frac{\gamma}{1-\gamma}}}{1-\gamma}\left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\gamma}}. \tag{5.B.16}$$

À ce stade là, l'équation aux dérivées partielles est simplifiée au maximum. Au final, nous avons posé

$$J(W,V,t) = \mathrm{sgn}(\gamma)\theta \mathrm{e}^{-\lambda t}\left[W + K_1 \frac{1 - \mathrm{e}^{-r(T-t)}}{r}\right]^{\gamma} K^{1-\gamma} F(V,t)^{1-\gamma}, \tag{5.B.17}$$

avec $K_1 := \frac{(1-\gamma)\eta}{\beta}$ et où $F \geq 0$ pour tout V et t. En remplaçant J_W dans les équations (5.B.7) et (5.B.8) à partir de son expression dans l'équation (5.B.17), la consommation optimale et la proportion optimale d'actif risqué s'écrivent

$$\begin{cases} c^*(W,V,t) = \dfrac{W}{F(V,t)} + K_1\left(\dfrac{1 - \mathrm{e}^{-r(T-t)}}{r}\dfrac{1}{F(V,t)} - 1\right), & (5.B.18) \\ w^*(W,V,t) = \dfrac{\mu - r}{(1-\gamma)V^2} + K_1 \dfrac{\mu - r}{(1-\gamma)V^2}\dfrac{1 - \mathrm{e}^{-r(T-t)}}{r}\dfrac{1}{W} \\ \qquad\qquad + \rho\sigma\dfrac{F_V(V,t)}{F(V,t)}\left(1 + K_1\dfrac{1 - \mathrm{e}^{-r(T-t)}}{r}\dfrac{1}{W}\right), & (5.B.19) \end{cases}$$

Dans le cadre d'un modèle à volatilité constante, les dérivées partielles de F par rapport à V dans l'équation aux dérivées partielles (5.B.15) s'annulent et on obtient alors une simple équation différentielle ordinaire

$$0 = 1 - K_2 F + F_t, \tag{5.B.20}$$

avec la condition limite $F(T) = \frac{\beta^{-\frac{\gamma}{1-\gamma}}}{1-\gamma}\left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\gamma}}$, et dont la solution analytique est donnée par

$$F(t) = \frac{1}{K_2} + \left[\frac{\beta^{-\frac{\gamma}{1-\gamma}}}{1-\gamma}\left(\frac{1-\theta}{\theta}\right)^{\frac{1}{1-\gamma}} - \frac{1}{K_2}\right] \mathrm{e}^{-K_2(T-t)}. \tag{5.B.21}$$

□

5.C Démonstration de la section 5.2.3

Nous considérons le cas d'un investisseur ayant pour objectif de maximiser l'utilité espérée de sa richesse finale au temps T. Dans ce cas de figure, il n'y a pas de consommation, et il faut alors reécrire l'équation différentielle stochastique de la richesse

$$dW(t) = [(\mu - r)w(t)W(t) + rW(t)]\,dt + w(t)W(t)V(t)dB(t). \tag{5.C.1}$$

Le système devient

$$\begin{cases} dW(t) = [(\mu - r)w(t)W(t) + rW(t)]\,dt + w(t)W(t)V(t)dB(t), & (5.C.2) \\ dV(t) = \left(\dfrac{\sigma^2}{2} + \kappa(\alpha - \ln V(t))\right) V(t)dt + \sigma V(t)dZ(t), & (5.C.3) \end{cases}$$

En posant $\theta = 0$ et $\lambda = 0$ (puisqu'il n'y a plus de consommation), le critère de gestion devient

$$\max \mathrm{E}\left[H(W(T), T)\right], \tag{5.C.4}$$

où la fonction héritage $H(W(T), T) = \frac{W(T)^\gamma}{\gamma}$.

De même, la fonction d'utilité indirecte et l'équation de Hamilton-Jacobi-Belman deviennent respectivement

$$J(W, V, t) = \max_{w(t)} \mathrm{E}_t\left[\frac{W(T)^\gamma}{\gamma}\right], \tag{5.C.5}$$

et

$$0 = \max_{w}\left[\mathcal{L}J(W, t) + J_t\right], \tag{5.C.6}$$

où l'opérateur différentiel générateur s'écrit

$$\begin{aligned} \mathcal{L}J(W, V, t) = {} & [(\mu - r)w(t)W + rW]\,J_W + V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] J_V \\ & + \frac{1}{2}w(t)^2 W^2 V^2 J_{WW} + \frac{\sigma^2}{2}V^2 J_{VV} + \rho\sigma w(t) W V^2 J_{VW}, \end{aligned} \tag{5.C.7}$$

La condition du premier ordre nous fournit le poids optimal de l'actif risqué :

$$w^* = -\frac{(\mu - r)}{W(t)V(t)^2}\frac{J_W}{J_{WW}} - \frac{\rho\sigma}{W(t)}\frac{J_{VW}}{J_{WW}}, \tag{5.C.8}$$

tandis que la condition du second ordre qui impose la concavité de la fonction J est admise.

En remplaçant $w(t)$ dans l'équation (5.C.7) par son expression dans (5.C.8), l'équation aux dérivées partielles générale du problème s'écrit

$$\begin{aligned} 0 = J_t + rWJ_W + V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right]J_V + \frac{\sigma^2}{2}V^2 J_{VV} \\ - \frac{(\mu - r)^2}{2V^2}\frac{J_W^2}{J_{WW}} - \frac{\sigma^2}{2}\rho^2 V^2\frac{J_{VW}^2}{J_{WW}} - (\mu - r)\rho\sigma\frac{J_W J_{VW}}{J_{WW}}, \end{aligned} \tag{5.C.9}$$

avec la condition limite $J(W, V, T) = \frac{W^\gamma}{\gamma}$ pour tout W, V. En remarquant que, en employant la même méthode que pour l'annexe précédente, nous pouvons écrire

$$J(W, V, t) = \mathrm{sgn}(\gamma)W^\gamma F(V, t)^{1-\gamma}, \tag{5.C.10}$$

avec $F \geq 0$. Nous obtenons l'expression du poids optimal de l'actif risqué

$$w^*(V, t) = \frac{\mu - r}{(1 - \gamma)V(t)^2} + \rho\sigma\frac{F_V(V, t)}{F(V, t)}. \tag{5.C.11}$$

et l'équation aux dérivées partielles dont F est la solution

$$\begin{aligned} 0 = \frac{\gamma\nu(V)}{1 - \gamma}F + F_t + \left\{V\left[\frac{\sigma^2}{2} + \kappa(\alpha - \log V)\right] + \frac{\gamma}{1 - \gamma}(\mu - r)\rho\sigma\right\}F_V \\ + \frac{\sigma^2}{2}V^2 F_{VV} - \gamma\frac{\sigma^2}{2}V^2(1 - \rho^2)\frac{F_V^2}{F}, \end{aligned} \tag{5.C.12}$$

avec la condition limite $F(V, T) = |\gamma|^{-\frac{1}{1-\gamma}}$ pour tout V, et où $\nu(V) := r + \frac{(\mu - r)^2}{2(1-\gamma)V^2}$. □

Chapitre 6

Modèle de choix de portefeuille en temps continu avec volatilité stochastique à mémoire longue

Dans le chapitre précédent, nous avons proposé un modèle de choix de portefeuille avec une volatilité stochastique. Or, à nos yeux, ce modèle n'est pas complet puisqu'il ne permet pas de tenir compte d'une caractéristique très importante de la volatilité : la mémoire longue. Dans notre esprit, c'était un modèle intermédiaire qui nous a permis d'étudier l'impact du caractère stochastique de la volatilité sur l'allocation optimale et qui nous sert de modèle de référence dans ce chapitre. En effet, nous améliorons celui-ci pour lui permettre de prendre en compte la mémoire longue de la volatilité d'actions.

La grande nouveauté de ce modèle est l'utilisation du mouvement brownien fractionnaire, en remplacement du mouvement brownien standard, pour caractériser la mémoire longue de la volatilité. Cela a une conséquence majeure sur la stratégie de gestion. En effet, l'utilisation du mouvement brownien fractionnaire dont les incréments sont corrélés, fait perdre au système son caractère markovien. Dans un cadre markovien, nous avons l'habitude de dire que "le futur ne dépend du passé que par le présent". Or ici, nous opérons dans un cadre non-markovien puisque l'état économique à un instant donné dépend de tout le passé. Par conséquent, l'approche classique par programmation dynamique et équation d'Hamilton-Jacobi-Bellman n'est plus applicable et nous avons été contraints de faire appel à une autre méthode de résolution. Par ailleurs, pour des raisons de simplification du problème à traiter, nous nous sommes intéressés à un cas simple de gestion : la maximisation de l'utilité espérée de la richesse finale à la date T. Sur le plan théorique, le cas avec consommation aurait pu être traité de la même manière, mais la mise en oeuvre est déjà extrêmement lourde du cas avec héritage seul.

La définition d'un tel modèle nous permet d'étudier l'impact de la mémoire longue de la volatilité sur l'allocation optimale en gestion de portefeuille et de mesurer l'importance de modéliser au mieux la volatilité dans ce type de problème. Pour mener à bien cette étude, nous nous penchons sur un pur cas de choix de portefeuille : nous utilisons le critère de gestion

précédemment cité associé à une fonction d'utilité CRRA.

Ce chapitre est composé de trois sections. Dans la première nous abordons l'aspect théorique du modèle. Puis, nous commençons par étudier le processus de la log-volatilité à mémoire longue afin de nous faire une première idée de l'importance de la prise en compte de la mémoire longue (section 2). Enfin, la troisième section est consacrée aux simulations et à l'étude principale de ce chapitre : l'impact de la mémoire longue sur l'allocation optimale.

6.1 Présentation du modèle avec volatilité stochastique à mémoire longue

Pour introduire la mémoire longue dans le modèle à volatilité stochastique précédent, nous faisons appel au mouvement brownien fractionnaire [48]. Il s'agit de la forme la plus simple d'un processus stochastique dans lequel la mémoire longue est prise en compte. Mandelbrot et Van Ness (1968) ont désigné par mouvement brownien fractionnaire, une famille de fonctions aléatoires gaussiennes définies de la manière suivante : "$B(t)$ being a ordinary Brownian motion and d a parameter satisfying $-1/2 < d < 1/2$, fBm of exponent d is a moving average of $dB(t)$, in which the past increments of $B(t)$ are weighted by the kernel $(t-s)^d$".

Remarque 5. Dans l'article original, Mandelbrot et Van Ness utilisent la notation B^H, où H est le paramètre de Hurst. Cette notation se retrouve également dans de nombreux autres articles. Afin d'éviter la confusion avec la fonction d'héritage d'une part et afin d'uniformiser les notations avec la partie précédente d'autre part, nous adaptons les notations pour ne faire apparaître que le paramètre de mémoire longue d grâce à la relation $H = d + 1/2$.

La définition du mouvement brownien fractionnaire est la suivante :

Définition 15. *Soit la filtration $\mathcal{F}_t = \sigma(B^d(s), 0 \leq s \leq t)$. Un mouvement brownien fractionnaire de paramètre $d \in]-1/2, 1/2[$ est un processus gaussien réel centré, adapté à la filtration $\mathcal{F}_t$, noté $\left\{B^d(t), \mathcal{F}_t, 0 \leq t < \infty\right\}$ défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, ayant les propriétés suivantes :*

(i) $B^d(0) = 0 \qquad \mathbb{P} - p.s.,$

(ii) $E\left[\left(B^d(t)\right)^2\right] = |t|^{2d+1} \quad \forall\, t \in \mathbb{R},$

(iii) Pour tout $0 \leq s \leq t$, les accroissements de $B^d(t) - B^d(s)$ sont stationnaires.

Lorsque $d > 0$, le mouvement brownien fractionnaire est capable de générer de la mémoire longue. En effet, en posant $r(k) = \mathrm{E}\left[B(1)^d(B(k+1)^d - B(k)^d)\right]$, nous avons, conformément à la définition 2 de la section 1.3.1,

$$\sum_{k=1}^{\infty} r(k) = +\infty. \tag{6.1}$$

48. Le lecteur est invité à se référer à Mandelbrot et Van Ness (1968), Hog et Frederiksen (2006) et Oksendal (2004).

La figure 6.1 illustre l'impact de d sur l'allure du mouvement brownien fractionnaire. Nous avons tracé, pour les mêmes réalisations du brownien sous-jacent, les browniens fractionnaires correspondant à différentes valeurs de d : 0.05 (brownien standard), 0.25 et 0.45 sur un intervalle de temps $[0, T]$. Nous remarquons que la mémoire longue a tendance à lisser la trajectoire du mouvement brownien fractionnaire. Cela a pour effet d'augmenter la prévisibilité à court terme. Le lissage de la trajectoire provient du fait que la corrélation produit un effet d'entraînement (chaque incrément dépend des incréments passés) et que la variance des incréments dépend de d. Par ailleurs, une propriété et les théorèmes importants concernant le mouvement brownien fractionnaire que nous auront à utiliser par la suite sont réunis en annexe 6.A.

6.1.1 Définition du modèle

Pour introduire de la mémoire longue dans le processus de la log-volatilité, nous remplaçons le mouvement brownien $Z(t)$ du modèle avec volatilité stochastique sans mémoire longue, présenté dans le chapitre 5, par $Z^d(t)$ composé d'un mouvement brownien $B(t)$ et d'un mouvement brownien fractionnaire $Y^d(t)$. Notre modèle est défini par

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (6.2) \\ dS_1(t) = \mu S_1(t)dt + V(t)S_1(t)dB(t), & (6.3) \\ d(\ln V(t)) = \kappa(\alpha - \ln V(t))dt + \sigma dZ^d(t), & (6.4) \end{cases}$$

où $dZ^d(t) = \rho dB(t) + \sqrt{1-\rho^2}dY^d(t)$. Nous supposons qu'il n'y a pas de corrélation entre le mouvement brownien et le mouvement brownien fractionnaire, ce qui d'un point de vue mathématique s'exprime $\mathrm{E}\left[dB(t)dY^d(t)\right] = 0$. En revanche, la définition de $dZ^d(t)$ entraîne une corrélation entre le brownien de l'actif risqué et le brownien fractionnaire composé de la volatilité.

$$\mathrm{E}\left[dB(t)dZ^d(t)\right] = \rho dt \tag{6.5}$$

Remarque 6. $Z^d(t)$ n'est pas un mouvement brownien fractionnaire car il ne vérifie pas le deuxième point de la définition 15. En effet, $\mathrm{E}[Z^d(t)] = \rho^2 t + (1-\rho^2)|t|^{2d+1} \neq |t|^{2d+1}$. Ainsi, nous n'utiliserons pas un processus d'Ornstein-Uhlenbeck fractionnaire pour décrire la log-volatilité, mais un processus voisin.

Remarque 7. Si $d = 0$, alors $Y^d(t)$ devient un mouvement brownien et $Z^d(t)$ devient également un mouvement brownien tel que $\mathrm{Corr}(dB(t), dZ^d(t)) = \rho$. Par conséquent, la log-volatilité suit un processus d'Ornstein-Ulhenbeck.

A l'image de ce que nous avons fait pour le modèle à volatilité stochastique, nous allons, dans un souci de lisibilité, exprimer l'équation différentielle stochastique de la volatilité. La théorie stochastique classique ne peut pas s'appliquer au mouvement brownien fractionnaire. En effet, comme le rappelle Duncan, Hu, et Pasik-Duncan (2000), un processus stochastique possède la propriété suivante : si l'intégrale stochastique $\int F(t)d\pi_t$ est bien définie pour une grande classe d'intégrant $(F(t), t \geq 0)$, alors le processus $(\pi(t), t \geq 0)$ est une semi-martingale (Dellacherie et Meyer 1982). Or, le mouvement brownien fractionnaire n'est pas une semi-martingale.

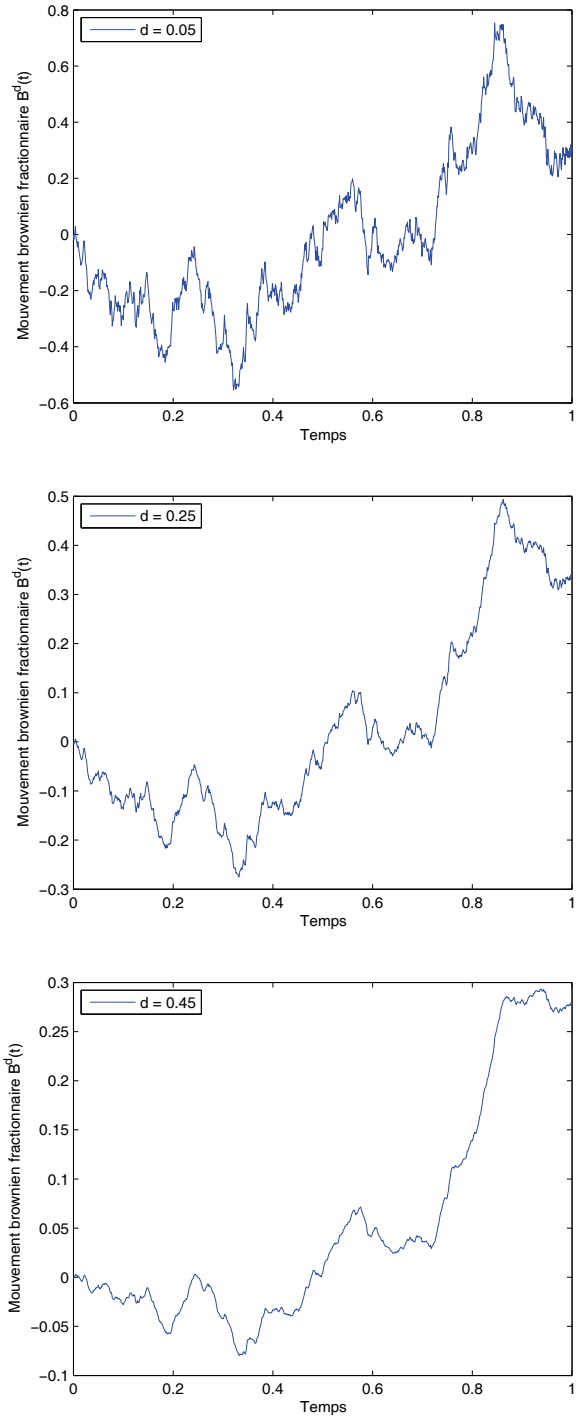

FIGURE 6.1 – Mouvements browniens fractionnaires ayant les mêmes réalisations du brownien sous-jacent mais différentes valeurs de d, 1000 points. En haut $d = 0.05$, au milieu $d = 0.25$ et en bas $d = 0.45$.

Il faut donc étendre la théorie classique de l'analyse stochastique au cas particulier du mouvement brownien fractionnaire. Dans ce sens, Lin (1995) Dai et Heyde (1996) ont introduit les intégrales stochastiques et la formule d'Itō adaptées au brownien fractionnaire. En faisant appel au lemme d'Itō fractionnaire multidimensionnel de Biagini, Hu, Oksendal, et Sulem (2002) et au théorème de dérivation des fonctions composées (la démonstration se situe en annexe 6.B), nous obtenons l'expression de $dV(t)$ et les équations permettant de décrire l'état de l'économie à un instant t se résument à

$$\begin{cases} dS_0(t) = rS_0(t)dt, & (6.6)\\ dS_1(t) = \mu S_1(t)dt + V(t)S_1(t)dB(t), & (6.7)\\ dV(t) = \left\{\sigma^2\left[\dfrac{\rho^2}{2} + \left(d+\dfrac{1}{2}\right)(1-\rho^2)f(t)\right] + \kappa(\alpha - \ln V(t))\right\}V(t)dt & \\ \qquad\qquad +\sigma V(t)dZ^d(t), & (6.8)\end{cases}$$

avec

$$f(t) := \begin{cases} 1, & \text{si } d = 0,\\ 2d\int_0^t \mathrm{e}^{-\kappa(t-r)}|t-r|^{2d-1}dr, & \text{si } 0 < d < 1/2.\end{cases} \tag{6.9}$$

Remarque 8. Si $d = 0$, alors on retrouve l'équation de la dynamique de la volatilité en l'absence de mémoire longue.

$$dV(t) = \left[\frac{\sigma^2}{2} + \kappa(\alpha - \ln V(t))\right]V(t)dt + V(t)dZ(t).$$

Dans le cas mémoire longue, on ne peut plus utiliser la même théorie que précédemment. En effet, le système n'est plus markovien donc J ne peut pas être une simple fonction de W et V à l'instant considéré, mais une fonctionnelle de $W(s)$ et $V(s)$ pour $s \leq t$, ce qui complique singulièrement les choses. Nous avons donc abandonné la méthode par programmation dynamique et considérons que l'investisseur cherche à maximiser l'utilité espérée de sa richesse à l'horizon de gestion T.

6.1.2 Étude du critère de gestion sans consommation de l'investisseur

La dynamique de la richesse du portefeuille s'écrit

$$\begin{cases} dW(t) = [(\mu - r)w(t) + r]W(t)dt + w(t)V(t)W(t)dB(t), & (6.10)\\ dV(t) = \left\{\sigma^2\left[\dfrac{\rho^2}{2} + \left(d+\dfrac{1}{2}\right)(1-\rho^2)f(t)\right] + \kappa(\alpha - \ln V(t))\right\}V(t)dt & \\ \qquad\qquad +\sigma V(t)dZ^d(t), & (6.11)\end{cases}$$

où $B(t)$ est un brownien standard, $Z^d(t)$ est tel que $dZ^d(t) = \rho dB(t) + \sqrt{1-\rho^2}dY^d(t)$ où $Y^d(t)$ est un Brownien fractionnaire de paramètre d.

Sous forme intégrale, la richesse s'écrit

$$W(t) = W(0)\exp\left\{\int_0^t \left[(\mu - r)w(s) + r - \frac{1}{2}w(s)^2V(s)^2\right]ds\right\} \times \exp\left\{\int_0^t w(s)V(s)dB(s)\right\}. \quad (6.12)$$

A ce niveau là, le fait que $Z^d(t)$ soit un Brownien fractionnaire, ou non, et donc que $V(t)$ soit un processus à mémoire longue, ou non, ne change rien.

L'objectif de l'investisseur se résume à maximiser sa richesse terminale, soit

$$\max_{w(t)} \mathrm{E}\left[H(W(T))\right]. \quad (6.13)$$

Nous considérons la même fonction héritage que dans les deux sections précédentes, à savoir

$$H(W) = \frac{W^\gamma}{\gamma}, \qquad \gamma < 1, \gamma \neq 0. \quad (6.14)$$

A l'aide du lemme d'Itō fractionnaire (voir théorème 2 en annexe 6.A)combiné à la dérivée de Malliavin fractionnaire, le poids optimal de l'actif risqué, dont la démonstration est détaillée en annexe 6.C, s'écrit

$$w^*(t) = \frac{\mu - r}{(1-\gamma)V(t)^2} - \frac{\sigma\rho\gamma}{(1-\gamma)V(t)} \times \frac{\mathrm{E}\left[W(T)^\gamma\left(\int_t^T w(s)^2V(s)^2\mathrm{e}^{-\kappa(s-t)}ds - \int_t^T w(s)V(s)\mathrm{e}^{-\kappa(s-t)}dB(s)\right)|\mathcal{F}_t\right]}{\mathrm{E}\left[W(T)^\gamma|\mathcal{F}_t\right]}, \quad (6.15)$$

où $\mathcal{F}_t$ est la filtration naturelle, associée à $B(t)$ et $Y^d(t)$. La mémoire longue a un impact sur la valeur de w^* puisqu'elle est implicitement contenue dans le poids optimal de l'actif risqué via $V(t)$.

Le poids optimal de l'actif risqué a deux composantes. La première est la demande du portefeuille moyenne-variance, et la deuxième est la demande de couverture intertemporelle de Merton. Contrairement au modèle avec volatilité stochastique de Chacko et Viceira (2005), où il est clair que le rapport entre la demande du portefeuille moyenne-variance et la demande de couverture intertemporelle est indépendante du niveau de la volatilité (voir équation (4.17)), ici, nous ne pouvons rien conclure en raison de l'expression dans l'espérance au numérateur. A noter que ce n'est pas une spécificaté de la mémoire longue puisque cela était déjà le cas avec le modèle à volatilité stochastique.

En regardant de plus près l'équation (6.15), on constate que la demande de couverture intertemporelle est toujours nulle dans trois cas. Premièrement, lorsque la volatilité de la log-volatilité est nulle ($\sigma = 0$). Deuxièmement, lorsque l'investisseur ne peut avoir recours à l'actif risqué pour se couvrir des changements de la log-volatilité ($\rho = 0$). Et, Troisièmement, lorsque l'aversion au risque relatif de l'investisseur vaut 1 ($\gamma = 0$). Ce sont les mêmes conditions que celles que trouvent Chacko et Viceira (2005) dans le cas où l'élasticité de la substitution intertemporelle de la consommation de l'investisseur est unitaire .

6.2 Étude du processus de la log-volatilité à mémoire longue

Dans cette section, nous étudions le processus de la log-volatilité (voir équation (6.4)) afin d'évaluer l'impact du paramètre de mémoire longue sur la volatilité et ainsi de mieux cerner les valeurs de d intéressantes pour mener à bien l'étude de l'impact de la mémoire longue sur la gestion de portefeuille. Nous étudions également l'autocovariance de la log-volatilité dans le but de déterminer le ou les paramètres impliqués dans la mémoire courte et la mémoire longue et donc pour paramétrer correctement nos modèles.

6.2.1 Impact du paramètre de mémoire longue

Nous commençons par observer l'effet de d sur l'allure de la trajectoire de la log-volatilité. Sur la figure 6.2, nous traçons la volatilité pour un même jeu de paramètre et différentes valeurs de d : $d \in [0, 0.1,]$ dans le graphique du haut et $d \in [0.1, 0.3,]$ dans le graphique du bas.

Dans le graphique du haut la différence est très nette alors que d ne varie que d'un dixième. En revanche, dans le graphique du bas et pour une variation deux fois supérieure, les courbes sont beaucoup plus proches et parfois confondues. Nous constatons un phénomène de saturation de l'impact de d sur l'allure de la trajectoire de la volatilité : au-delà de $d = 0.1$, l'écart entre les courbes de volatilité devient faible.

6.2.2 Autocovariance de la log-volatilité

Pour une question de stationnarité du processus, nous réécrivons le processus de la log-volatilité (voir équation (6.B.4) en annexe 6.B), comme suit

$$\ln V_t = \alpha + \mathrm{e}^{-\kappa t}(\ln V_0 - \alpha) + \sigma\rho \int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)} dB_s + \sigma\sqrt{1-\rho^2} \int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)} dY_s^d. \tag{6.16}$$

L'espérance et la variance de la log-volatilité valent respectivement

$$\mathrm{E}\left[\ln V_t\right] = \alpha, \tag{6.17}$$

$$\mathrm{Var}\left[\ln V_t\right] = \frac{\sigma^2\rho^2}{2\kappa} + \frac{\sigma^2(1-\rho^2)}{2\kappa^{2d+1}} V_1, \tag{6.18}$$

avec, après un changement de variable pour faire sortir κ de l'intégrale double, puis deux intégrations par parties successives,

$$V_1 = \int_0^{+\infty} \int_0^{+\infty} \mathrm{e}^{-(x+y)} |y-x|^{2d+1} dx dy. \tag{6.19}$$

Remarque 9. Lorsque $d = 0$, alors $V_1 = 1$ et nous retrouvons bien la variance de la log-volatilité dans le cadre du modèle à volatilité stochastique sans mémoire longue.

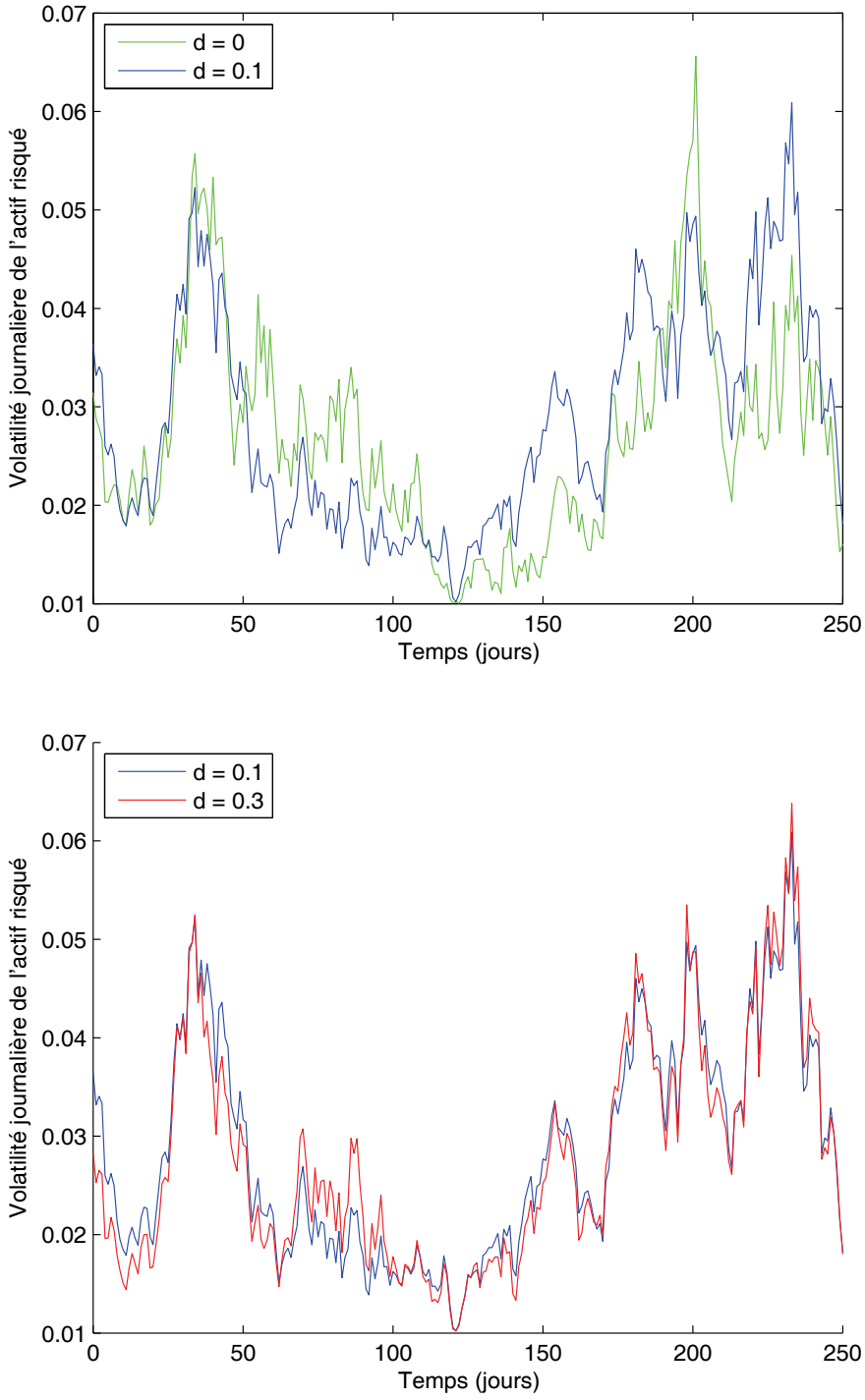

FIGURE 6.2 – Illustration de l'impact du paramètre à mémoire longue sur l'allure de la trajectoire de la volatilité pour $\kappa = 12$, $sigma = 2.5$ et $\alpha = -0.92$.

L'autocovariance de la log-volatilité $\gamma(\tau) = \mathrm{Cov}(\ln V_t, \ln V_{t+\tau})$, dont la démonstration est fournie en annexe 6.D, s'écrit pour τ grand

$$\gamma(\tau) \simeq \sigma^2 \frac{\mathrm{e}^{-\kappa\tau}}{2\kappa} + \sigma^2(1-\rho^2)\left[\frac{\mathrm{e}^{-\kappa\tau}}{2\kappa}\left(\frac{\Gamma(2d+1)}{2\kappa^{2d}} - 1\right) + 2d(2d+1)\frac{\tau^{2d-1}}{2\kappa^2}\right]. \tag{6.20}$$

Le terme $\frac{\tau^{2d-1}}{2\kappa^2}$ indique, conformément nos attentes, que la décroissance de la fonction d'autocovariance est hyperbolique, ce qui confirme la présence de mémoire longue.

6.3 Étude de l'impact de la mémoire longue de la volatilité stochastique en gestion de portefeuille

Dans cette section, nous commençons par définir le critère de gestion étudié. Puis nous détaillons la méthode numérique employée pour approcher la solution de notre problème. Ensuite, nous fournissons et justifions les paramètres utilisés dans notre étude de l'impact de la mémoire longue sur l'allocation optimale. Enfin, nous exposons les résultats de cette étude.

6.3.1 Définition du problème

Contrairement au modèle précédent où nous avions le choix du critère de gestion et de la fonction d'utilité, avec le modèle avec volatilité stochastique à mémoire longue, ce n'est plus le cas puisque nous avons dérivé les résultats analytiques seulement lorsque le gestionnaire cherche à maximiser l'utilité espérée de sa richesse terminale pour une utilité de type CRRA. La consommation de l'investisseur étant absente du problème d'optimisation, seul le poids optimal de l'actif risqué a été obtenu. Nous rappelons que celui-ci est donné par

$$\begin{aligned} w^*(t) = {} & \frac{\mu - r}{(1-\gamma)V(t)^2} - \frac{\sigma\rho\gamma}{(1-\gamma)V(t)} \\ & \times \frac{\mathrm{E}\left[W(T)^\gamma \left(\int_t^T w^*(s)^2 V(s)^2 \mathrm{e}^{-\kappa(s-t)} ds - \int_t^T w^*(s) V(s) \mathrm{e}^{-\kappa(s-t)} dB(s)\right) | \mathcal{F}_t\right]}{\mathrm{E}\left[W(T)^\gamma | \mathcal{F}_t\right]}. \end{aligned} \tag{6.21}$$

6.3.2 Méthodologie

Pour déterminer le poids optimal de l'actif risqué $w^*(t)$ pour tout $t \in [0, T[$, nous devons discrétiser les intégrales dans le deuxième terme de l'équation (6.21) avec un pas τ suffisamment petit. Nous utilisons la méthode des rectangles : nous supposons la volatilité constante sur l'intervalle $[T-(k+1)\tau, T-k\tau[$ et égale à $V(T-(k+1)\tau)$ et nous faisons tendre τ vers zéro. Nous nous affranchissons de la rigueur mathématique.

Posons $\tau = \frac{T}{N}$ avec N suffisamment grand. Nous supposons que $w^*(t)$ est constant sur l'intervalle $[T-(k+1)\tau, T-k\tau[$ et notons $w^*_{T-k\tau}$ la valeur approchée de $w^*(T-k\tau)$. Notre

objectif est de calculer numériquement $w^*_{T-\tau}$, $w^*_{T-2\tau}$, et ainsi de suite jusqu'à w^*_0. Nous montrons dans l'annexe 6.E que

Proposition 5. La valeur approchée du poids optimal de l'actif risqué (6.21) est donné par

$$w^*_{T-k\tau} = \frac{\mu - r}{(1-\gamma)V^2_{T-k\tau}} - \frac{\sigma\rho\gamma}{(1-\gamma)V_{T-k\tau}} \frac{\tilde{B}_{T-k\tau}}{\tilde{A}_{T-k\tau}}, \tag{6.22}$$

avec

$$\begin{aligned}\tilde{A}_{T-k\tau} \approx \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau} + r - \frac{1}{2}w^2_{T-k\tau}V^2_{T-k\tau}\right]\tau\right\} \\ \times \mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right],\end{aligned} \tag{6.23}$$

et

$$\begin{aligned}\tilde{B}_{T-k\tau} = \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau} + r - \frac{1}{2}w^2_{T-k\tau}V^2_{T-k\tau}\right]\tau\right\} \\ \times\Big(\mathrm{e}^{-\kappa\tau}\mathrm{E}\left[\tilde{B}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ + w^2_{T-k\tau}V^2_{T-k\tau}\frac{1-\mathrm{e}^{-\kappa\tau}}{\kappa}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ -w_{T-k\tau}V_{T-k\tau}\sqrt{\frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}\eta_{T-k\tau}|\mathcal{F}_{T-k\tau}\right]\Big),\end{aligned} \tag{6.24}$$

avec pour condition terminale respective $\tilde{A}_T = 1$ et $\tilde{B}_T = 0$, et où

$$\xi_{T-k\tau} \sim N(0,1), \tag{6.25}$$

$$\eta_{T-k\tau} \sim N(0,1), \tag{6.26}$$

$$\mathrm{Corr}(\xi_{T-k\tau}, \eta_{T-k\tau}) = \sqrt{\frac{2(1-\mathrm{e}^{-\kappa\tau})}{\kappa\tau(1+\mathrm{e}^{-\kappa\tau})}}. \tag{6.27}$$

Les variables ξ_t et η_t n'ont pas de corrélation sérielle.

Remarque 10. Lorsque $\kappa\tau$ tend vers 0, le développement limité de $\mathrm{e}^{-\kappa\tau}$ à l'ordre 3 nous donne $\mathrm{Corr}(\xi_{T-k\tau}, \eta_{T-k\tau}) = 1$.

Pour calculer $w_{T-k\tau}$ il faut connaître $V_{T-k\tau}$ et également $\ln V_{T-k\tau}$. Nous montrons (dans l'annexe 6.F) que

Proposition 6. La log-volatilité peut être approchée par

$$\begin{aligned}\ln V_{T-(k-1)\tau} = \alpha + \mathrm{e}^{-\kappa\tau}(\ln V_{T-k\tau} - \alpha) \\ + \sigma\left(\rho\sqrt{\frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}}\zeta_{T-k\tau} + \sqrt{1-\rho^2}\sqrt{V_1}\chi_{T-k\tau}\right),\end{aligned} \tag{6.28}$$

où,

$$V_1 = \tau^{2d+1}\left\{\mathrm{e}^{-\kappa\tau} + \frac{(\kappa\tau\mathrm{e}^{-\kappa\tau})^2}{2}\int_0^1\int_0^1 \mathrm{e}^{\kappa\tau(x+y)}|x-y|^{2d+1}dxdy\right\}, \tag{6.29}$$

et pour tout $t \in [0,T]$,

$$\chi_t \sim N(0,1), \tag{6.30}$$
$$\zeta_t \sim N(0,1), \tag{6.31}$$
$$\mathrm{Corr}(\zeta_t, \chi_t) = \frac{2\kappa\tau}{\mathrm{e}^{\kappa\tau} - \mathrm{e}^{-\kappa\tau}}. \tag{6.32}$$

Les variables ζ_t sont indépendantes, tandis que χ_t sont sériellement corrélées et

$$\mathrm{E}\left[\chi_t, \chi_{t+\tau}\right] = 2d\left(d+\frac{1}{2}\right)\mathrm{e}^{-2\tau}\int_0^\tau\int_0^\tau \mathrm{e}^{\kappa(u+v)}|u-v-\tau|^{2d-1}dudv. \tag{6.33}$$

Puis nous montrons (dans l'annexe 6.G) que

Proposition 7. La richesse totale du portefeuille et le cours de l'actif risqué sont calculés par

$$W_{T-(k-1)\tau} \approx W_{T-k\tau}\exp\left\{\left[(\mu-r)w_{T-k\tau} + r - \frac{1}{2}w_{T-k\tau}^2 V_{T-k\tau}^2\right]\tau + \sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}, \tag{6.34}$$

$$S_{1,T-(k-1)\tau} \approx S_{1,T-k\tau}\exp\left\{\left(\mu - \frac{1}{2}V_{T-k\tau}^2\right)\tau + \sqrt{\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}. \tag{6.35}$$

Pour obtenir une approximation du poids optimal de l'actif risqué, nous proposons de procéder comme suit :

1. Tirer une trajectoire de volatilité sur $[0,T]$,
2. Puis N trajectoires de ξ et η sur $[0,T]$, conformément aux indications de la proposition 5,
3. Pour $k = 0$,
   (a) Partir des conditions finales $\tilde{A}_T = 1$ et $\tilde{B}_T = 0$,
   (b) En déduire d'après l'expression (6.22) que $w_T^* = \frac{\mu - r}{(1-\gamma)V_T^2}$,
4. Pour $k = 1$,
   (a) Simplifier les expressions $\tilde{A}_{T-\tau}$ et $\tilde{B}_{T-\tau}$ (respectivement (6.23) et (6.24)) grâce aux conditions finales $\tilde{A}_T = 1$ et $\tilde{B}_T = 0$,
   (b) Calculer une valeur approchée de chaque espérance de $\tilde{A}_{T-\tau}$ et $\tilde{B}_{T-\tau}$ par la technique de Monte-Carlo, par exemple

$$\mathrm{E}\left[\exp\left\{\gamma\sqrt{\tau}w_{T-\tau}V_{T-\tau}\xi_{T-\tau}\right\}\eta_{T-\tau}\right] \approx \frac{1}{N}\sum_{i=1}^{N}\exp\left\{\gamma w_{T-\tau}V_{T-\tau}\xi_{T-\tau}^i\right\}\eta_{T-\tau}^i, \tag{6.36}$$

   puisque $w_{T-\tau}$ et $V_{T-\tau}$ sont supposées connues à l'instant $T-\tau$.
   (c) Obtenir $w_{T-\tau}^*$ en résolvant l'équation (6.22) simplifiée par la méthode de Newton-Raphson,
5. Pour $k \geq 2$ jusqu'à $k = N$,

(a) Calculer une valeur approchée de chaque espérance de $\tilde{A}_{T-k\tau}$ et $\tilde{B}_{T-k\tau}$ par la technique de Monte-Carlo, en utilisant $\tilde{A}_{T-(k-1)\tau}$ et $\tilde{B}_{T-(k-1)\tau}$ calculée précédemment,

(b) Obtenir $w^*_{T-k\tau}$ en résolvant l'équation (6.22) par la méthode de Newton-Raphson,

(c) Tirer un nombre L suffisamment grand de trajectoires de volatilité (par exemple 200) sur $[T-k\tau, T-(k-1)\tau[$,

(d) Pour chacune de ces trajectoires, déterminer $w^{*(j)}_{T-(k-1)\tau}$, $j = 1, \cdots, L$ par la méthode de Newton-Raphson,

(e) En déduire $\tilde{A}_{T-(k-1)\tau}$ et $\tilde{B}_{T-(k-1)\tau}$.

Remarque 11. En raison de la présence de $\tilde{A}_{T-(k-1)\tau}$ et $\tilde{B}_{T-(k-1)\tau}$ dans les espérances conditionnelles à l'information disponible au temps $T-k\tau$, nous avons développé un algorithme récursif.

6.3.3 Paramétrage

L'investisseur, moyennement risquophobe ($\gamma = -2$) a pour objectif de maximiser uniquement l'utilité de sa richesse terminale à un horizon de gestion de $t = 1$ an. Il peut répartir sa richesse entre un actif non risqué, un compte bancaire rémunéré au taux annuel sans risque $r = 5\%$, et un actif risqué, une action dont le rendement annuel espéré est $\mu = 15\%$. La log-volatilité des rendements de cette action a une moyenne de $\alpha = \log(0.40)$, une volatilité de 0.2 et un retour à la moyenne de l'ordre de quatre mois ($\kappa = 3$). La corrélation entre le rendement de l'action et sa volatilité est fixée à $\rho = -0.25$. En $t = 0$, la valeur de la log-volatilité vaut $\log(0.40)$. Ces paramètres sont résumés dans le tableau 6.1

Actif non risqué	Taux sans risque	$r = 5\%$
Actif risqué	Rendement moyen	$\mu = 15\%$
Log-volatilité	Retour à la moyenne	$1/\kappa = 1/3$
	Valeur initiale	$\log V_0 = \log(0.40)$
	Moyenne à long terme	$\alpha = \log(0.40)$
	Volatilité	$\sigma = 1$
	Corrélation	$\rho = -0.25$
Fonction d'utilité	CRRA	$\gamma = -2$

TABLE 6.1 – Paramètres du modèle de choix de portefeuille avec volatilité stochastique à mémoire longue pour un horizon de gestion de 1 an.

Puisque d'une part nous avons montré qu'au delà de $d = 0.1$ nous observons un phénomène de saturation de l'impact du paramètre de mémoire longue sur l'allure de la volatilité, et d'autre part les estimations de d dans la première partie de la thèse ont montré que dans la majorité des cas $0.2 \leq d \leq 0.45$, alors nous fixerons $d = 0$ dans le cas de l'absence de mémoire longue et $d = 0.3$ dans le cas contraire. Cela nous permettra d'étudier l'impact de la mémoire longue sur l'allocation optimale.

6.3.4 Résultats

Convergence des modèles avec volatilité stochastiques

Afin de nous assurer de la bonne modélisation, nous observons la convergence des deux modèles avec volatilité stochastique pour le jeu de données utilisées dans l'étude précédente couplée au critère de gestion de l'étude actuelle, à savoir la maximisation de l'utilité espérée de l'héritage.

Nous traçons sur la figure 6.3 le poids optimal de l'actif risqué obtenu grâce au modèle à volatilité stochastique (voir section 5.2) et celui obtenu au moyen du modèle avec volatilité stochastique à mémoire longue (voir section 6.1). Pour que la comparaison ait un sens, les tirages sous-jacents de loi normale des variables ξ, η et ζ sont identiques dans les deux optimisations et d est fixé à 0 dans le modèle à mémoire longue. Si les deux méthodes d'approximation numérique sont correctes, nous devrions donc obtenir des courbes confondues. Nous constatons que les courbes, bien qu'elle ne soient pas totalement superposées, sont très proche l'une de l'autre. Ces légères différences sont dues aux approximations des deux méthodes numériques utilisées pour résoudre notre problème de choix de portefeuille.

Sur le graphique du bas de cette même figure, nous montrons l'erreur relative (en %) du poids optimal de l'actif risqué. Nous observons que l'erreur absolue varie entre 0 et 3.5% et a une valeur moyenne de 1.33%. Nous pouvons donc conclure que pour un même problème donné, les deux modèles convergent plutôt bien. En outre, l'erreur pourrait être réduite en prenant une grille plus fine mais en contrepartie le temps de calcul augmenterait.

Allocation optimale pour un jeu de paramètres donnés

Pour commencer, nous paramétrons notre modèle à partir des données fournies dans le tableau 6.1. La période de gestion est d'un an. Nous avons joint à la figure 6.4 deux graphiques. Dans celui du haut, nous avons tracé deux trajectoires de volatilité, l'une est à mémoire longue ($d = 0.3$), et l'autre à mémoire courte($d = 0$). Quant au graphique du bas, il montre le poids optimal de l'actif risqué obtenu au moyen du modèle avec volatilité stochastique à mémoire longue (graphique du bas) lorsque $d = 0.3$ et également lorsque $d = 0$.

Nous constatons que la plupart du temps les deux trajectoires de volatilité sont éloignées l'une de l'autre. Parfois cet écart est supérieur à 40%. Cette différence se retrouve au niveau du poids optimal de l'actif risqué. Pour une aversion au risque égale à 3, et pour les tirages sous-jacents de loi normale qui ont conduit à ces trajectoires de volatilité, le poids optimal de l'actif risqué est compris entre 0 et 0.8 quelle que soit la valeur de d. De plus, pendant les deux tiers du temps, ce poids est compris entre 0.1 et 0.3, ce qui est relativement faible et compte tenu du niveau de volatilité moyennement élevé au même moment dénote une attitude assez risquophobe de l'investisseur. Enfin, il est raisonnable d'observer que le moment où le portefeuille est majoritairement constitué d'actif risqué correspond à celui où la volatilité est particulièrement faible.

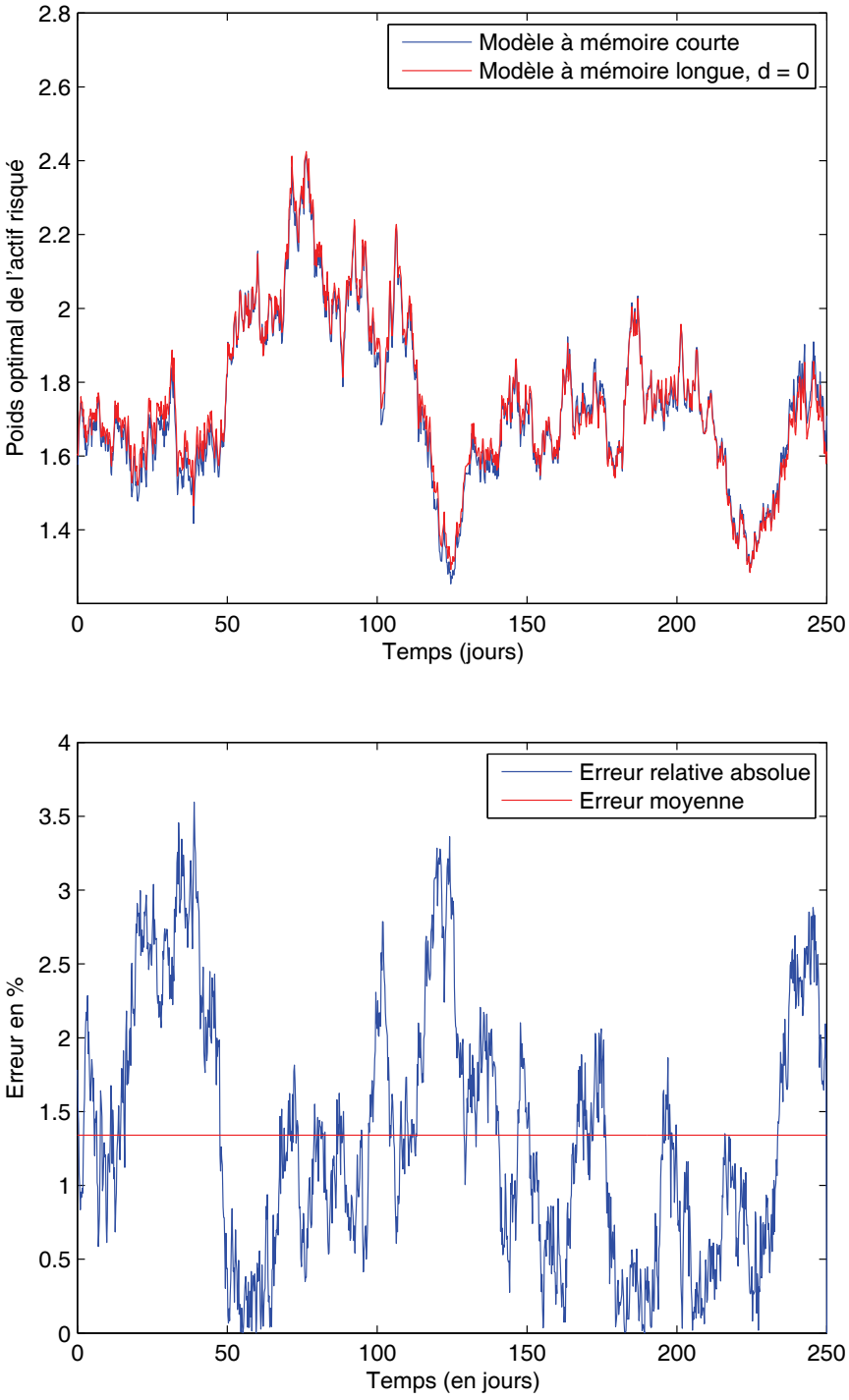

FIGURE 6.3 – Convergence du modèle avec volatilité stochastique et du modèle avec volatilité stochastique à mémoire longue, lorsque $d = 0$.

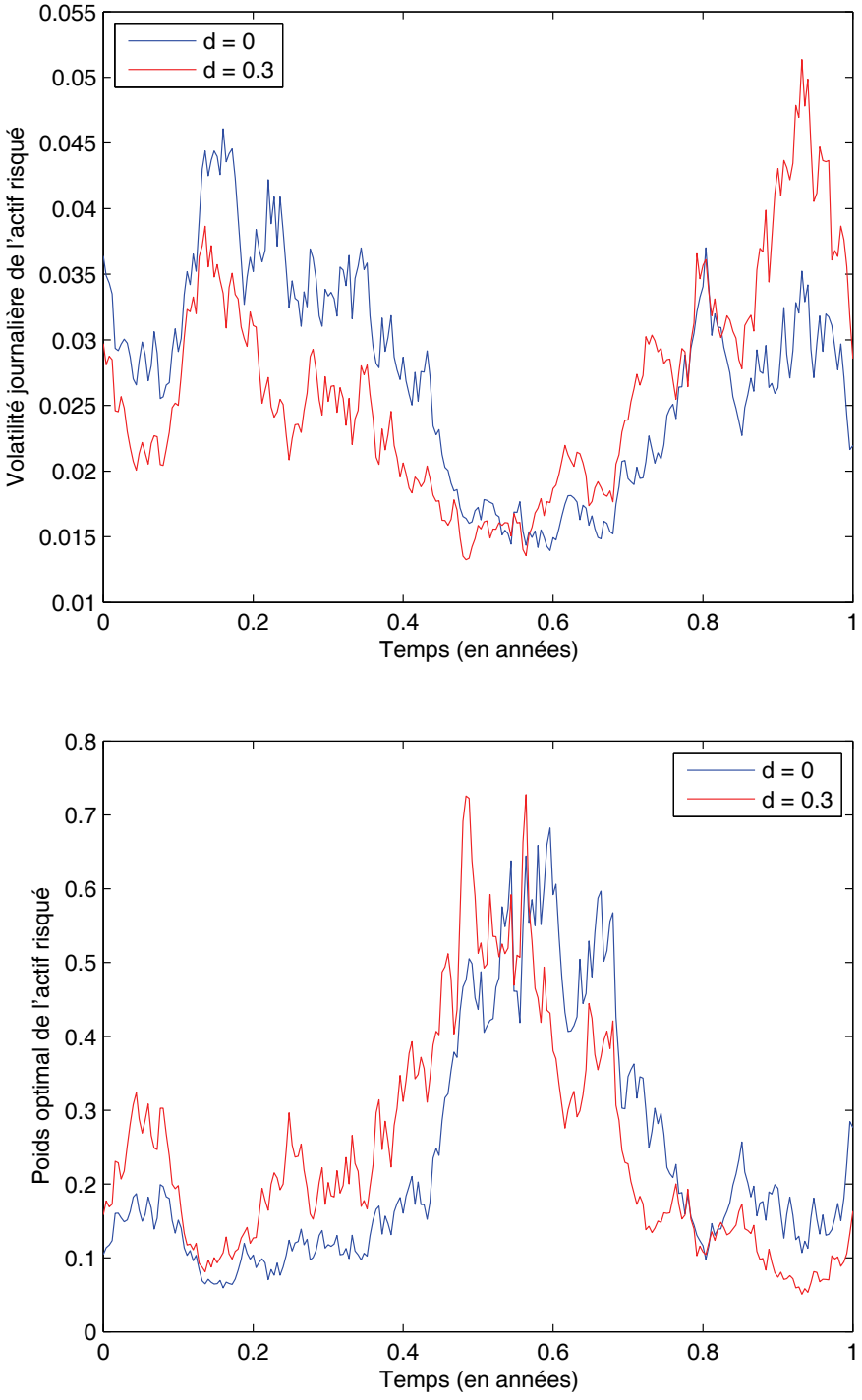

FIGURE 6.4 – Trajectoires de la volatilité de l'actif risqué et son poids optimal correspondant lorsqu'il y a mémoire longue ou au contraire mémoire courte, pour une période de gestion de 1 an.

Nous avons regroupé sur la figure 6.5 les trajectoires du cours de l'actif risqué lorsque $d = 0.3$ et $d = 0$ (graphique du haut) et celles de la richesse du portefeuille que nous avons déduites à partir du poids optimal de la volatilité.

Nous observons naturellement que l'allure de la trajectoire du prix de l'actif risqué se retrouve en partie dans celle de la richesse du portefeuille. Dans notre exemple, nous remarquons également que lorsque $d = 0.3$ le rendement du portefeuille sur un an est meilleur : il atteint 5.77% contre 3.50% lorsque $d = 0$ alors que le taux annuel sans risque est de 5%.

Nous avons dit plusieurs fois que la mémoire longue est un fait stylisé important de la volatilité. Donc, il paraît logique de dire que la volatilité à mémoire longue est la plus proche de la véritable volatilité de l'actif risqué. Par conséquent, nous supposons à présent que la véritable volatilité est la volatilité à mémoire longue lorsque $d = 0.3$. Ainsi, dans le cas où $d = 0$, nous traçons la richesse du portefeuille déduite du poids optimal de l'actif risqué obtenu pour $d = 0$, mais avec la véritable volatilité ($d = 0.3$). Cela revient à illustrer l'erreur de spécification du modèle. Dans cet esprit, nous avons superposé sur la figure 6.6 la trajectoire de la richesse pour $d = 0.3$ que nous avions déjà montré sur la figure précédente et celle de la richesse que nous qualifions de corrigée que nous venons d'expliquer.

Nous constatons immédiatement que la correction de la richesse du portefeuille a un impact très net sur le rendement du portefeuille : le rendement du portefeuille sur 1 an est descendu à 1.25%. Par ailleurs, si l'on regarde de plus près les rendements journaliers moyens du portefeuille, les différences sont également très nettes. A partir de la richesse corrigée, nous déduisons un rendement journalier moyen de 0.005 contre 0.0231 pour la richesse avec $d = 0.3$.

En réalité, ce raisonnement n'est pas tout à fait exact. En effet, comme nous l'avons montré précédemment, dans le cadre du modèle à mémoire courte, la décroissance de la fonction d'autocorrélation est contrôlée par κ. Lorsque $\kappa = 2$ par exemple, les corrélations persistent au delà d'une année. Par conséquent, l'absence de paramètre de mémoire longue dans le modèle à volatilité stochastique sans mémoire longue est en partie compensée par κ et donc il ne suffit pas de reprendre les mêmes paramètres que le modèle à volatilité stochastique à mémoire longue en posant $d = 0$. Toutefois, cela constitue une première étape permettant d'illustrer le rôle de la prise en compte de la mémoire longue dans les méthodes d'allocation optimale. C'est pourquoi, nous utilisons par la suite ce procédé.

En outre, à ce stade de notre étude, nous nous devons de mettre en avant le fait que par nature ces simulations sont des cas particuliers. Les résultats obtenus n'ont pas valeur de conclusion générale puisque d'autres tirages sous-jacents auraient pu aboutir à des conclusions différentes. Par exemple nous ne pouvons pas conclure que la richesse du portefeuille lorsque $d = 0.3$ est toujours supérieure à celle lorsque $d = 0$. Pour pouvoir tirer des conclusions, il faudrait avoir les moyens de réaliser un grand nombre de simulations pour pouvoir établir un comportement moyen. Hélas, en raison du temps de calcul nécessaire pour chaque optimisation, cela n'est pas réalisable avec le matériel informatique dont nous disposons. Toutefois, cela nous permet déjà de souligner l'importance de bien spécifier le modèle et dans notre cas de modéliser le plus justement possible la volatilité de l'actif risqué en prenant en compte le phénomène de mémoire longue.

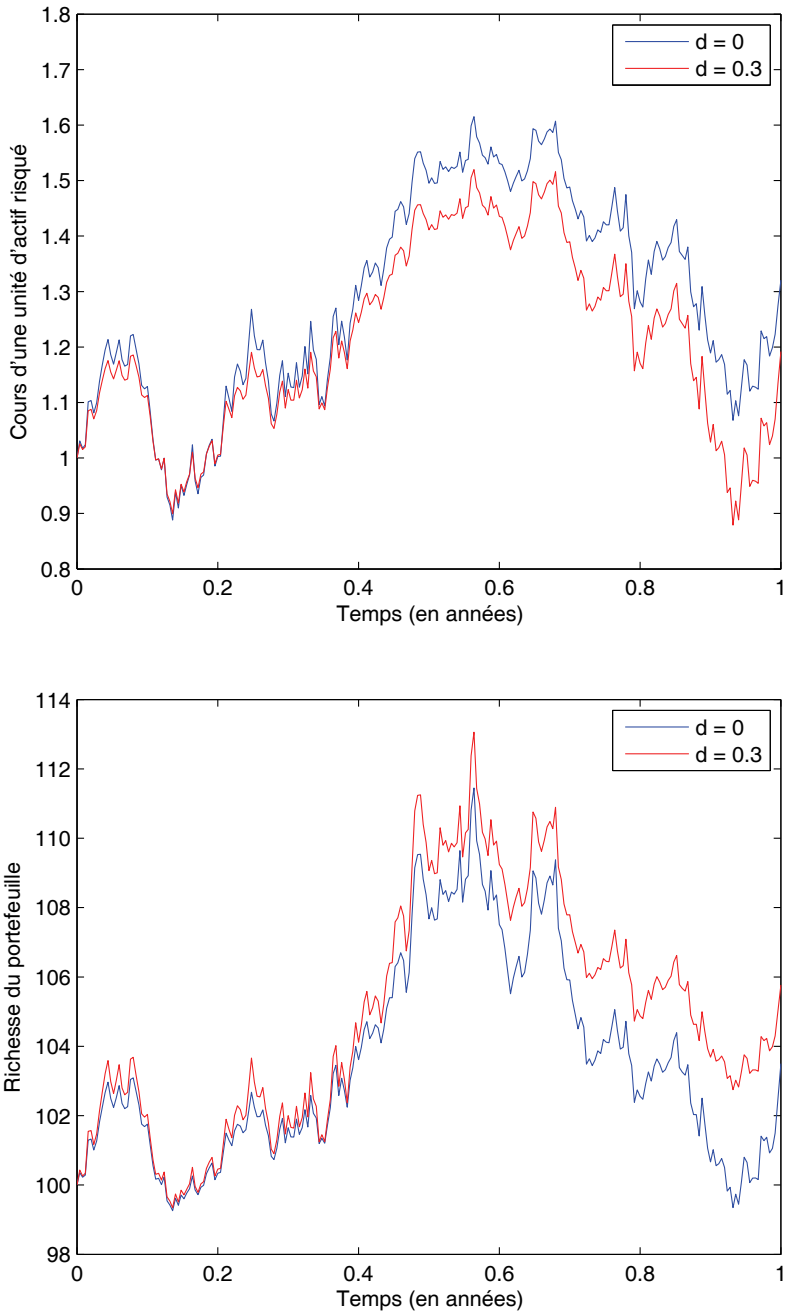

FIGURE 6.5 – Trajectoires du cours de l'actif risqué et de la richesse du portefeuille dans le cas où la volatilité de l'actif risqué est à mémoire longue ou à mémoire courte, pour une période de gestion de 1 an.

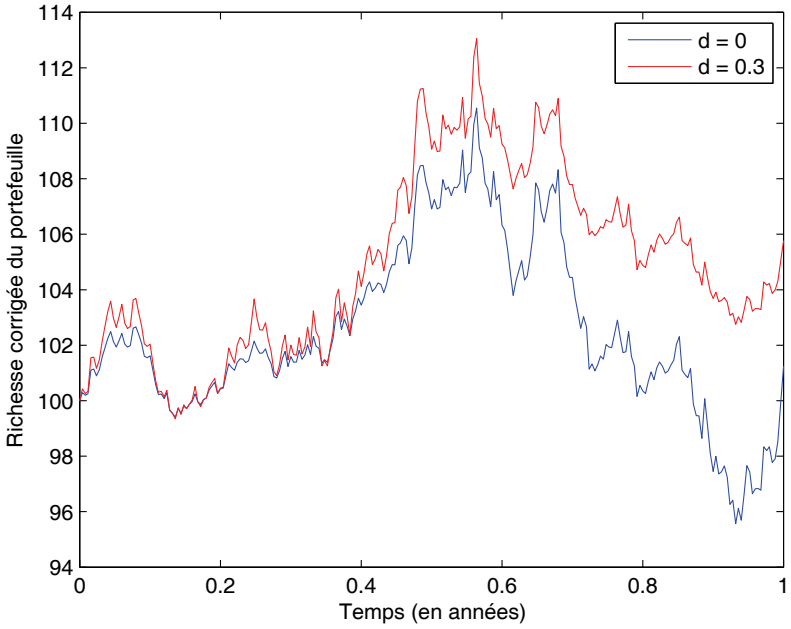

FIGURE 6.6 – Richesse du portefeuille lorsque $d = 0.3$ et richesse corrigée du portefeuille lorsque $d = 0$, pour une période de gestion de 1 an.

A présent, nous allons étudier successivement l'impact de différents paramètres tels que l'horizon de gestion, l'aversion au risque de l'investisseur, le niveau de volatilité sur l'allocation optimale.

Impact de l'horizon de gestion

L'intérêt de l'incorporation de la mémoire longue via un paramètre concret dans une méthode de choix de portefeuille est atténué par un horizon de gestion à court terme, comme ceci a été le cas jusqu'à présent ($T = 1$ an). En effet, puisque le paramètre d contrôle les corrélations à long terme, il parait plus intéressant d'observer les résultats d'une gestion à plus long terme comme par exemple 5 ou 10 ans.

La figure 6.7 est consacrée à la gestion avec un horizon à 5 ans. Sur le graphique du haut sont tracées les évolutions de la richesse dans le cas où $d = 0.3$ et dans le cas où $d = 0$. Sur le graphique du bas on trace à nouveau la richesse dans le cas où $d = 0.3$ avec cette fois la richesse corrigée pour $d = 0$.

Nous observons les deux mêmes phénomènes que pour la gestion à 1 an, à savoir : le rendement global du portefeuille est meilleur lorsque $d = 0.3$: $+41.49\%$ (pour un rendement journalier moyen de 0.0332), et le fait de corriger le calcul de la richesse entraîne une détérioration de la performance du portefeuille $+37.07\%$ (pour un rendement journalier moyen de 0.0297) contre $+40.29\%$. A titre de comparaison, le même capital initial placé au taux sans

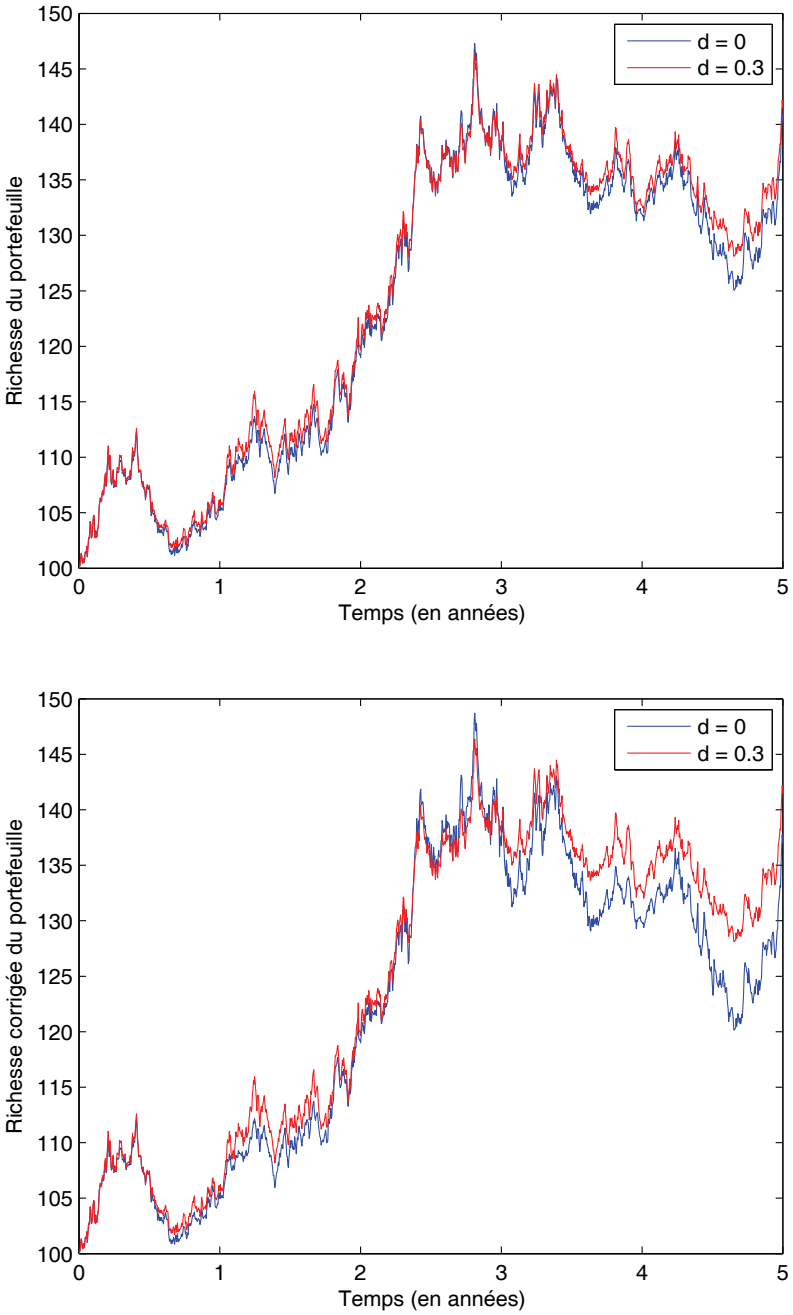

FIGURE 6.7 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) pour une période de gestion de 5 ans.

risque annuel de 5% aurait généré un rendement de +27.63%.

La figure 6.8 est identique à la figure 6.7, à une exception près : l'horizon de gestion est passé de 5 à 10 ans.

Les observations et conclusions sont les mêmes que celles de la figure précédente. Le rendement global du portefeuille lorsque la mémoire longue est prise en compte est de 100.16% contre 96.44% lorsqu'elle ne l'est pas (richesse corrigée) et un rendement journalier moyen de 0.0401 contre 0.0386. A titre de comparaison, le même capital placé au taux annuel sans risque de 5% aurait généré un rendement global de 62.89%.

Par la suite, pour ne pas surcharger cette section de figures, nous nous intéressons uniquement à la période de gestion de 5 ans.

Impact de l'aversion au risque de l'investisseur

A présent, nous voulons mesurer l'impact de l'aversion au risque de l'investisseur sur l'allocation optimale. Nous allons comparer le poids qu'accorde un investisseur très averse au risque ($\gamma = -19$), moyennement averse au risque ($\gamma = -2$) ou encore faiblement averse au risque ($\gamma = 0.5$) à l'actif risqué et à la richesse du portefeuille.

Le caractère très peu averse au risque de l'investisseur a une conséquence immédiate : la proportion d'actif risqué dans son portefeuille va être beaucoup plus grande. Au contraire, le caractère très averse au risque l'incitera à placer la grande majorité de ses ressources dans l'actif non risqué afin de ne pas subir un revers de fortune dans le cas où le cours de l'actif risqué se mettrait à chuter. C'est ce qu'illustre la figure 6.9.

Nous commençons par étudier le cas de l'investisseur très peu risquophobe au moyen de la figure 6.10 où ont été placés les graphiques de la richesse dans le cas où $d = 0.3$ et $d = 0$ (en haut) et ceux de la richesse dans le cas où $d = 0.3$ et la richesse corrigée dans le cas où $d = 0$ (en bas).

Les proportions d'actif risqué étant très importantes pour ce type de comportement de l'investisseur, cela entraîne un accroissement de l'amplitude des rendements journaliers et par conséquent cela démultiplie les erreurs de spécification du modèle. Dès lors, il n'est pas étonnant d'obtenir de telles différences de rendement de portefeuille : +39.06% dans le cas où $d = 0.3$ (avec un rendement journalier moyen de 0.0312), −2.72% dans le cas où $d = 0$ avec correction (avec un rendement journalier moyen de −0.0022). La différence est également énorme dans le cas où $d = 0$ si l'on apporte ou non la correction : −2.72% au lieu de +23.26%.

Par ailleurs, le hasard du tirage de chacune des trajectoires sous-jacentes entraîne que pour une période de gestion de 5 ans, le rendement global du portefeuille d'un investisseur assez moyennement averse au risque ($\gamma = -2$) et très peu averse au risque ($\gamma = 0.5$) est du même ordre de grandeur dans le cas où $d = 0.3$ avec respectivement +41.49% et +39.06%.

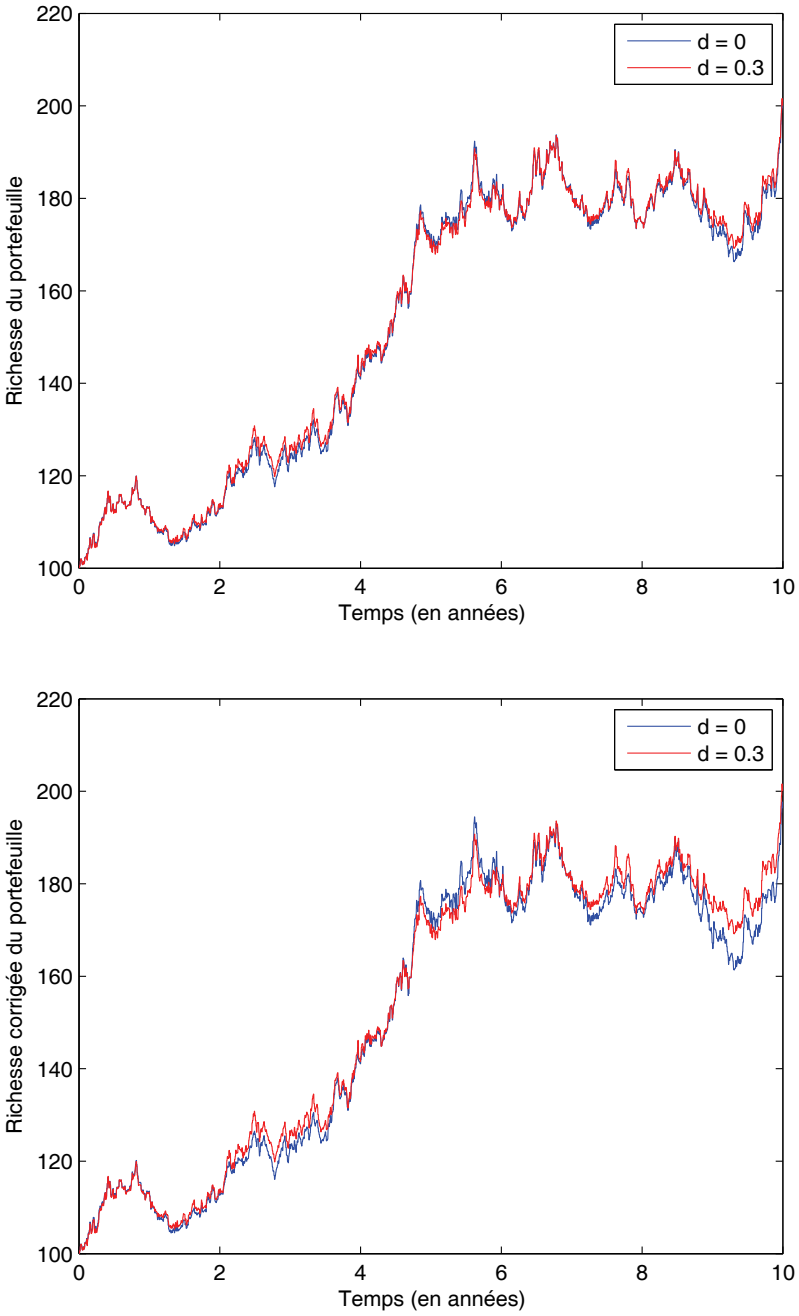

FIGURE 6.8 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) pour une période de gestion de 10 ans.

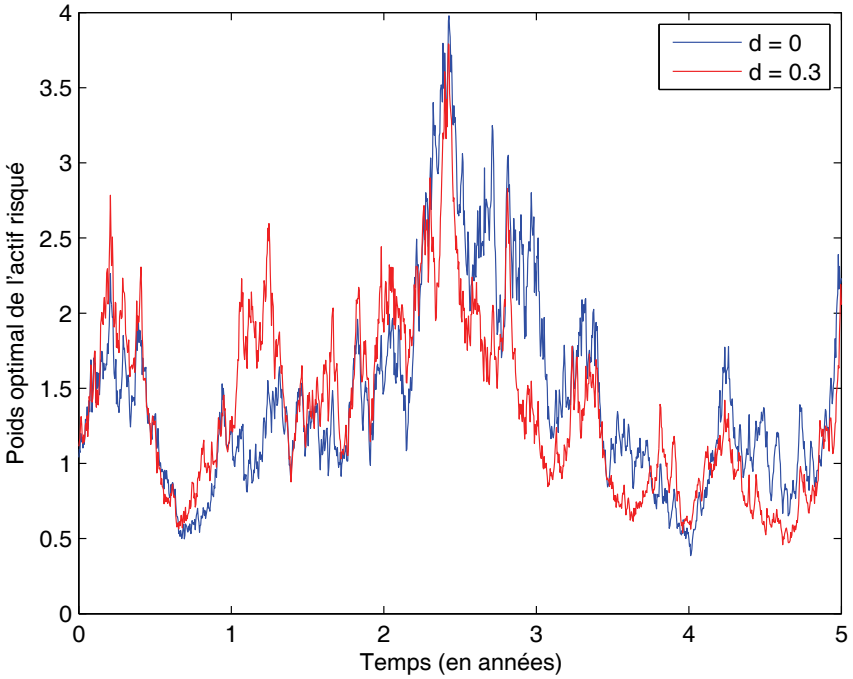

FIGURE 6.9 – Poids optimal de l'actif risqué pour un investisseur très averse au risque ($\gamma = -19$) et au contraire faiblement au risque ($\gamma = 0.5$), pour un horizon de gestion de 5 ans.

A l'opposée de la figure 6.10, la figure 6.11 traite du caractère très averse au risque de l'investisseur (le coefficient d'aversion vaut 20). Le graphique du haut représente la richesse du portefeuille obtenue pour $d = 0.3$ et $d = 0$ et le graphique du bas reprend la richesse du portefeuille pour $d = 0.3$ et superpose la richesse corrigée pour $d = 0$.

Compte tenu du caractère très averse au risque de l'investisseur, il n'est pas étonnant de constater que la trajectoire de la richesse du portefeuille, qu'elle soit corrigée ou pas, suit quasiment une droite. Au bout de 5 ans, le rendement est de 30.54% dans le cas où $d = 0.3$, et de 30.01% dans le cas où $d = 0$ avec correction (sans la correction le rendement est de 30.41%. Toutes ces valeurs sont très proches les unes des autres puisque du fait de sa grande aversion au risque, la proportion de richesse investie dans l'actif risqué sera très faible devant celle placée au taux sans risque. C'est la raison pour laquelle en plus d'être très proches les un des autres, les rendements sont aussi proches du rendement qu'aurait obtenu l'investisseur s'il avait placé la totalité de son capital au taux sans risque : 27.63%.

Moins l'investisseur va être averse au risque, plus il aura tendance à investir dans l'actif risqué. Par conséquent plus les différences au niveau des trajectoires de volatilité et du poids de l'actif risqué seront grandes et plus les conséquences des erreurs de spécification du modèle pénaliseront la rentabilité du portefeuille. En conclusion, moins l'investisseur est risquophobe, plus grand est son intérêt à avoir recours à un modèle de choix de portefeuille prenant en compte la mémoire longue de la volatilité de l'actif risqué.

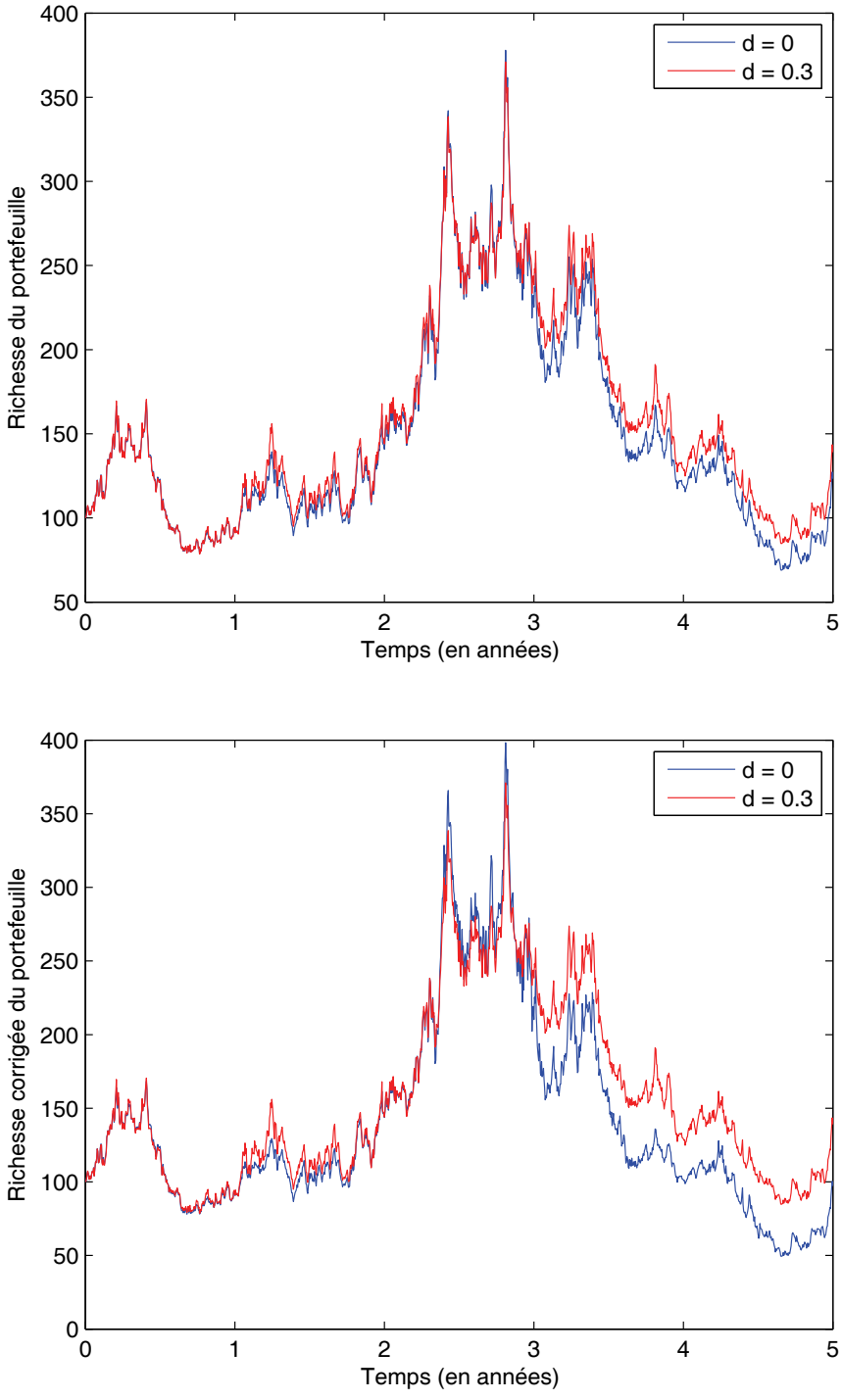

FIGURE 6.10 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) lorsque l'investisseur est faiblement averse au risque ($\gamma = 0.05$). La période de gestion est de 5 ans.

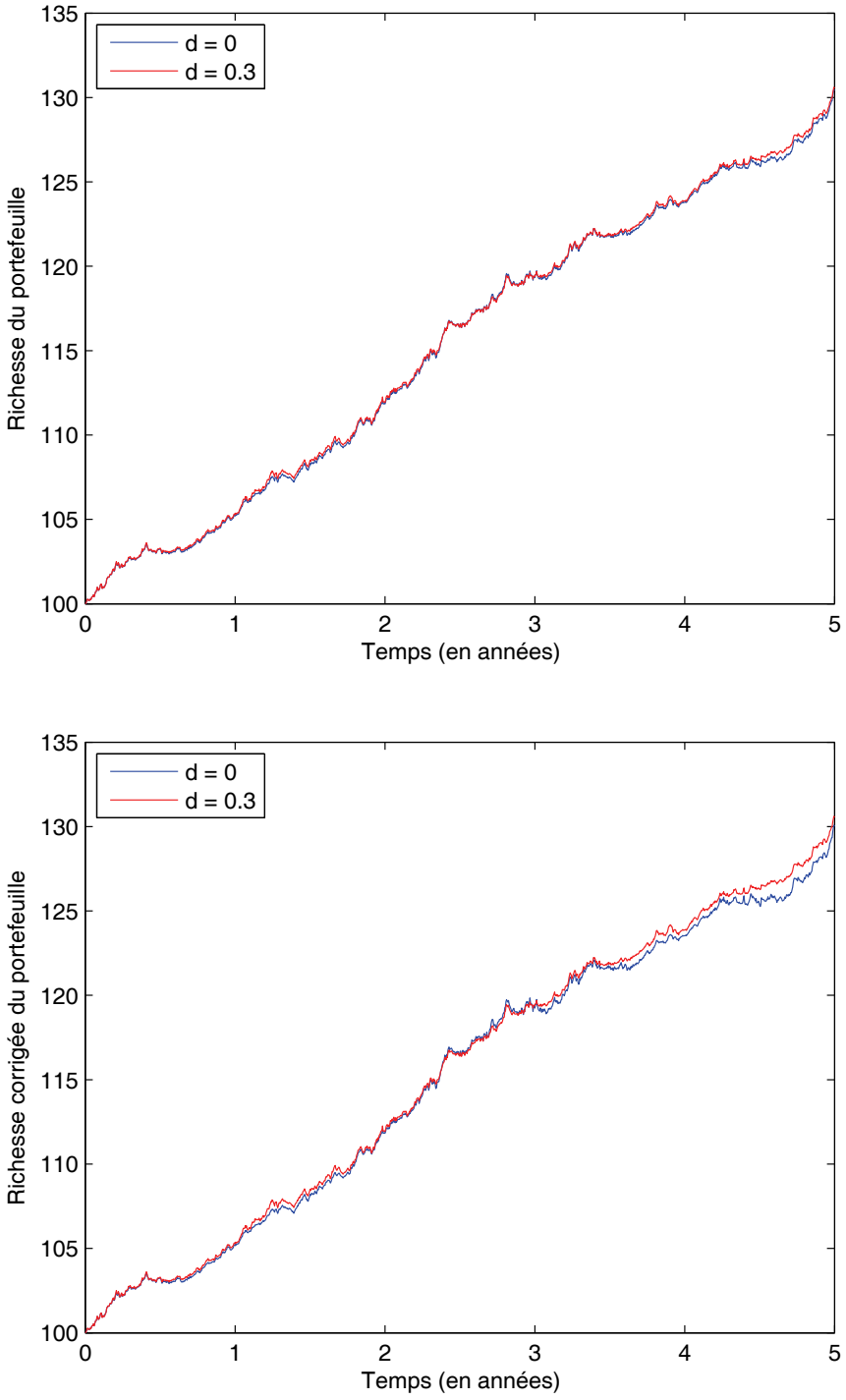

FIGURE 6.11 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) lorsque l'investisseur est fortement averse au risque ($\gamma = -19$). La période de gestion est de 5 ans.

Impact du retour à la moyenne du processus de la log-volatilité

A présent, nous nous intéressons au paramètre de retour à la moyenne. Dans l'étude du processus de la log-volatilité du début du chapitre, nous avons montré que ce paramètre est responsable de la mémoire courte. Jusqu'à présent, nous avons fixé ce retour à quatre mois. Que se passe-t-il si le retour à la moyenne est plus court, par exemple de l'ordre d'un mois ou d'une semaine ?

Pour commencer, observons l'impact d'une accélération du retour à la moyenne du processus de la log-volatilité sur la trajectoire de la volatilité. La figure 6.12 montre la volatilité réalisée journalière pour un retour à la moyenne fixé à un mois dans le cas où $d = 0.3$ et $d = 0$ (graphique du haut). Ce même graphique, mais pour un retour à la moyenne d'une semaine est placé en bas de cette figure. Nous constatons que l'accélération de la vitesse de retour à la moyenne a une conséquence visuelle sur les trajectoires de volatilité. En effet, plus κ est grand, plus la volatilité présente des fluctuations importantes.

Les figures 6.13 et 6.14 montrent l'évolution des richesses (graphique du haut) et richesses corrigées (graphique du bas) respectivement pour une retour à la moyenne d'un mois et d'une semaine.

Nous observons ici encore que quel que soit la vitesse de retour à la moyenne, le fait de corriger la richesse lorsque $d = 0$ entraîne une diminution du rendement du portefeuille. A noter également que le rendement est supérieur lorsque la mémoire longue est prise en compte. Par ailleurs, le rendement du portefeuille n'est que faiblement influencé par le changement de la vitesse de retour à la moyenne.

Impact de la volatilité de l'actif risqué

Pour finir, nous étudions l'importance de la prise en compte de la mémoire longue lorsque les marchés financiers sont soumis à de fortes turbulences comme c'est le cas par exemple lors d'une crise majeure. Dans la section 1.2.1 de la première partie de la thèse, nous avons appris que lorsqu'un tel évènement survenait, d'une part la volatilité était bien supérieure à sa valeur moyenne en absence de crise et d'autre part que ce phénomène ne durait pas aussi longtemps que la crise. Par conséquent, nous allons revenir à un horizon de gestion de 1 an et multiplier la volatilité moyenne par 2.5. La figure 6.15 montre les trajectoires de la volatilité, l'une à mémoire longue et l'autre à mémoire courte, (graphique du haut) et le poids optimal de l'actif risqué correspondant obtenu lors de la résolution numérique (graphique du bas).

Etant donné le caractère moyennement averse au risque de l'investisseur, le fait que la volatilité connaisse des niveaux très élevés a une conséquence directe sur le poids accordé à l'actif risqué : il est très faible. Par conséquence, les grandes différences que l'on observe entre les trajectoires de volatilité à mémoire longue et à mémoire courte (parfois de l'ordre de 40%), n'auront que peu d'impact sur le rendement du portefeuille. En effet, que la proportion d'action dans le portefeuille atteigne 2 ou 4%, cela ne fait finalement pas une très grande différence.

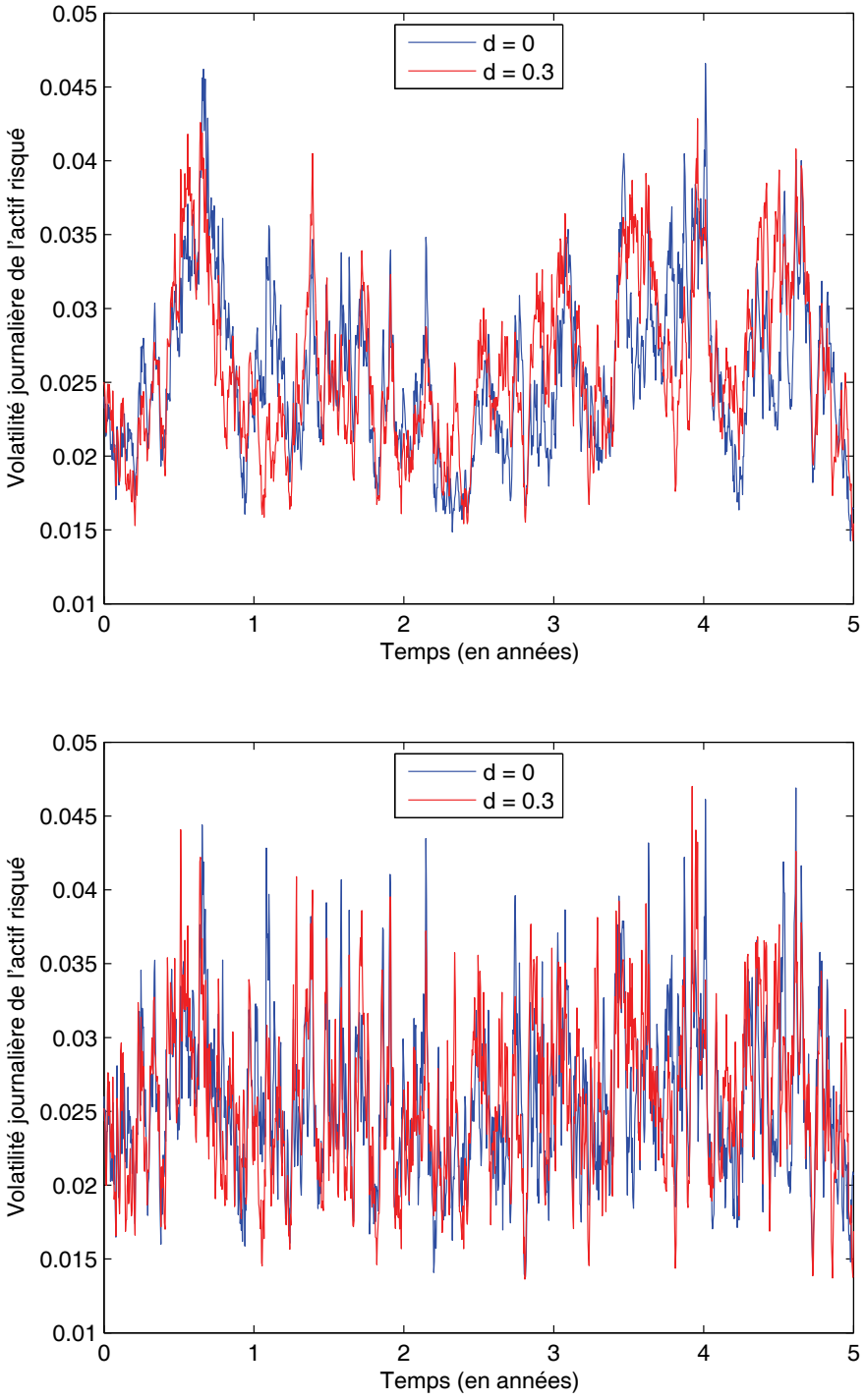

FIGURE 6.12 – Trajectoires de volatilité avec et sans mémoire longue pour un retour à la moyenne du processus de log-volatilité fixé à 1 mois (en haut) et fixé à 1 semaine (en bas), sur une période de 5 ans.

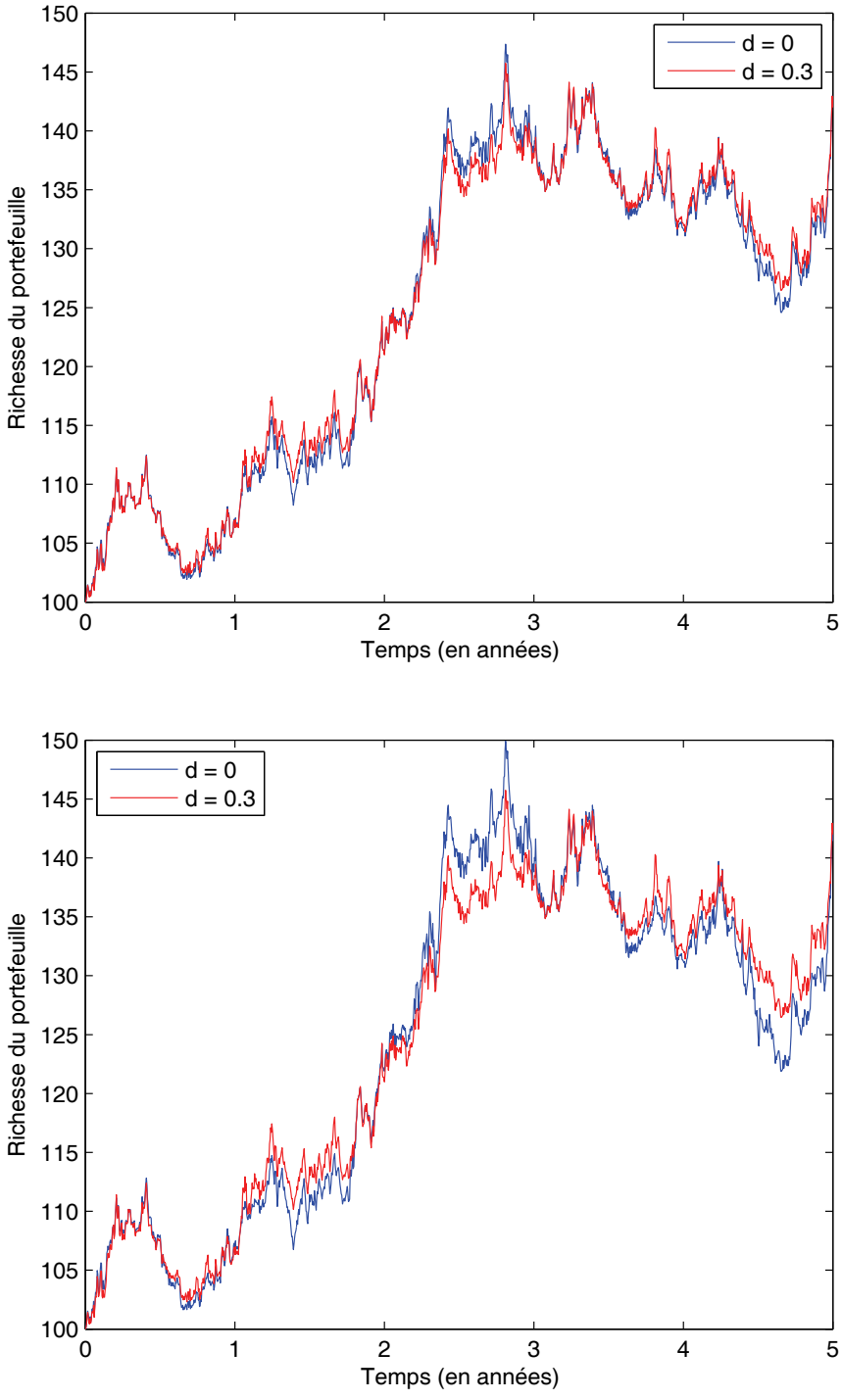

FIGURE 6.13 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) lorsque le retour à la moyenne du processus de log-volatilité est fixé à 1 mois. La période de gestion est de cinq ans.

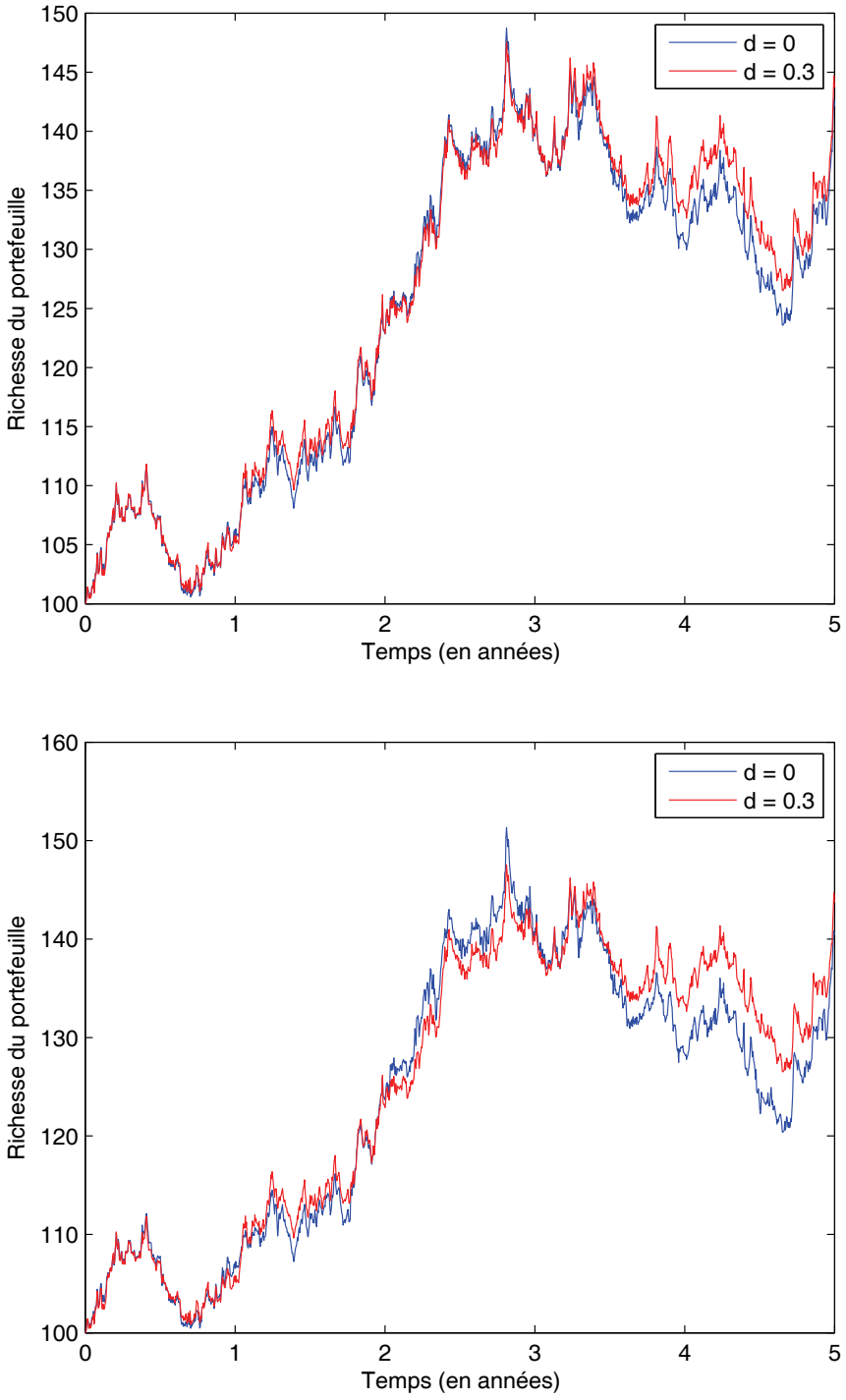

FIGURE 6.14 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) lorsque le retour à la moyenne du processus de log-volatilité est fixé à 1 semaine. La période de gestion est de cinq ans.

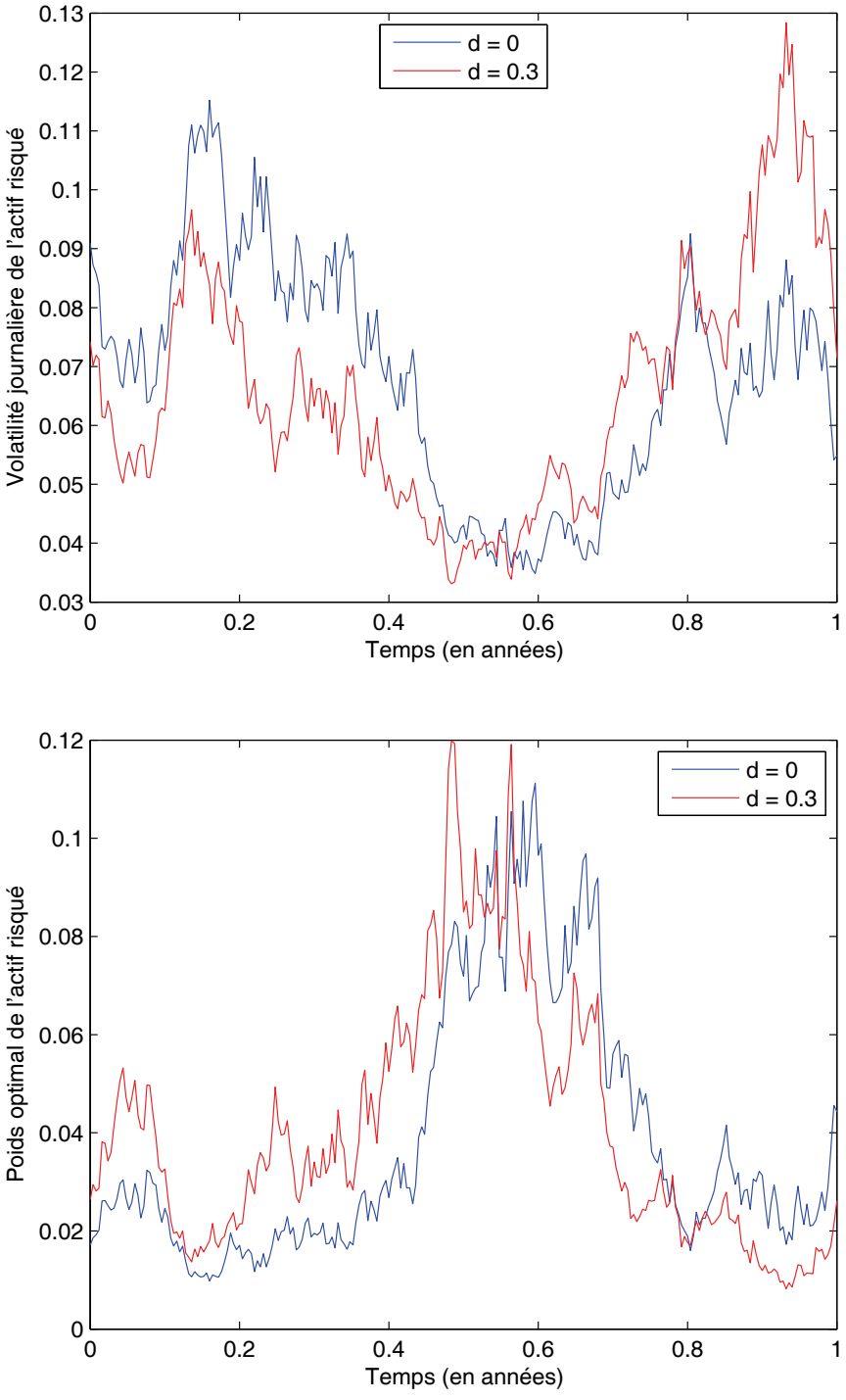

FIGURE 6.15 – Simulation d'une trajectoire de la volatilité de l'actif risque avec ($d = 0.3$) et sans ($d = 0$) mémoire longue et poids optimal de l'actif risqué correspondant obtenu sur une période de gestion de 1 an.

Afin de confirmer ces premières impressions, regardons à présent la richesse et la richesse corrigée du portefeuille que l'on peut déduire des résultats numériques de notre problème (voir figure 6.16). Nous constatons que le rendement moyen du portefeuille dans le cas où $d = 0.3$ est très proche du taux annuel sans risque : $+4.99\%$ contre $+5\%$. Le rendement corrigé du portefeuille dans le cas où $d = 0$ n'est pas très éloigné : $+3.11\%$. C'est en tout cas beaucoup plus proche que lorsque la volatilité moyenne est à niveau "normal" (voir figure 6.6). Comme nous l'avions pressenti au regard de la figure précédente, le fait que l'investisseur ait investi entre 90 et 98% de son capital dans l'actif non risqué et nivelé l'impact de la non prise en compte de la mémoire longue.

Conclusion

Dans ce chapitre, nous avons développé un modèle de choix de portefeuille en temps continu, composé de deux actifs (l'un risqué et l'autre non), avec une volatilité stochastique à mémoire longue. Nous avons modélisé cette mémoire longue au moyen d'un mouvement brownien fractionnaire. L'avantage d'un tel procédé est qu'il nous permet de contrôler la mémoire longue grâce à un paramètre, ce que d'autres modèles traitant du même sujet et rencontrés dans la littérature, tel que le modèle de Benth, Karlsen, et Reikvam (2003) et Barndorff-Nielsen et Shephard (2001), ne permettent pas de faire. L'autre aspect très intéressant de ce modèle réside dans le caractère non markovien du système économique décrit. Celui-ci nous a empêché d'employer la méthode classique par programme dynamique et l'équation d'Hamilton-Jacobi-Bellman. Par conséquent, nous avons procédé différemment et utilisé un critère de gestion nous permettant de simplifier le problème : la maximisation de l'utilité espérée de la richesse finale à un horizon de gestion fixé. Nous avons fourni un algorithme de résolution récursif original basé sur la technique de Monte Carlo et faisant intervenir l'approximation par la méthode des rectangles.

Enfin, en ce qui concerne l'étude destinée à montrer l'importance de la prise en compte de la mémoire longue de la volatilité de l'actif risqué dans la gestion de portefeuille, ses résultats montrent clairement que pour percevoir correctement le risque du portefeuille, il faut modéliser le plus fidèlement possible la volatilité de l'actif risqué. Or, parmi les faits stylisés les plus importants de la volatilité figure le phénomène de mémoire longue. Par conséquent, l'intégration de cette caractéristique dans la modélisation de la volatilité est primordiale. En particulier, l'analyse des différentes simulations obtenues durant notre étude nous enseigne que

- Les conséquences d'une mauvaise spécification du modèle, en terme de rendement de portefeuille sont d'autant plus pénalisantes, voire désastreuses que l'aversion au risque de l'investisseur est faible.
- L'impact de l'horizon de gestion n'est pas très concluant dans notre étude. Toutefois, il s'agit d'une seule simulation et nous pouvons penser qu'en règle générale, les erreurs s'accumulent et ont un effet boule de neige sur un horizon de gestion long.

Une fois encore, pour être tout à fait honnête, il est bon de rappeler que les résultats concernant le modèle avec volatilité stochastique à mémoire longue doivent être interprétés avec la plus grande précaution. Dans le cas présent, il ne s'agit que d'un faible nombre de

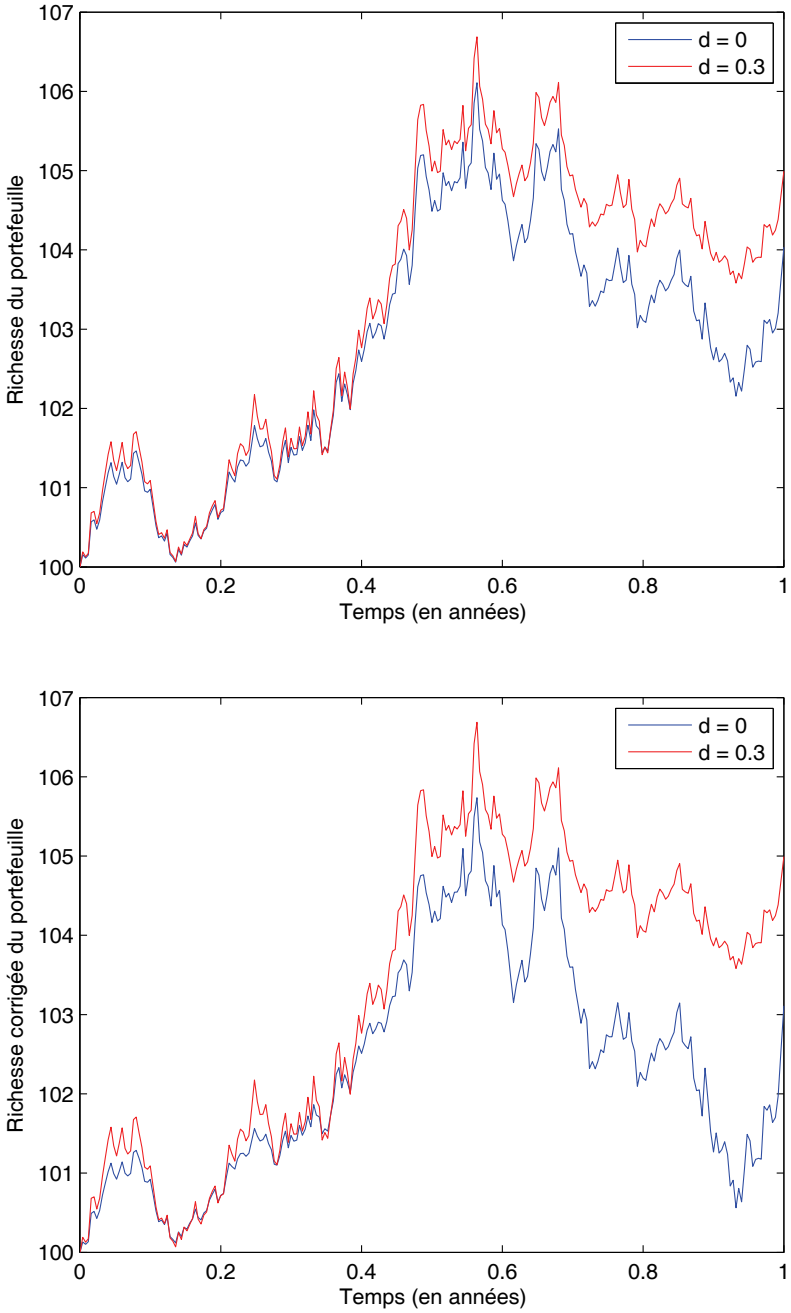

FIGURE 6.16 – Richesse du portefeuille lorsque $d = 0.3$ et $d = 0$ (en haut) et richesse corrigée du portefeuille (en bas) lorsque la volatilité a un niveau 2.5 fois plus élevé qu'en moyenne, synonyme de crise financière sur une période d'un an.

simulations. Nous ne pouvons donc pas tirer de conclusion générale concernant l'apport de la volatilité stochastique à mémoire longue sur les rendements du portefeuille. Pour pouvoir faire cela, il faudrait tirer un grand nombre de trajectoires de volatilité pour chaque hypothèse testée, un millier par exemple. Hélas, étant donné le temps de calcul nécessaire pour chaque simulation, cela n'est pas réalisable avec les ressources informatiques dont nous disposons.

Annexes

6.A Définition et théorèmes importants du mouvement brownien fractionnaire

Nous présentons dans cette annexe, tous les outils nécessaires à la résolution de notre problème d'allocation d'actifs.

Le mouvement brownien fractionnaire est un processus stochastique à mémoire longue défini par Mandelbrot et Wallis (1968) de la manière suivante :

Définition 16. *Un mouvement brownien fractionnaire de paramètre $d \in]-1/2, 1/2[$ est un processus gaussien réel centré, noté $\left\{B^d(t) : t \in \mathbb{R}\right\}$ défini sur un espace probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$ et vérifiant :*

(i) $B^d(0) = 0 \qquad \mathbb{P} - p.s.,$

(ii) $E\left[\left(B^d(t)\right)^2\right] = |t|^{2d+1} \quad \forall\, t \in \mathbb{R},$

(iii) Pour tout $0 \leq s < t$, les accroissements $B^d(t) - B^d(s)$ sont stationnaires.

Parmi les propriétés intéressantes du mouvement brownien, citons celles-ci

Propriété 1. Le mouvement brownien fractionnaire admet pour fonction d'autocovariance

$$\mathrm{E}\left[B^d(t)B^d(s)\right] = \frac{1}{2}\left(|t|^{2d+1} + |s|^{2d+1} - |t-s|^{2d+1}\right) \tag{6.A.1}$$

A présent, nous définissons la fonction $\phi(s,t)$:

Définition 17. *Pour tout $d \in]0, 1/2[$ et pour tout $s, t \in \mathbb{R}_+$, définissons*

$$\phi(s,t) = 2d\left(d + \frac{1}{2}\right)|s-t|^{2d-1}. \tag{6.A.2}$$

Puis, définissons la dérivée de Malliavin fractionnaire :

Définition 18. *Si $F : \Omega \to \mathbb{R}$ est une fonction donnée, la dérivée de Malliavin fractionnaire*[49] *de F en t (si elle existe) est donnée par*

$$D_t^{\phi} F = \int_{\mathbb{R}} D_r F \cdot \phi(r,t) dr, \tag{6.A.3}$$

où $D_r F$ est la dérivée classique de Malliavin.

49. Dans la littérature, nous la rencontrons sous le nom *Malliavin ϕ-derivative*.

Nous ne rappelons pas ce qu'est la dérivée de Malliavin [50].

Définissons par $\mathcal{L}_\phi^{1,2}(m)$ l'ensemble des processus (mesurables) $g(t,\omega) : \mathbb{R}_+ \times \Omega \longrightarrow \mathbb{R}$ tel que $D_s^\phi g(s)$ existe pour tout $s \in \mathbb{R}_+$ et

$$\| g \|^2_{\mathcal{L}_\phi^{1,2}} := \mathrm{E}\left[\int_{\mathbb{R}}\int_{\mathbb{R}} g(s)g(t)\phi(s,t)dsdt + \left(\int_{\mathbb{R}} D_s^\phi g(s)ds\right)^2\right] < \infty. \tag{6.A.4}$$

A présent, considérons le cas général d'un mouvement brownien fractionnaire multidimensionnel. Soit $d = (d_1, \cdots, d_m)$, avec $0 < d_j < 1/2$, $j = 1, \cdots, m$. Et soit $B^d(t) = (B_1^{d_1}(t), \cdots, B_m^{d_m}(t))$, $t \in \mathbb{R}$, un mouvement brownien fractionnaire de dimension m. Puis, définissons $X(t) = (X_1(t), \cdots, X_n(t))$ tel que

$$X_i(t,\omega) = \sum_{j=1}^m \int_0^t \sigma_{ij}(s,\omega) dB_j^{d_j}(s), \tag{6.A.5}$$

pour tout $i = 1, \cdots, n$ et avec $\sigma_{ij} \in \mathcal{L}_\phi^{1,2}(m)$.

Théorème 2 (Biagini, Hu, Oksendal, et Sulem (2002), théorème 2.3 adapté à notre problème)**.** Soit $f \in \mathcal{C}^{2,1}(\mathbb{R}^n \times \mathbb{R}_+)$ ayant des dérivées du second ordre bornées. Alors, pour $t \geq 0$,

$$\begin{aligned} f(X(t),t) = f(X(0),0) &+ \int_0^t \frac{\partial f}{\partial s}(X(s),s)ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X(s),s)dX_i(s) \\ &+ \sum_{i,j=1}^n \int_0^t \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X(s),s) \sum_{k=1}^m \sigma_{ik}(s) D_{k,s}^\phi X_j(s)\right] ds, \end{aligned} \tag{6.A.6}$$

où $D_{k,s}^\phi$ représente la dérivée de Malliavin fractionnaire par rapport à la k-ième composante du brownien X_j au temps s.

Théorème 3 (Dérivation des fonction composées)**.** Soit $f : \mathbb{R}^m \to \mathbb{R}$ une fonction donnée et $F_i : \Omega \to \mathbb{R}$ pour tout $i = 1, \cdots, m$. Alors

$$D_{k,s}^\phi f(F_1, \cdots, F_m) = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(F_1, \cdots, F_m) D_{k,s}^\phi F_i \tag{6.A.7}$$

Une application de la dérivée de Malliavin fractionnaire nous donne

Théorème 4 (Biagini, Hu, Oksendal, et Sulem (2002), théorème 2.4)**.** Soit

$$X(t) = \sum_{j=1}^m \int_0^t \sigma_j(r,w) dB_j^{d_j}(r), \tag{6.A.8}$$

avec $\sigma = (\sigma_1, \cdots, \sigma_m) \in \mathcal{L}_\phi^{1,2}(m)$. Alors

$$D_{k,s}^\phi X(t) = \sum_{j=1}^m \int_0^t D_{k,s}^\phi \sigma_j(r) dB_j^{d_j}(r) + \int_0^t \sigma_k(r)\phi_{d_k}(s,r)dr, \tag{6.A.9}$$

50. Le lecteur est invité à se référer à Malliavin et Thalmaier (2005) et Bell (2007).

pour $k = 1, \cdots, m$. En particulier, si σ_j est déterministe pour tout $j \in 1, \cdots, m$,

$$D^{\phi}_{k,s} X(t) = \int_0^t \sigma_k(r) \phi_{d_k}(s, r) dr. \tag{6.A.10}$$

Cette autre propriété, concernant les intégrales stochastiques, nous sera utile par la suite.

Propriété 2.

$$\mathrm{E}\left[\int_{\mathbb{R}} f(X(s), s) dB^d(s)\right] = 0, \tag{6.A.11}$$

$$\mathrm{E}\left[\int_0^{+\infty} f(s) dB^d(s) \int_0^{+\infty} g(u) dB^d(u)\right] = \int_0^{+\infty} \int_0^{+\infty} f(s) g(u) \phi(s, u) ds du, \tag{6.A.12}$$

Enfin, nous aurons besoin du théorème de Clark-Ocone dans le cadre classique. Ce théorème a aussi été adapté au cadre fractionnaire (théorème 4.14 (b), Hu et Oksendal (1999)), mais il ne nous sera pas nécessaire, donc nous nous contentons de l'énoncer dans le cadre classique.

A partir du théorème de représentation des martingales (voir Karatzas et Shreve (1991), annexe A), nous savons que pour toute fonction de brownien $G(B)$, il existe un processus $\mathcal{F}_t$-mesurable unique ψ tel que

$$G(B) = \mathrm{E}\left[G(B)\right] + \int_0^T \psi(t) dB(t). \tag{6.A.13}$$

Théorème 5 (Théorème de Clark-Ocone)**.** Pour toute fonction de brownien 6.A.13, si $G \in \mathbb{D}_{1,2}$ (voir Nualart (2006)), alors

$$\psi(t) := \mathrm{E}\left[D_t G(B) | \mathcal{F}_t\right], \tag{6.A.14}$$

où D_t est la dérivée de Malliavin.

6.B Démonstration de l'équation (6.6)

Le but est d'exprimer la dynamique de la volatilité $dV(t)$ à partir de l'équation différentielle stochastique de la log-volatilité,

$$d \ln V(t) = \kappa(\alpha - \ln V(t)) dt + \sigma \rho dB(t) + \sigma \sqrt{1 - \rho^2} dY^d(t). \tag{6.B.1}$$

Posons, pour plus de clarté, $X(t) = \ln V(t)$. Le processus $X(t)$ vérifie l'équation différentielle stochastique suivante :

$$dX(t) = \kappa(\alpha - X(t)) dt + \sigma \rho dB(t) + \sigma \sqrt{1 - \rho^2} dY^d(t). \tag{6.B.2}$$

Appliquons le théorème 2 à la fonction $f(X(t)) = \exp X(t)$ (ici, $n = 1$, $m = 2$) :

$$f(X(t)) = f(X(0)) + \int_0^t \frac{\partial f}{\partial x}(X(s), s)dX(s) + \int_0^t \frac{\partial^2 f}{\partial x^2}(X(s), s)\left[\sigma\rho D^\phi_{1,s}X(s) + \sigma\sqrt{1-\rho^2}D^\phi_{2,s}X(s)\right]ds,$$

d'où

$$V(t) = V(0) + \int_0^t V(s)\left[\kappa(\alpha - \ln V(s))ds + \sigma\rho dB(s) + \sigma\sqrt{1-\rho^2}dY^d(s)\right] + \sigma\rho\int_0^t V(s)D^\phi_{1,s}\ln V(s)ds + \sigma\sqrt{1-\rho^2}\int_0^t V(s)D^\phi_{2,s}\ln V(s)ds. \quad (6.B.3)$$

Il nous faut donc calculer $D^\phi_{1,s}\ln V(s)$ et $D^\phi_{2,s}\ln V(s)$. Par conséquent, nous devons commencer par déterminer l'expression de $\ln V(t)$. Appliquons le lemme d'Itō fractionnaire (théorème 2) à la fonction $f(X(t), t) = X(t)\mathrm{e}^{\kappa t}$:

$$f(X(t),t) = f(X(0),0) + \int_0^t \frac{\partial f}{\partial s}(X(s),s)ds + \int_0^t \frac{\partial f}{\partial x}(X(s),s)dX(s) + \int_0^t \underbrace{\frac{\partial^2 f}{\partial x^2}(X(s),s)}_{=0}\sigma(s)D^\phi_s X(s)ds$$

$$X(t)\mathrm{e}^{\kappa t} = X(0) + \int_0^t \kappa X(s)\mathrm{e}^{\kappa s}ds + \int_0^t \mathrm{e}^{\kappa s}\left[\kappa(\alpha - X(s)ds + \sigma\rho\int_0^t \mathrm{e}^{\kappa s}dB(s) + \sigma\sqrt{1-\rho^2}\int_0^t \mathrm{e}^{\kappa s}dY^d(s)\right]$$

$$\ln V(t) = \alpha + \mathrm{e}^{-\kappa t}(\ln V(0) - \alpha) + \sigma\rho\int_0^t \mathrm{e}^{\kappa(u-t)}dB(u) + \sigma\sqrt{1-\rho^2}\int_0^t \mathrm{e}^{\kappa(u-t)}dY^d(u). \quad (6.B.4)$$

A présent, en appliquant le théorème 4 à l'équation (6.B.4) [51], nous obtenons les expressions de $D^\phi_{1,s}\ln V(s)$ et $D^\phi_{2,s}\ln V(s)$:

$$\begin{aligned} D^\phi_{1,s}\ln V(t) &= \int_0^t \sigma_1(u)\phi_{d_1}(s,u)du, \\ &= \sigma\rho\int_0^t \mathrm{e}^{\kappa(u-t)}\phi_{d_1}(s,u)du. \end{aligned}$$

Nous rappelons que

$$\phi_d(s,u) = \begin{cases} \frac{1}{2}\delta(s-u) & \text{si } d = 0, \\ 2d\left(d+\frac{1}{2}\right)|s-u|^{2d-1} & \text{si } 0 < d < 1/2, \end{cases} \quad (6.B.5)$$

51. Dans notre cas, $\sigma_1(u) = \sigma\rho\mathrm{e}^{\kappa(u-t)}$ et $\sigma_2(u) = \sigma\sqrt{1-\rho^2}\mathrm{e}^{\kappa(u-t)}$ sont déterministes.

avec $\delta(x)$ est la fonction de Dirac. Pour simplifier la notation, introduisons la fonction

$$f(s,t) = \begin{cases} \mathbb{1}_{\{s\leq t\}} & \text{si } d = 0, \\ \int_0^t \mathrm{e}^{\kappa(u-s)} 2d|u-s|^{2d-1} du & \text{si } 0 < d < 1/2. \end{cases} \tag{6.B.6}$$

La dérivée de Malliavin s'écrit alors

$$D^{\phi}_{1,s} \ln V(t) = \left(d_1 + \frac{1}{2}\right) \sigma\rho \mathrm{e}^{\kappa(s-t)} f(s,t), \tag{6.B.7}$$

Puisque $d_1 = 0$, alors

$$D^{\phi}_{1,s} \ln V(t) = \frac{1}{2}\sigma\rho \mathrm{e}^{\kappa(s-t)} \mathbb{1}_{\{s\leq t\}}. \tag{6.B.8}$$

On en déduit que

$$D^{\phi}_{1,s} \ln V(s) = \frac{1}{2}\sigma\rho. \tag{6.B.9}$$

D'autre part,

$$D^{\phi}_{2,s} \ln V(t) = \int_0^t \sigma_2(u)\phi_{d_2}(s,u)du. \tag{6.B.10}$$

Puisque $d_2 = d$, alors

$$D^{\phi}_{2,s} \ln V(t) = \left(d + \frac{1}{2}\right) \sigma\sqrt{1-\rho^2} \mathrm{e}^{\kappa(s-t)} f(s,t),$$

On en déduit que

$$D^{\phi}_{2,s} \ln V(s) = \left(d + \frac{1}{2}\right) \sigma\sqrt{1-\rho^2} f(s), \tag{6.B.11}$$

avec la notation $f(s) = f(s,s)$.

En remplaçant $D^{\phi}_{1,s} \ln V(s)$ et $D^{\phi}_{2,s} \ln V(s)$ par leurs expressions respectives (6.B.9) et (6.B.11) dans l'équation (6.B.3), on obtient

$$\begin{aligned} V(t) = V(0) + \int_0^t V(s)\kappa(\alpha - \ln V(s))ds + \sigma\rho \int_0^t V(s) dB(s) \\ + \sigma\sqrt{1-\rho^2} \int_0^t V(s) dY^d(s) + \frac{1}{2}\sigma^2\rho^2 \int_0^t V(s)ds \\ + \left(d + \frac{1}{2}\right) \sigma^2(1-\rho^2) \int_0^t V(s)f(s)ds. \end{aligned} \tag{6.B.12}$$

En notation différentielle, cela donne

$$dV(t) = \left\{ \sigma^2 \left[\frac{\rho^2}{2} + \left(d + \frac{1}{2}\right)(1-\rho^2) f(t) \right] + \kappa(\alpha - \ln V(t)) \right\} V(t)dt + \sigma V(t) dZ^d(t), \tag{6.B.13}$$

□

6.C Démonstration de la section 6.1.2

Nous rappelons que l'objectif de l'investisseur est de maximiser l'utilité espérée de sa richesse finale, soit

$$\max_{w(t)} \mathrm{E}\left[H(W(T))\right]. \tag{6.C.1}$$

Le processus $w(t)^*$ est optimal si et seulement si, quel que soit le processus $\mathcal{F}_t$-adapté $\delta(t)$ et quel que soit ε

$$\mathrm{E}\left[H(W(T,w^*))\right] \geq \mathrm{E}\left[H(W(T,w^*+\varepsilon\delta))\right], \tag{6.C.2}$$

ce qui se traduit par le fait que la fonction

$$\varepsilon \longrightarrow \mathrm{E}\left[H(W(T,w^*+\varepsilon\delta))\right] \tag{6.C.3}$$

admet un maximum en $\varepsilon = 0$. Sous des conditions de régularité raisonnables, que nous ne cherchons pas à exprimer, la condition du premier ordre s'exprime comme

$$\partial_\varepsilon \mathrm{E}\left[H(W(T,w^*+\varepsilon\delta))\right]|_{\varepsilon=0} = 0, \tag{6.C.4}$$

soit

$$\mathrm{E}\left[H'(W(T))W(T)\left(\int_0^T [(\mu-r)-w^*(s)V(s)^2]\delta(s)ds + \int_0^T V(s)\delta(s)dB(s)\right)\right] = 0. \tag{6.C.5}$$

Focalisons-nous sur le dernier terme de l'équation ci-dessus :

$$\mathrm{E}\left[H'(W(T))W(T)\int_0^T V(s)\delta(s)dB(s)\right]. \tag{6.C.6}$$

En vertu du théorème de Clark-Ocone (voir Oksendal (1997), théorème 5.8 de Clark-Ocone), par dualité,

$$\mathrm{E}\left[H'(W(T))W(T)\int_0^T V(s)\delta(s)dB(s)\right] = \mathrm{E}\left[\int_0^T D_s[H'(W(T))W(T)]V(s)\delta(s)ds\right], \tag{6.C.7}$$

$$= \mathrm{E}\left[\int_0^T (H'(W(T)) + H''(W(T))W(T))D_sW(T)\cdot V(s)\delta(s)ds\right], \tag{6.C.8}$$

où $D_sW(T)$ désigne la dérivée de Malliavin de $W(T)$ par rapport au Brownien B.
En reportant ceci dans l'équation (6.C.5), il vient

$$\mathrm{E}\left[\int_0^T ([(\mu-r)-w^*(s)V(s)^2]H'(W(T))W(T) + V(s)(H'(W(T)) + H''(W(T))W(T))D_sW(T))\cdot\delta(s)ds\right] = 0, \tag{6.C.9}$$

quel que soit le processus δ $\mathcal{F}_t$-adapté. On en déduit donc le terme entre parenthèses dans l'intégrale :

$$\Big[(\mu-r)-w^*(t)V(t)^2\Big]\mathrm{E}\left[H'(W(T))W(T)|\mathcal{F}_t\right] + V(t)\cdot\mathrm{E}\left[(H'(W(T)) + H''(W(T))W(T))D_tW(T)|\mathcal{F}_t\right] = 0. \tag{6.C.10}$$

Il reste à calculer $D_tW(T)$. D'après l'expression intégrale de la richesse (6.12) et en vertu du théorème de dérivation des fonctions composées appliqué à la dérivée de Malliavin classique,

$$D_tW(T) = W(T)D_t \cdot \left[\int_0^T \left[(\mu - r)w^*(s) + r - \frac{1}{2}w^*(s)^2V(s)^2\right]ds + \int_0^T w^*(s)V(s)dB(s)\right], \tag{6.C.11}$$

$$= W(T)\left[-\int_0^T w^*(s)^2V(s)D_tV(s)ds + \int_0^T w^*(s)D_tV(s)dB_s + w^*(t)V(t)\right]. \tag{6.C.12}$$

A partir de l'expression de $\ln V(s)$. En appliquant le théorème de dérivation des fonctions composées et en remarquant que $\sigma\rho e^{-\kappa(u-t}$ est une fonction déterministe,

$$D_tV(s) = \sigma\rho V(s)e^{-\kappa(s-t)}\mathbb{1}_{s\geq t}, \tag{6.C.13}$$

donc

$$D_tW(T) = W(T)\left[-\int_t^T \sigma\rho w^*(s)^2V(s)^2e^{-\kappa(s-t)}ds \right.$$
$$\left. + \int_t^T \sigma\rho w^*(s)V(s)e^{-\kappa(s-t)}dB(s) + w^*(t)V(t)\right]. \tag{6.C.14}$$

Donc en substituant l'expression de $D_tW(T)$ de l'équation (6.C.14) dans l'équation (6.C.10), nous obtenons

$$\begin{aligned}
&\mathrm{E}\left[(H'(W(T)) + H''(W(T))W(T))D_tW(T)|\mathcal{F}_t\right] = \\
&\qquad w^*(t)V(t)\cdot\mathrm{E}\left[(H'(W(T))W(T) + H''(W(T))W(T)^2)|\mathcal{F}_t\right] \\
&\quad - \mathrm{E}\left[(H'(W(T))W(T) + H''(W(T))W(T)^2)\cdot\int_t^T \sigma\rho w^*(s)^2V(s)^2e^{-\kappa(s-t)}ds|\mathcal{F}_t\right] \\
&+ \mathrm{E}\left[(H'(W(T))W(T) + H''(W(T))W(T)^2)\cdot\int_t^T \sigma\rho w^*(s)V(s)e^{-\kappa(s-t)}dB(s)|\mathcal{F}_t\right].
\end{aligned} \tag{6.C.15}$$

Dans le cas d'une fonction d'utilité CRRA, la fonction héritage $H(W) = \frac{1}{\gamma}W^\gamma$, et il vient

$$\begin{aligned}
w^*(t) &= \frac{\mu - r}{(1-\gamma)V(t)^2} - \frac{\sigma\rho\gamma}{(1-\gamma)V(t)} \\
&\quad\times \frac{\mathrm{E}\left[W(T)^\gamma\left(\int_t^T w^*(s)^2V(s)^2e^{-\kappa(s-t)}ds - \int_t^T w^*(s)V(s)e^{-\kappa(s-t)}dB(s)\right)|\mathcal{F}_t\right]}{\mathrm{E}\left[W(T)^\gamma|\mathcal{F}_t\right]}.
\end{aligned} \tag{6.C.16}$$

Nous constatons qu'en l'absence de corrélation ($\rho = 0$), l'expression de $w^*(t)$ se simplifie considérablement :

$$w^*(t) = \frac{\mu - r}{(1-\gamma)V(t)^2}. \tag{6.C.17}$$

□

6.D Démonstration de l'équation (6.20)

L'autocovariance de la log-volatilité $\gamma(\tau) = \mathrm{Cov}(\ln V_t, \ln V_{t+\tau})$ s'écrit

$$\gamma(\tau) = \mathrm{E}\left[\left(\sigma\rho\int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)}dB_s + \sigma\sqrt{1-\rho^2}\int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)}dY_s^d\right) \times \left(\sigma\rho\int_{-\infty}^{t+\tau} \mathrm{e}^{\kappa(s-t-\tau)}dB_s + \sigma\sqrt{1-\rho^2}\int_{-\infty}^{t+\tau} \mathrm{e}^{\kappa(s-t-\tau)}dY_s^d\right)\right]. \quad (6.D.1)$$

Par définition, les incréments d'un mouvement brownien sont indépendants, donc

$$\mathrm{E}\left[\int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)}dB_s \int_{-\infty}^{t+\tau} \mathrm{e}^{\kappa(s-t-\tau)}dB_s\right] = \mathrm{e}^{-\kappa\tau}\cdot\mathrm{E}\left[\left(\int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)}dB_s\right)^2\right]. \quad (6.D.2)$$

Par ailleurs, par hypothèse les incréments du mouvement brownien B_s et ceux du mouvement brownien fractionnaire Y_s^d sont indépendants, d'où

$$\gamma(\tau) = \sigma^2\rho^2\mathrm{e}^{-\kappa\tau}\cdot\mathrm{E}\left[\left(\int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)}dB_s\right)^2\right] + \sigma^2(1-\rho^2)\cdot\mathrm{E}\left[\int_{-\infty}^{t} \mathrm{e}^{\kappa(s-t)}dY_u^d \int_{-\infty}^{t+\tau} \mathrm{e}^{\kappa(s-t-\tau)}dY_v^d\right], \quad (6.D.3)$$

ou encore

$$\gamma(\tau) = \sigma^2\mathrm{e}^{-\kappa\tau}\left\{\rho^2\frac{1}{2\kappa} + (1-\rho^2)V(\tau)\right\}, \quad (6.D.4)$$

avec

$$V(\tau) = 2d\left(d+\frac{1}{2}\right)\int_{-\infty}^{t}\int_{-\infty}^{t+\tau} \mathrm{e}^{\kappa(u+v-2t)}|u-v|^{2d-1}dudv. \quad (6.D.5)$$

Etudions le comportement de $\gamma(\tau)$.

En effectuant les changements de variable $x = -u+t+\tau$ et $y = -v+t+\tau$, nous obtenons

$$V(\tau) = 2d\left(d+\frac{1}{2}\right)\mathrm{e}^{2\kappa\tau}(A+B), \quad (6.D.6)$$

avec

$$A = \int_{\tau}^{+\infty}\int_{0}^{y} \mathrm{e}^{-\kappa(x+y)}(y-x)^{2d-1}dxdy, \quad (6.D.7)$$

$$B = \int_{\tau}^{+\infty}\int_{y}^{+\infty} \mathrm{e}^{-\kappa(x+y)}(x-y)^{2d-1}dxdy. \quad (6.D.8)$$

Commençons par calculer B. En effectuant successivement les changements de variable $u = x-y$ puis $v = \kappa u$, nous obtenons

$$B = \frac{\mathrm{e}^{-2\kappa\tau}}{2\kappa^{2d+1}}\Gamma(2d). \quad (6.D.9)$$

A présent intéressons nous au calcul de A. Nous effectuons successivement les changements de variable $u = y - x$, $v = \kappa u$, $u = \kappa y$ et $t = v/u$. Nous obtenons

$$A = \frac{1}{\kappa^{2d+1}} \int_{\kappa\tau}^{+\infty} \mathrm{e}^{-2u} u^{2d} \int_0^1 \mathrm{e}^{tu} t^{2d-1} dt du, \tag{6.D.10}$$

qui, d'après Abramowitz et Stegun (1965, (13.2.1)), peut s'écrire

$$A = \frac{1}{2d\kappa^{2d+1}} \int_{\kappa\tau}^{+\infty} \mathrm{e}^{-2u} u^{2d} M(2d, 2d+1, u) du, \tag{6.D.11}$$

où $M(\cdot,\cdot,\cdot)$ est la fonction Kummer.

D'après Abramowitz et Stegun (1965, (13.1.4)),

$$M(a,b,z) \underset{\tau\to\infty}{\simeq} \frac{\Gamma(b)}{\Gamma(a)} \mathrm{e}^z z^{-1}. \tag{6.D.12}$$

D'où, pour τ grand, nous avons

$$A \simeq \frac{1}{\kappa^{2d+1}} \int_{\kappa\tau}^{+\infty} \mathrm{e}^{-u} u^{2d-1} du. \tag{6.D.13}$$

En effectuant le changement de variable $t = u/\kappa\tau$ et en utilisant la formule (13.2.6) de Abramowitz et Stegun (1965), nous aboutissons, pour τ grand, à

$$A \simeq \frac{\tau^{2d}}{\kappa} \mathrm{e}^{-\kappa\tau} U(1, 2d+1, \kappa\tau), \tag{6.D.14}$$

où $U(\cdot,\cdot,\cdot)$ est définie par

$$U(a,b,z) = \frac{\pi}{\sin \pi b} \left\{ \frac{M(a,b,z)}{\Gamma(a-b+1)\Gamma(b)} - z^{1-b} \frac{M(a-b+1, 2-b, z)}{\Gamma(a)\Gamma(2-b)} \right\}, \tag{6.D.15}$$

et possède, d'après Abramowitz et Stegun (1965, (13.1.8)) la propriété suivante

$$U(a,b,z) \underset{z\to\infty}{\simeq} z^{-a}. \tag{6.D.16}$$

Par conséquent, pour τ grand,

$$A \simeq \frac{\tau^{2d-1}}{\kappa^2} \mathrm{e}^{-\kappa\tau}. \tag{6.D.17}$$

En conclusion, en substituant A et B dans l'équation (6.D.6) par leurs expressions respectives (6.D.17) et (6.D.9), pour obtenons, pour τ grand :

$$\gamma(\tau) \underset{\tau\to\infty}{\simeq} \sigma^2 \frac{\mathrm{e}^{-\kappa\tau}}{2\kappa} + \sigma^2(1-\rho^2) \left[\frac{\mathrm{e}^{-\kappa\tau}}{2\kappa} \left(\frac{\Gamma(2d+1)}{2\kappa^{2d}} - 1 \right) + 2d(2d+1) \frac{\tau^{2d-1}}{2\kappa^2} \right]. \tag{6.D.18}$$

□

6.E Démonstration de la proposition 5

.

A partir de l'équation (6.C.16), nous voulons obtenir une valeur approchée du poids optimal de l'actif risqué. Nous discrétisons le temps $t = [0, \tau, 2\tau, \cdots, T-\tau, T]$, avec $\tau = \frac{T}{N}$ où N est le nombre de points de la discrétisation. Nous pouvons alors écrire

$$w^*_{T-k\tau} = \frac{\mu - r}{(1-\gamma)V^2_{T-k\tau}} - \frac{\sigma\rho\gamma}{(1-\gamma)V_{T-k\tau}} \frac{B_{T-k\tau}}{A_{T-k\tau}}, \tag{6.E.1}$$

avec

$$A_{T-k\tau} = \mathrm{E}\left[W_T^\gamma | \mathcal{F}_{T-k\tau}\right], \tag{6.E.2}$$

$$B_{T-k\tau} = \mathrm{E}\left[W_T^\gamma \left(\int_{T-k\tau}^{T} w_s^2 V_s^2 \mathrm{e}^{-\kappa(s-T+k\tau)} ds - \int_{T-k\tau}^{T} w_s V_s \mathrm{e}^{-\kappa(s-T+k\tau)} dB_s\right) | \mathcal{F}_{T-k\tau}\right], \tag{6.E.3}$$

où

$$W_T = W_0 \exp\left\{\int_0^T \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \int_0^T w_s V_s dB_s\right\}. \tag{6.E.4}$$

Les coefficients $A_{T-k\tau}$ et $B_{T-k\tau}$ ne sont pas exploitables en l'état. Pour y remédier, remplaçons W_T dans les équations (6.E.2) et (6.E.3) par son expression dans l'équation (6.E.4) et développons.

6.E.1 Nouvelle formulation du poids optimal de l'actif risqué

Nous voulons reformuler $A_{T-k\tau}$ et $B_{T-k\tau}$ dans (6.E.2) et (6.E.3) pour pouvoir en donner une valeur approchée. Commençons par nous intéresser à $A_{T-k\tau}$:

$$\begin{aligned}
A_{T-k\tau} &= \mathrm{E}\left[W^\gamma_{T-k\tau} \exp\left\{\gamma \int_{T-k\tau}^{T} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \gamma \int_{T-k\tau}^{T} w_s V_s dB_s\right\} | \mathcal{F}_{T-k\tau}\right] \\
&= W^\gamma_{T-k\tau} \mathrm{E}\left[\exp\left\{\gamma \int_{T-k\tau}^{T-(k-1)\tau} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \gamma \int_{T-k\tau}^{T-(k-1)\tau} w_s V_s dB_s\right\}\right. \\
&\qquad \left. \times \exp\left\{\gamma \int_{T-(k-1)\tau}^{T} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \gamma \int_{T-(k-1)\tau}^{T} w_s V_s dB_s\right\} | \mathcal{F}_{T-k\tau}\right]
\end{aligned}$$

Puisque si $\mathcal{H} \subset \mathcal{G}$, alors $\mathrm{E}[X|\mathcal{H}] = \mathrm{E}[\mathrm{E}[X|\mathcal{G}]|\mathcal{H}]$ ps, on peut donc écrire

$$\begin{aligned}
A_{T-k\tau} = W_{T-k\tau}^{\gamma} \mathrm{E}\Bigg[& \exp\left\{\gamma \int_{T-k\tau}^{T-(k-1)\tau} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \gamma \int_{T-k\tau}^{T-(k-1)\tau} w_s V_s dB_s\right\} \\
& \times \mathrm{E}\Bigg[\exp\left\{\gamma \int_{T-(k-1)\tau}^{T} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds \right. \\
& \qquad \left. + \ \gamma \int_{T-(k-1)\tau}^{T} w_s V_s dB_s\right\} |\mathcal{F}_{T-(k-1)\tau}\Bigg] |\mathcal{F}_{T-k\tau}\Bigg] \\
\approx W_{T-k\tau}^{\gamma} \mathrm{E}\Bigg[& \exp\left\{\gamma \left[(\mu - r)w_{T-k\tau} + r - \frac{1}{2}w_{T-k\tau}^2 V_{T-k\tau}^2\right]\tau\right\} \\
& \times \exp\left\{\gamma w_{T-k\tau} V_{T-k\tau} \tilde{\xi}_{T-k\tau}\right\} \tilde{A}_{T-(k-1)\tau} |\mathcal{F}_{T-k\tau}\Bigg],
\end{aligned}$$

où $\tilde{A}_t = \frac{A_t}{W_t^{\gamma}}$, $\tilde{\xi}_t = \int_t^{t+\tau} dB_s$ est indépendant de $\mathcal{F}_t$ pour tout $t \in [0, T]$. Or, w_t et V_t sont $\mathcal{F}_t$-mesurables pour tout $t \in [0, T]$. Or, puisque si Y est $\mathcal{G}$-mesurable, alors $\mathrm{E}[XY|\mathcal{G}] = Y\mathrm{E}[X|\mathcal{G}]$ et en particulier $\mathrm{E}[Y|\mathcal{G}] = Y$, on obtient finalement

$$\begin{aligned}
A_{T-k\tau} \approx W_{T-k\tau}^{\gamma} & \exp\left\{\gamma \left[(\mu - r)w_{T-k\tau} + r - \frac{1}{2}w_{T-k\tau}^2 V_{T-k\tau}^2\right]\tau\right\} \\
& \times \mathrm{E}\left[\tilde{A}_{T-(k-1)\tau} \exp\left\{\gamma w_{T-k\tau} V_{T-k\tau} \tilde{\xi}_{T-k\tau}\right\} |\mathcal{F}_{T-k\tau}\right], \quad (6.E.5)
\end{aligned}$$

A présent, concentrons nous sur $B_{T-k\tau}$

$$\begin{aligned}
B_{T-k\tau} = \mathrm{E}\Bigg[& W_{T-k\tau}^{\gamma} \exp\left\{\gamma \int_{T-k\tau}^{T-(k-1)\tau} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \gamma \int_{T-k\tau}^{T-(k-1)\tau} w_s V_s dB_s\right\} \\
& \times \exp\left\{\gamma \int_{T-(k-1)\tau}^{T} \left[(\mu - r)w_s + r - \frac{1}{2}w_s^2 V_s^2\right] ds + \gamma \int_{T-(k-1)\tau}^{T} w_s V_s dB_s\right\} \\
& \times \Bigg(\int_{T-(k-1)\tau}^{T} w_s^2 V_s^2 \mathrm{e}^{-\kappa(s-T+k\tau)} ds - \int_{T-(k-1)\tau}^{T} w_s V_s \mathrm{e}^{-\kappa(s-T+k\tau)} dB_s \\
& \quad + \int_{T-k\tau}^{T-(k-1)\tau} w_s^2 V_s^2 \mathrm{e}^{-\kappa(s-T+k\tau)} ds - \int_{T-k\tau}^{T-(k-1)\tau} w_s V_s \mathrm{e}^{-\kappa(s-T+k\tau)} dB_s\Bigg) |\mathcal{F}_{T-k\tau}\Bigg]
\end{aligned}$$

$$
\begin{aligned}
&= W_{T-k\tau}^{\gamma}\mathrm{E}\left[\exp\left\{\gamma\int_{T-k\tau}^{T-(k-1)\tau}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-k\tau}^{T-(k-1)\tau}w_sV_sdB_s\right\}\right.\\
&\quad\times\exp\left\{\gamma\int_{T-(k-1)\tau}^{T}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-(k-1)\tau}^{T}w_sV_sdB_s\right\}\\
&\quad\times\left(\int_{T-(k-1)\tau}^{T}w_s^2V_s^2\mathrm{e}^{-\kappa(s-T+k\tau)}ds-\int_{T-(k-1)\tau}^{T}w_sV_s\mathrm{e}^{-\kappa(s-T+k\tau)}dB_s\right)\\
&\quad+\exp\left\{\gamma\int_{T-k\tau}^{T-(k-1)\tau}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-k\tau}^{T-(k-1)\tau}w_sV_sdB_s\right\}\\
&\quad\times\exp\left\{\gamma\int_{T-(k-1)\tau}^{T}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-(k-1)\tau}^{T}w_sV_sdB_s\right\}\\
&\quad\left.\times\left(\int_{T-k\tau}^{T-(k-1)\tau}w_s^2V_s^2\mathrm{e}^{-\kappa(s-T+k\tau)}ds-\int_{T-k\tau}^{T-(k-1)\tau}w_sV_s\mathrm{e}^{-\kappa(s-T+k\tau)}dB_s\right)|\mathcal{F}_{T-k\tau}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= W_{T-k\tau}^{\gamma}\mathrm{E}\left[\mathrm{E}\left[\exp\left\{\gamma\int_{T-k\tau}^{T-(k-1)\tau}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-k\tau}^{T-(k-1)\tau}w_sV_sdB_s\right\}\right.\right.\\
&\quad\times\exp\left\{\gamma\int_{T-(k-1)\tau}^{T}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-(k-1)\tau}^{T}w_sV_sdB_s\right\}\\
&\quad\left.\times\,\mathrm{e}^{-\kappa\tau}\left(\int_{T-(k-1)\tau}^{T}w_s^2V_s^2\mathrm{e}^{-\kappa(s-T+(k-1)\tau)}ds-\int_{T-(k-1)\tau}^{T}w_sV_s\mathrm{e}^{-\kappa(s-T+(k-1)\tau)}dB_s\right)|\mathcal{F}_{T-(k-1)\tau}\right]\\
&\quad+\mathrm{E}\left[\exp\left\{\gamma\int_{T-k\tau}^{T-(k-1)\tau}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-k\tau}^{T-(k-1)\tau}w_sV_sdB_s\right\}\right.\\
&\quad\times\exp\left\{\gamma\int_{T-(k-1)\tau}^{T}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-(k-1)\tau}^{T}w_sV_sdB_s\right\}\\
&\quad\left.\left.\times\left(\int_{T-k\tau}^{T-(k-1)\tau}w_s^2V_s^2\mathrm{e}^{-\kappa(s-T+k\tau)}ds-\int_{T-k\tau}^{T-(k-1)\tau}w_sV_s\mathrm{e}^{-\kappa(s-T+k\tau)}dB_s\right)|\mathcal{F}_{T-(k-1)\tau}\right]|\mathcal{F}_{T-k\tau}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= W_{T-k\tau}^{\gamma}\mathrm{E}\left[\exp\left\{\gamma\int_{T-k\tau}^{T-(k-1)\tau}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-k\tau}^{T-(k-1)\tau}w_sV_sdB_s\right\}\right.\\
&\quad\times\mathrm{e}^{-\kappa\tau}\mathrm{E}\left[\exp\left\{\gamma\int_{T-(k-1)\tau}^{T}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-(k-1)\tau}^{T}w_sV_sdB_s\right\}\right.\\
&\quad\left.\times\left(\int_{T-(k-1)\tau}^{T}w_s^2V_s^2\mathrm{e}^{-\kappa(s-T+(k-1)\tau)}ds-\int_{T-(k-1)\tau}^{T}w_sV_s\mathrm{e}^{-\kappa(s-T+(k-1)\tau)}dB_s\right)|\mathcal{F}_{T-(k-1)\tau}\right]\\
&\quad+\exp\left\{\gamma\int_{T-k\tau}^{T-(k-1)\tau}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-k\tau}^{T-(k-1)\tau}w_sV_sdB_s\right\}\\
&\quad\times\left(\int_{T-k\tau}^{T-(k-1)\tau}w_s^2V_s^2\mathrm{e}^{-\kappa(s-T+k\tau)}ds-\int_{T-k\tau}^{T-(k-1)\tau}w_sV_s\mathrm{e}^{-\kappa(s-T+k\tau)}dB_s\right)\\
&\quad\left.\times\,\mathrm{E}\left[\exp\left\{\gamma\int_{T-(k-1)\tau}^{T}\left[(\mu-r)w_s+r-\frac{1}{2}w_s^2V_s^2\right]ds+\gamma\int_{T-(k-1)\tau}^{T}w_sV_sdB_s\right\}|\mathcal{F}_{T-(k-1)\tau}\right]|\mathcal{F}_{T-k\tau}\right]
\end{aligned}
$$

En reprenant les mêmes arguments que pour $A_{T-k\tau}$, on obtient

$$\begin{aligned} B_{T-k\tau} \approx W_{T-k\tau}^{\gamma} \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau}+r-\frac{1}{2}w_{T-k\tau}^2V_{T-k\tau}^2\right]\tau\right\} \\ \times\Big(\mathrm{e}^{-\kappa\tau}\mathrm{E}\left[\tilde{B}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ + w_{T-k\tau}^2V_{T-k\tau}^2\frac{1-\mathrm{e}^{-\kappa\tau}}{\kappa}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ -w_{T-k\tau}V_{T-k\tau}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}\tilde{\eta}_{T-k\tau}|\mathcal{F}_{T-k\tau}\right]\Big), \end{aligned} \tag{6.E.6}$$

où $\tilde{\eta}_{T-k\tau}=\int_{T-k\tau}^{T-(k-1)\tau}\mathrm{e}^{-\kappa(s-T+k\tau)}dB_s$.

En résumé, l'équation du poids optimal de l'actif risqué dans le portefeuille vérifie

$$w_{T-k\tau}^{*}=\frac{\mu-r}{(1-\gamma)V_{T-k\tau}^2}-\frac{\sigma\rho\gamma}{(1-\gamma)V_{T-k\tau}}\frac{\tilde{B}_{T-k\tau}}{\tilde{A}_{T-k\tau}}, \tag{6.E.7}$$

avec

$$\begin{aligned} \tilde{A}_{T-k\tau} \approx \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau}+r-\frac{1}{2}w_{T-k\tau}^2V_{T-k\tau}^2\right]\tau\right\} \\ \times\,\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right], \end{aligned} \tag{6.E.8}$$

et

$$\begin{aligned} \tilde{B}_{T-k\tau} = \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau}+r-\frac{1}{2}w_{T-k\tau}^2V_{T-k\tau}^2\right]\tau\right\} \\ \times\Big(\mathrm{e}^{-\kappa\tau}\mathrm{E}\left[\tilde{B}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ + w_{T-k\tau}^2V_{T-k\tau}^2\frac{1-\mathrm{e}^{-\kappa\tau}}{\kappa}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ -w_{T-k\tau}V_{T-k\tau}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma w_{T-k\tau}V_{T-k\tau}\tilde{\xi}_{T-k\tau}\right\}\tilde{\eta}_{T-k\tau}|\mathcal{F}_{T-k\tau}\right]\Big), \end{aligned} \tag{6.E.9}$$

avec pour condition finale respectivement $\tilde{A}_T=1$ et $\tilde{B}_T=0$, et où

$$\tilde{\xi}_{T-k\tau}=\int_{T-k\tau}^{T-(k-1)\tau}dB_s, \tag{6.E.10}$$

$$\tilde{\eta}_{T-k\tau}=\int_{T-k\tau}^{T-(k-1)\tau}\mathrm{e}^{-\kappa(s-T+k\tau)}dB_s. \tag{6.E.11}$$

6.E.2 Simulation des mouvements browniens

Dans le cadre d'un mouvement brownien standard, nous savons que

$$\mathrm{E}\left[\int_a^b f(t)dW_t\right]=0, \tag{6.E.12}$$

$$\mathrm{E}\left[\int_a^b f(t)dW_t\int_a^b g(t)dW_t\right]=\int_a^b\mathrm{E}\left[f(t)g(t)\right]dt. \tag{6.E.13}$$

nous en déduisons d'une part que

$$\mathrm{E}\left[\tilde{\xi}_{T-k\tau}\right] = 0, \tag{6.E.14}$$

$$\mathrm{Var}\left[\tilde{\xi}_{T-k\tau}\right] = \int_{T-k\tau}^{T-(k-1)\tau} ds = \tau. \tag{6.E.15}$$

et d'autre part

$$\mathrm{E}\left[\tilde{\eta}_{T-k\tau}\right] = 0, \tag{6.E.16}$$

$$\mathrm{Var}\left[\tilde{\eta}_{T-k\tau}\right] = \int_{T-k\tau}^{T-(k-1)\tau} \mathrm{e}^{-2\kappa(s-T+k\tau)} ds = \frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}. \tag{6.E.17}$$

Les variables $\tilde{\xi}$ et $\tilde{\eta}$ sont reliées entre elles par

$$\mathrm{Corr}(\tilde{\xi}_{T-k\tau}, \tilde{\eta}_{T-k\tau}) = \frac{\mathrm{E}\left[\tilde{\xi}_{T-k\tau}\tilde{\eta}_{T-k\tau}\right]}{\sqrt{\mathrm{Var}(\tilde{\xi}_{T-k\tau})}\sqrt{\mathrm{Var}(\tilde{\eta}_{T-k\tau})}} = \sqrt{\frac{2(1-\mathrm{e}^{-\kappa\tau})}{\kappa\tau(1+\mathrm{e}^{-\kappa\tau})}}. \tag{6.E.18}$$

En somme, $\tilde{\xi}_t \sim N(0,\tau)$ et $\tilde{\eta}_t \sim N\left(0, \frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}\right)$ pour tout $t \in [0,T]$.

Choisissons $\xi_t \sim N(0,1)$ et $\eta_t \sim N(0,1)$ pour tout $t \in [0,T]$, telles que

$$\tilde{\xi}_t = \sqrt{\tau}\xi_t, \tag{6.E.19}$$

$$\tilde{\eta}_t = \sqrt{\frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}}\eta_t, \tag{6.E.20}$$

avec

$$\mathrm{Corr}(\xi_{T-k\tau}, \eta_{T-k\tau}) = \mathrm{Corr}(\tilde{\xi}_{T-k\tau}, \tilde{\eta}_{T-k\tau}) = \sqrt{\frac{2(1-\mathrm{e}^{-\kappa\tau})}{\kappa\tau(1+\mathrm{e}^{-\kappa\tau})}}. \tag{6.E.21}$$

6.E.3 Formulation définitive

L'équation du poids optimal de l'actif risqué (6.C.16) est finalement donnée par

$$w^*_{T-k\tau} = \frac{\mu - r}{(1-\gamma)V^2_{T-k\tau}} - \frac{\sigma\rho\gamma}{(1-\gamma)V_{T-k\tau}}\frac{\tilde{B}_{T-k\tau}}{\tilde{A}_{T-k\tau}}, \tag{6.E.22}$$

avec

$$\tilde{A}_{T-k\tau} \approx \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau} + r - \frac{1}{2}w^2_{T-k\tau}V^2_{T-k\tau}\right]\tau\right\} \times \mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right], \tag{6.E.23}$$

et

$$\begin{aligned}\tilde{B}_{T-k\tau} = \exp\left\{\gamma\left[(\mu-r)w_{T-k\tau} + r - \frac{1}{2}w_{T-k\tau}^2 V_{T-k\tau}^2\right]\tau\right\} \\ \times\left(\mathrm{e}^{-\kappa\tau}\mathrm{E}\left[\tilde{B}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right]\right. \\ + w_{T-k\tau}^2 V_{T-k\tau}^2\frac{1-\mathrm{e}^{-\kappa\tau}}{\kappa}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}|\mathcal{F}_{T-k\tau}\right] \\ \left.-w_{T-k\tau}V_{T-k\tau}\sqrt{\frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}}\mathrm{E}\left[\tilde{A}_{T-(k-1)\tau}\exp\left\{\gamma\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}\eta_{T-k\tau}|\mathcal{F}_{T-k\tau}\right]\right),\end{aligned} \tag{6.E.24}$$

avec pour condition terminale respectives $\tilde{A}_T = 1$ et $\tilde{B}_T = 0$, et où

$$\xi_{T-k\tau} \sim N(0,1), \tag{6.E.25}$$

$$\eta_{T-k\tau} \sim N(0,1), \tag{6.E.26}$$

$$\mathrm{Corr}(\xi_{T-k\tau}, \eta_{T-k\tau}) = \sqrt{\frac{2(1-\mathrm{e}^{-\kappa\tau})}{\kappa\tau(1+\mathrm{e}^{-\kappa\tau})}}. \tag{6.E.27}$$

□

6.F Démonstration de la proposition 6

Nous avons pour objectif de discrétiser l'expression de la log-volatilité. Nous avons montré (voir équation (6.B.4) dans l'annexe 6.B) qu'elle pouvait s'écrire

$$\ln V(t) = \alpha + \mathrm{e}^{-\kappa t}(\ln V(0) - \alpha) + \sigma\rho\int_0^t \mathrm{e}^{\kappa(s-t)}dB_s + \sigma\sqrt{1-\rho^2}\int_0^t \mathrm{e}^{\kappa(s-t)}dY_s^d. \tag{6.F.1}$$

Nous en déduisons aisément que

$$\ln V(t+\delta) = \alpha + \mathrm{e}^{-\kappa\delta}(\ln V(t) - \alpha) + \sigma\rho\int_t^{t+\delta} \mathrm{e}^{\kappa(s-t-\delta)}dB_s + \sigma\sqrt{1-\rho^2}\int_t^{t+\delta} \mathrm{e}^{\kappa(s-t-\delta)}dY_s^d. \tag{6.F.2}$$

Avec la même discrétisation que pour le poids optimal,

$$\ln V_{T-(k-1)\tau} = \alpha + \mathrm{e}^{-\kappa\tau}(\ln V_{T-k\tau} - \alpha) + \sigma\rho\tilde{\zeta}_{T-k\tau} + \sigma\sqrt{1-\rho^2}\tilde{\chi}_{T-k\tau}, \tag{6.F.3}$$

où

$$\tilde{\zeta}_{T-k\tau} = \int_{T-k\tau}^{T-(k-1)\tau} \mathrm{e}^{\kappa(s-T+(k-1)\tau)}dB_s, \tag{6.F.4}$$

$$\tilde{\chi}_{T-k\tau} = \int_{T-k\tau}^{T-(k-1)\tau} \mathrm{e}^{\kappa(s-T+(k-1)\tau)}dY_s^d. \tag{6.F.5}$$

En suivant le même raisonnement que pour le poids optimal de l'actif risqué,

$$\tilde{\zeta}_{T-k\tau} = \sqrt{\frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}}\zeta_{T-k\tau}, \tag{6.F.6}$$

où, pour tout $t \in [0,T]$,

$$\zeta_t \sim N(0,1), \tag{6.F.7}$$

$$\mathrm{Corr}(\zeta_t, \eta_t) = \frac{2\kappa\tau}{\mathrm{e}^{\kappa\tau} - \mathrm{e}^{-\kappa\tau}}. \tag{6.F.8}$$

Remarque 12. La corrélation est de la forme $x/\sinh(x)$. Ici aussi, lorsque τ tend vers 0, la corrélation tend vers 1.

Dans le cadre d'un mouvement brownien fractionnaire, nous ne pouvons plus utiliser les équations (6.E.12) et (6.E.13). Les règles ont changé et nous devons recourir à la propriété 2 (voir annexe 6.A). Dans un premier temps, en posant $f(s) = \mathrm{e}^{\kappa(s-T+(k-1)\tau)}\mathbb{1}_{[T-k\tau,T-(k-1)\tau]}(s)$ et $dB^d(s) = dY^d(s)$, alors

$$\mathrm{E}\left[\tilde{\chi}_{T-k\tau}\right] = 0, \tag{6.F.9}$$

puis $g(u) = \mathrm{e}^{\kappa(u-T+(k-1)\tau)}\mathbb{1}_{T-k\tau,T-(k-1)\tau}(u)$,

$$\mathrm{Var}(\tilde{\chi}_{T-k\tau}) = \int_{T-k\tau}^{T-(k-1)\tau}\int_{T-k\tau}^{T-(k-1)\tau} \mathrm{e}^{\kappa(s-T+(k-1)\tau)}\mathrm{e}^{\kappa(u-T+(k-1)\tau)} 2d\left(d+\frac{1}{2}\right)|s-u|^{2d-1}dsdu. \tag{6.F.10}$$

en effectuant les changements de variables $x = s - T + k\tau$ et $y = u - T + k\tau$, on obtient

$$\mathrm{Var}(\tilde{\chi}_{T-k\tau}) = \mathrm{e}^{-2\kappa\tau}\int_0^\tau\int_0^\tau \mathrm{e}^{\kappa x}\mathrm{e}^{\kappa y} 2d\left(d+\frac{1}{2}\right)|x-y|^{2d-1}dxdy. \tag{6.F.11}$$

À nouveau, en procédant aux changements de variables $u = x/\tau$ et $v = y/\tau$,

$$\mathrm{Var}(\tilde{\chi}_{T-k\tau}) = \tau^{2d+1}\mathrm{e}^{-2\kappa\tau}\int_0^1\int_0^1 \mathrm{e}^{\kappa\tau(x+y)} 2d\left(d+\frac{1}{2}\right)|u-v|^{2d-1}dudv. \tag{6.F.12}$$

En conclusion, la log-volatilité est approchée par

$$\ln V_{T-(k-1)\tau} = \alpha + \mathrm{e}^{-\kappa\tau}(\ln V_{T-k\tau} - \alpha) + \sigma\left(\rho\sqrt{\frac{1-\mathrm{e}^{-2\kappa\tau}}{2\kappa}}\zeta_{T-k\tau} + \sqrt{1-\rho^2}\sqrt{\mathrm{Var}(\tilde{\chi})}\chi_{T-k\tau}\right). \tag{6.F.13}$$

avec χ_t des variables aléatoires sériellement corrélées vérifiant

$$\mathrm{Var}(\tilde{\chi}) = \tau^{2d+1}\left\{\mathrm{e}^{-\kappa\tau} + \frac{(\kappa\tau\mathrm{e}^{-\kappa\tau})^2}{2}\int_0^1\int_0^1 \mathrm{e}^{\kappa\tau(x+y)}|x-y|^{2d+1}dxdy\right\}. \tag{6.F.14}$$

□

6.G Démonstration de la proposition 7

La richesse du portefeuille s'écrit

$$W_t = W_0 \exp\left\{\int_0^t\left[(\mu-r)w_s + r - \frac{1}{2}w_s^2V_s^2\right]ds + \int_0^t w_sV_sdB_s\right\}. \tag{6.G.1}$$

nous en déduisons

$$W_{t+\delta} = W_t \exp\left\{\int_t^{t+\delta}\left[(\mu-r)w_s + r - \frac{1}{2}w_s^2V_s^2\right]ds + \int_t^{t+\delta} w_sV_sdB_s\right\}, \tag{6.G.2}$$

$$\approx W_t \exp\left\{\left[(\mu-r)w_t + r - \frac{1}{2}w_t^2V_t^2\right]\delta + \sqrt{\delta}w_tV_t\xi_t\right\}. \tag{6.G.3}$$

Avec la même discrétisation que pour le poids optimal et la log-volatilité

$$W_{T-(k-1)\tau} \approx W_{T-k\tau}\exp\left\{\left[(\mu-r)w_{T-k\tau} + r - \frac{1}{2}w_{T-k\tau}^2V_{T-k\tau}^2\right]\tau\right\} \times \exp\left\{\sqrt{\tau}w_{T-k\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}. \tag{6.G.4}$$

□

Quant à la dynamique du prix de l'actif risqué,elle est donnée par

$$dS_1(t) = \mu S_1(t)dt + V(t)S_1(t)dB(t). \tag{6.G.5}$$

En appliquant le lemme d'Itō à la fonction $f(S_1(t)) = \exp(S_1(t))$,

$$S_1(t) = S_1(0)\exp\left\{\int_0^t\left(\mu - \frac{1}{2}V(s)^2\right)ds + \int_0^t V(s)dB(s)\right\}. \tag{6.G.6}$$

Nous en déduisons

$$S_1(t+\delta) = S_1(t)\exp\left\{\int_t^{t+\delta}\left(\mu - \frac{1}{2}V(s)^2\right)ds + \int_t^{t+\delta} V(s)dB(s)\right\}, \tag{6.G.7}$$

$$\approx S_1(t)\exp\left\{\left(\mu - \frac{1}{2}V(t)^2\right)\delta + \sqrt{\delta}V(t)\xi_t\right\}. \tag{6.G.8}$$

Avec la même discrétisation que pour le poids optimal et la log-volatilité, le prix de l'actif risqué s'écrit

$$S_{1,T-(k-1)\tau} \approx S_{1,T-k\tau}\exp\left\{\left(\mu - \frac{1}{2}V_{T-k\tau}^2\right)\tau + \sqrt{\tau}V_{T-k\tau}\xi_{T-k\tau}\right\}. \tag{6.G.9}$$

□

Conclusion générale

Durant cette thèse, nous nous sommes intéressés à une caractéristique très importante de la volatilité des rendements d'action, à savoir la *mémoire longue*. Elle se traduit par une lente décroissance, hyperbolique, de la fonction d'autocorrélation de la volatilité. Cette persistance à long terme indique que le niveau de volatilité actuel a un impact sur la volatilité future pouvant durer des mois. Par conséquent, la prise en compte de la mémoire longue est cruciale pour appréhender correctement le risque futur des actions. Notre intérêt s'est manifesté selon deux axes de recherche. Le premier a consisté à fournir une interprétation de la mémoire longue de la volatilité. Pour y parvenir, nous avons développé un modèle de volatilité dont les paramètres sont reliés aux comportements hétérogènes des agents. Le second axe visait à introduire la mémoire longue dans un modèle de choix de portefeuille en temps continu et à analyser les conséquences en matière de gestion.

Dans la première partie, nous avons proposé un modèle de volatilité d'action en tenant compte de l'hétérogénéité des comportements des investisseurs. Malgré l'accessibilité actuelle aux données intrajournalières, la volatilité n'est pas une donnée observable en direct sur le marché. Elle doit être estimée. Par conséquent, nous supposons que chaque acteur du marché fait une anticipation du niveau actuel de volatilité et que le marché se charge d'agréger l'ensemble de ces anticipations pour former la réalisation de la volatilité. Cette hypothèse est inspirée par le fait que le marché agrège les ordres d'achats et de ventes des investisseurs pour fixer le prix des actions à chaque instant. Nous supposons également que chaque agent, pour faire une anticipation du niveau de volatilité se base sur quatre critères, à savoir (i) la réalisation passée, (ii) l'anticipation précédente, (iii) une information publique et (iv) une information privée. Ces critères permettent aux agents de se différencier : certains basent leurs anticipation sur les informations et sont qualifiés de rationnels, tandis que d'autres négligent les informations pour se fonder sur leur anticipation passée ou la réalisation passée et font preuve d'une rationalité limitée.

Lors de l'étude théorique du modèle, nous avons mis en évidence un paramètre, qualifié de *paramètre d'hétérogénéité* qui caractérise la diversité du comportement des investisseurs. La divergence de la densité de ce paramètre est une condition nécessaire à la présence de mémoire longue : plus la divergence est rapide, plus la mémoire longue est importante. En termes de comportements, cela signifie que la mémoire longue est d'autant plus grande que la proportion relative d'agents fondant leur anticipation de la volatilité sur leur anticipation passée est grande. Autrement dit, plus l'hétérogénéité des comportements est grande, plus la mémoire longue est importante. Ce résultat apporte une interprétation de la signification de la mémoire longue différente de celle de Dacorogna, Gençay, Müller, Olsen, et Pictet (2001) et Di Matteo, Aste, et Dacorogna (2005), qui en considérant une autre forme d'hétérogénéité et en utilisant d'autres données que le cours d'actions, aboutissent à la conclusion contraire.

L'étape suivante a consisté à définir la densité du paramètres d'hétérogénéité. Comme nous venons de le souligner, pour traduire la mémoire longue, la densité doit diverger au voisinage de un. Nous optons pour une loi Béta car nous pouvons concrètement contrôler la vitesse de la divergence au moyen d'un paramètre. En terme de comportement, il s'agit de la proportion d'agents auto-référents, c'est à dire ceux qui fondent leur anticipation majoritairement sur leur anticipation passée. Hélas, l'utilisation d'une loi Béta ne permet pas de traduire d'autres comportements, tels que le comportement rationnel. Par conséquent, la densité du paramètre d'hétérogénéité a été définie comme le mélange d'une loi Beta et

d'une loi régulière, capable elle d'accorder du poids aux agents fondant leur anticipation sur les informations ou sur l'opposé de leur anticipation passée. Afin d'estimer les paramètres du modèle, nous avons utilisé l'estimateur classique minimisant la distance entre la fonction d'autocorrélation empirique et la fonction d'autocorrélation, obtenue pour un jeu de paramètres donnés, de la série de volatilité. Il a pour propriété de converger asymptotiquement vers une loi normale et présente l'avantage d'accorder un poids plus important aux lags dont l'estimation de la volatilité est la meilleure, en l'occurrence il s'agit des premiers lags. Enfin, pour la mise en œuvre, nous avons suggéré un algorithme fondé sur l'algorithme de Nelder Mead qui présente l'avantage d'avoir, pour une efficacité équivalente à celui d'un algorithme génétique, un temps de calcul nettement inférieur.

Ensuite, nous avons appliqué notre modèle aux données historiques, entre janvier 1994 et décembre 2003, de 24 actions américaines à moyennes et grandes capitalisations appartenant à différents secteurs d'activité. Nous avons constaté deux allures de densité bien distinctes à l'exception d'un point commun : la proportion d'agents auto-référents (qui utilisent leur anticipation passée). A noter que ces deux allures sont indépendantes *a priori* de la capitalisation et du secteur d'activité de l'entreprise. Pour être plus catégorique, il faudrait disposer des données d'un plus grand nombre d'actions.

Nous avons clôt la première partie en étudiant le comportement des agents en fonction de la phase du marché. Pour y parvenir, nous nous sommes focalisés sur la crise obligataire de 1998 ainsi que la bulle internet de la fin des années 1990. Conformément à nos attentes, nous avons observé un net changement de comportement entre la phase précédant l'éclatement de la bulle et celle la suivant. En effet, lors du développement d'une bulle, le comportement auto-référent des agents est bien plus prononcé que lors d'une phase de crise durant laquelle les agents ont plutôt tendance à être très attentifs aux informations.

Cette première partie de thèse ouvre la voie à de nombreuses perspectives. Tout d'abord, nous avons évoqué pour justifier nos travaux de recherche le possible gain en termes de prévisibilité d'un modèle de volatilité prenant en compte la mémoire longue de la volatilité d'action. Par conséquent, la première perspective consisterait à comparer l'efficacité prédictive de notre modèle avec d'autres modèles classiques existants. Dans un premier temps, nous pourrions le comparer à des modèles sans mémoire longue tels que les modèles ARMA de Box et Jenkins (1970) et GARCH de Bollerslev (1986). Puis dans un deuxième temps, à d'autres modèles prenant en considération, cette fois, la mémoire longue, tels que les modèles fractionnairement intégrés ARFIMA de Granger et Joyeux (1980) et FIGARCH de Baillie, Bollerslev, et Mikkelsen (1996).

Un deuxième axe de recherches futures concernerait l'origine de la mémoire longue. Nous avons opté pour la théorie de l'agrégation de Granger (1980). Or, une explication alternative fournie entre autres par Liu (2000) et Granger et Hyung (2004) fait état de changements structuraux résultant de chocs financiers de grande ampleur. Il conviendrait donc de se demander si la mémoire longue que nous observons ne provient pas aussi de changements structuraux éventuels. En d'autres termes, la présence de tels changements peut biaiser l'analyse de séries temporelles financières réalisée à l'aide de modèles fractionnairement intégrés. Or, puisque les modèles fractionnairement intégrés et les modèles à changements structurels supposent une mémoire longue ayant des origines différentes, une combinaison de ces deux modèles pourrait

être par conséquent plus efficace pour modéliser les séries temporelles financières et faire de la prédiction.

Dans la deuxième partie de la thèse, notre but principal était de proposer un modèle de choix de portefeuille, en temps continu, basé sur la théorie de l'utilité espérée, avec une volatilité stochastique à mémoire longue. La volatilité stochastique a été introduite au moyen d'un processus d'Ornstein-Uhlenbeck que nous avons utilisé pour caractériser le processus de log-volatilité. Un tel processus est tout à fait indiqué puisqu'il traduit un fait stylisé de la volatilité : le retour à la moyenne.

Nous avons dérivé les résultats analytiques de notre modèle intermédiaire pour trois critères de gestion couramment rencontrés dans la littérature. La fonction d'utilité choisie est une fonction HARA, classe qui englobe la fonction CRRA. La résolution analytique de notre problème par la méthode classique par programmation dynamique et équation d'Hamilton Jacobi-Bellman, nous a conduit à une équation aux dérivées partielles trop compliquée pour être résolue analytiquement, contrairement à celle obtenue dans le cadre du cas particulier du modèle de Merton. Nous avons toutefois été capable de la simplifier et avons proposé une méthode numérique basée sur la méthode des différences finies pour obtenir une approximation de la solution.

L'étude de l'impact du caractère stochastique sur l'allocation optimale a été menée en considérant un investisseur moyennement risquophobe, dont l'aversion au risque est caractérisée par une fonction d'utilité CRRA, consommant et souhaitant faire fructifier son patrimoine sur une période d'un an, et laisser un capital à l'échéance. Nous avons montré que l'amplitude de l'écart relatif entre la volatilité stochastique et la volatilité moyenne, que nous avons appelé *degré de stochasticité*, est contrôlée par deux paramètres. Le degré de stochasticité est synonyme de source d'incertitude quant au niveau futur de la volatilité. Par ailleurs, nous avons établi que le niveau d'incertitude a un impact sur le plan de consommation et d'investissement qui diminue au fur et à mesure que l'on se rapproche de l'horizon de gestion. En particulier, l'augmentation de l'incertitude induit une révision du plan de gestion.

A ce stade, nous nous devons d'attirer l'attention sur un point important lié à la modélisation de l'effet de levier ayant une grande influence sur les résultats. Dans de nombreux modèles comme dans le nôtre (voir Chacko et Viceira (2005), Liu (2007) ou encore Heston (1993) dans un autre domaine), l'effet de levier est pris en compte au moyen de la corrélation ρ entre le rendement de l'actif risqué et la volatilité. Donc prévoir la volatilité au temps t permet de déduire, en partie, le rendement au temps t.

Ensuite, nous avons apporté une amélioration à notre premier modèle pour lui permettre de prendre en compte la mémoire longue de la volatilité : nous avons remplacé le processus d'Ornstein-Uhlenbeck par un processus voisin du processus d'Ornstein-Uhlenbeck fractionnaire qui fait intervenir le mouvement brownien fractionnaire. Cette modélisation présente deux avantages. En premier lieu, elle a la particularité de relier explicitement la mémoire longue de la volatilité à un paramètre. En outre, elle permet de contrôler concrètement la corrélation entre la volatilité de l'actif risqué et son rendement.

La conséquence de la prise en compte de la mémoire longue de la volatilité, en l'occurrence

au moyen d'un mouvement brownien fractionnaire, est qu'à présent un grand nombre d'observations passées de la volatilité sont utilisées pour déterminer la volatilité actuelle, alors que dans les modèles avec volatilité stochastique à mémoire courte, seule la volatilité précédente importait. Nous sommes donc passé d'un cadre markovien à un cadre non-markovien. Par ailleurs, la théorie stochastique classique d'une part, et l'approche classique par programmation dynamique et équation d'Hamilton-Jacobi-Bellman d'autre part, ne peuvent plus s'appliquer. Nous avons donc eu recours à la théorie du calcul stochastique appliquée au mouvement brownien fractionnaire et nous avons fourni une méthode originale de résolution numérique de notre problème faisant intervenir la technique de Monte Carlo. D'ailleurs, afin de simplifier le problème, nous avons été contraint de dériver les résultats analytiques pour un seul critère qui s'avère être un pur cas de choix de portefeuille : la seule maximisation de l'utilité espérée de la richesse finale.

L'étude de l'impact de la mémoire longue de la volatilité de l'actif risqué dans la gestion de portefeuille a permis de mettre en évidence la nécessité de tenir compte de ce phénomène dans la modélisation du risque d'une action. Dans le cas contraire, l'allocation optimale obtenue au sens du modèle serait éloignée de l'allocation optimale réelle puisque le risque réel ne serait pas correctement appréhendé. En particulier, nous avons montré que les conséquences d'une mauvaise spécification du modèle, en termes de rendement de portefeuille, sont d'autant plus pénalisantes, voire désastreuses, que l'aversion au risque de l'investisseur est faible. En outre, l'impact de la durée de gestion n'est pas très concluant dans notre étude. Toutefois, il s'agit d'une seule simulation et nous pouvons penser qu'en règle générale, les erreurs s'accumulent et s'amplifient sur un horizon de gestion long.

Les perspectives de cette deuxième partie sont également nombreuses. En effet, pour commencer, nous pourrions traiter un plus grand nombre de critères de gestion. Nous avions commencé par traiter un pur cas de choix de portefeuille pour sa relative simplicité de résolution. Par conséquent, il serait intéressant de traiter d'autres critères de gestion, tels que le critère général de la maximisation de l'espérance de l'utilité de la consommation de l'investisseur sur $[0, T]$, avec héritage à l'échéance.

Une autre perspective consisterait à améliorer le modèle en faisant l'hypothèse d'un rendement espéré de l'actif risqué variable dans le temps, comme l'ont fait Campbell et Viceira (1999), et non plus constant comme c'était le cas jusqu'à présent. Nous pourrions traduire cette dynamique au moyen d'une simple fonction ou, de manière plus complexe, d'un processus stochastique.

Enfin, il serait intéressant de généraliser notre modèle à N actifs risqués. En effet, un portefeuille n'est jamais constitué que d'une action pour des raisons de diversification. Et même si l'on peut considérer que l'action dans notre modèle représente à elle seule une synthèse de l'ensemble des actions d'un portefeuille, il serait préférable d'utiliser la généralisation à N actifs. A noter toutefois que cela représenterait avant tout un intérêt théorique, car du point de vue pratique étant donné le temps de calcul nécessaire à l'optimisation dans le cadre d'un actif risqué et d'un actif non risqué, il est très probable que l'on rencontre assez rapidement des problèmes lorsque l'on souhaite traiter le cas de plusieurs actifs risqués. Par ailleurs, une perspective potentiellement importante serait de déduire de notre deuxième modèle à volatilité stochastique un modèle d'équilibre des actifs financiers tenant compte de la mémoire

longue, à l'image de ce qu'a proposé Breeden (1979).

Annexe A

Évaluation de la conduite du projet de thèse

Cette annexe s'inscrit dans le cadre d'un projet pilote développé par l'ancien ministère de l'éducation nationale, de la recherche et de la technologie (MENRT) afin d'aider les jeunes docteurs à s'insérer plus facilement dans la vie active, que ce soit dans le milieu de la recherche académique ou bien au sein d'entreprises privées. L'objet de ce chapitre [52] est de mener à bien une réflexion sur la conduite de la thèse, non seulement en termes de retombées scientifiques, mais également d'un point de vue comptable, avec l'évaluation du coup du projet, et de faire le bilan des expériences et compétences acquises au cours de ces années.

A.1 Bref résumé du sujet

Cette thèse porte sur la mémoire longue de la volatilité. Dans une première partie, nous avons proposé une interprétation de la mémoire longue en termes de comportements d'agents. Pour y parvenir, nous avons développé un modèle de volatilité ayant la particularité d'avoir des paramètres reliés aux comportements hétérogènes des agents. Dans une deuxième partie, nous avons pris en compte la mémoire longue en gestion de portefeuille. Nous avons procédé en deux temps. Nous avons commencé par proposer un modèle de choix de portefeuille en temps continu à volatilité stochastique et étudié l'impact du caractère stochastique de la volatilité sur la stratégie optimale. Puis, nous avons amélioré ce premier modèle en introduisant la mémoire longue au moyen d'un mouvement brownien fractionnaire et avons analysé l'impact de la mémoire longue sur l'allocation optimale.

52. Pour les thèses futures, le MENRT, devenu depuis le ministère de l'enseignement supérieur et de la recherche, a décidé de le remplacer par un nouveau chapitre intitulé "Valorisation des compétences" dont l'objectif est de "préparer le doctorant à la démarche de recherche d'emploi en le conduisant à identifier les compétences multiples mises en œuvre au cours des trois années de thèse et à les valoriser dans des situations professionnelles diversifiées. Ce travail, encadré par un tuteur généralement issu d'un cabinet de recrutement, donne lieu à l'élaboration d'un document de quelques pages, véritable bilan de compétences" (source : http ://www.enseignementsup-recherche.gouv.fr/cid20185/doctorat.html)

A.2 Éléments de contexte

A.2.1 Choix du sujet

Après avoir préparé le Diplôme d'Études Approfondies de Sciences Actuarielle et Financière de l'ISFA - Université Lyon 1 durant l'année universitaire 2003-2004, j'ai décidé de m'inscrire en thèse. Après l'obtention d'un tel diplôme, j'avais la possibilité de m'orienter vers l'assurance ou la finance. J'ai finalement opté pour la finance. Mon objectif initial étant d'aller travailler dans le privé, je savais, à condition de choisir un sujet porteur, que j'aurais la possibilité de décrocher un emploi dans une banque, en particulier à l'étranger où le diplôme de doctorat est plus valorisé.

En réalité, la finance m'attirait déjà depuis quelques années lorsque j'étais inscrit en licence de mathématiques. Assez tôt, je me suis passionné pour les mathématiques appliquées. En effet, contrairement aux mathématiques abstraites telles que la topologie, les probabilités et les statistiques sont des outils permettant d'apporter des réponses à des questions concrètes. Après une maîtrise de mathématique appliquées, j'ai fait le choix de m'inscrire en DESS de statistiques, Informatiques et Techniques Numériques afin consolider mes acquis en probabilités et statistiques et de progresser en informatique, un outil dont la maîtrise est primordiale de nos jours, notamment en finance. Bien qu'un DESS soit un diplôme professionnalisant, une fois celui-ci acquis, je me suis inscrit en DEA SAF puisque j'étais toujours bien décidé à découvrir le monde de finance.

Dans les mois qui ont précédé ma thèse, j'ai découvert différents domaines en finance et pu discuter avec un certain nombre de personnes, la plupart issues du monde académique, mais également du privé. Parmi elles, deux de mes professeurs, François Quittard-Pinon et Yannick Malevergne, qui sont devenus par la suite mon directeur et co-directeur de thèse. Le sujet de mon mémoire, qui portait sur la gestion de portefeuille a été l'élément décisif dans le choix de mon sujet de thèse. Il s'agissait d'étudier les facteurs qui influencent la diversification d'un portefeuille d'actions. Pour se prémunir d'un risque trop important, il est nécessaire de ne pas faire dépendre le rendement de son portefeuille d'un trop petit nombre d'actions car si leur cours chute, le rendement du portefeuille sera fortement négativement affecté. Cette première approche de la gestion de portefeuille m'a donné envie d'approfondir mes connaissances dans ce domaine, en particulier de mettre l'accent sur la problématique de l'allocation optimale.

A.2.2 Choix de l'encadrement et du laboratoire d'accueil

La thèse, c'est bien connu, est un projet de longue durée semé d'embûches souvent d'ordre scientifique (les aléas de la recherche) parfois humaines. J'ai pris connaissances de témoignages de doctorants s'étant retrouvés en situation d'échec entre autres à cause d'incompatibilité d'humeur ou de la divergence de points de vue avec leur directeur. Pour que mon aventure ne se termine pas en voie de garage, j'étais donc conscient qu'il était primordial de choisir un encadrement de qualité dont le mode de fonctionnement corresponde à ma façon de travailler,

à mon caractère et à mes besoins.

Par chance, je n'ai pas eu à chercher bien loin pour réunir toutes ces conditions. En effet, durant mon DEA, j'ai eu le privilège de rencontrer des enseignants motivants laissant une grande liberté d'action, mais également très disponibles lorsque le besoin s'en faisait sentir. J'ai d'abord discuté du projet avec Pr. Yannick Malevergne, sous la direction duquel j'ai effectué mon mémoire de DEA, qui m'a orienté vers Pr. François Quittard-Pinon. Après avoir discuté de mon intérêt pour la gestion de portefeuille, nous avons tous les trois défini le sujet de la thèse que j'effectuerai au sein du laboratoire de Sciences Actuarielle et Financière de l'ISFA - Université Lyon 1. Monsieur Quittard-Pinon a accepté d'être mon directeur de thèse et monsieur Malevergne mon co-directeur.

Leur seule condition, qui rejoignait ma priorité, était que j'obtienne un financement. De leur point de vue, il n'était pas concevable d'espérer réaliser une bonne thèse s'il avait fallu que je subvienne à mes besoins en parallèle. J'ai pu satisfaire à cette exigence en obtenant une allocation de recherche.

A.2.3 Financement du projet

Lorsque j'ai réfléchi aux différentes possibilités de financement, j'ai rapidement opté pour un financement public. Non pas que mon sujet ne pouvait pas intéresser des établissements financiers, mais je voulais être sûr de consacrer mon temps à la thèse uniquement. Un de mes amis qui avait débuté une thèse CIFRE au sein d'une banque regrettait de ne pouvoir y consacrer qu'un jour par semaine environ, tant et si bien qu'au bout d'un an, n'ayant presque pas avancé, il décida d'arrêter. J'ai donc rempli un dossier de demande d'allocation de recherche du ministère de l'éducation nationale et l'ai défendu avec succès devant le jury de l'école doctorale de Sciences Economiques et de Gestion (anciennement Management Information Finance). Étant sur le point de m'inscrire en thèse de sciences de gestion, mon sujet plutôt mathématique pouvait me desservir quant à l'obtention de cette allocation. Mais finalement, en mettant l'accent sur les bénéfices que l'on pouvait raisonnablement espérer obtenir en finance comportementale, celui-ci fut jugé intéressant et innovant, ce qui me permit d'obtenir l'attribution de cette allocation sur 3 ans, d'un montant de 23 000 euros annuel, charges patronales comprises.

A ce montant s'ajoute également un pourcentage du coût salarial de mes directeur et co-directeur de thèse. Naturellement, ce coût est difficilement estimable, mais on peut considérer que chacun de mes encadrants m'ont consacré en moyenne environ 10% de leur temps. Cela doit représenter environ 15000 euros par an charges comprises sur 4 ans. Par conséquent, j'estime que le coût total de ma thèse en ressources humaines s'élève à 129 000 euros.

Pour faire une thèse en finance quantitative, le coût du matériel s'avère être relativement modeste. Il faut un ordinateur, une imprimante (avec encre et feuilles), les licences des logiciels Matlab et Maple. Par ailleurs il a fallu se procurer des données financières hautes-fréquences plutôt onéreuses. Le montant total est d'environ 5 000 euros.

De plus, durant ma thèse j'ai effectué divers déplacements pour participer à des conférences. Ceci ont engendré des coûts de transport, d'hébergement, de nourriture et d'inscription d'un montant total de 3 000 euros. Lors de mon année passée à l'ETH, j'ai obtenu de la part de la chair "Entrepreneurial risk" une aide d'un montant de 8 500 francs suisses, soit environ 5 300 euros.

Enfin, à tout cela il reste à ajouter ma part dans les frais de fonctionnement du laboratoire qui m'a accueilli pendant 3 ans. Il est difficile d'évaluer avec précision ces frais étant donné qu'ils regroupent le local, l'électricité, les ressources documentaires... Il est raisonnable d'aboutir à une estimation de 6 000 euros.

Voici le tableau récapitulatif des frais engagés :

Ressources humaines	129 000 euros
Frais de déplacement	8 300 euros
Frais de matériel	5 000 euros
Frais de fonctionnement	6 000 euros
TOTAL	**148 300 euros**

TABLE A.1 – Récapitulatif des différents frais engagés pour la thèse

Le total des frais atteint 148 300 euros sur 4 ans, soit un coût annuel d'environ 37 000 euros. Dans ce total, les coûts en ressources humaines représente 87%, ce qui est à la fois considérable et tout à fait normal dans une discipline faisant appel aux mathématiques.

A.3 Évolution du projet

A.3.1 Élaboration du projet

Le projet initial visait à développer dans une première partie un modèle de volatilité d'actions permettant de relier le comportement des agents intervenant sur le marché aux paramètres du modèle afin de fournir une interprétation de la mémoire longue. Dès le départ nous nous sommes intéressés à la volatilité réalisée qui présente l'avantage de faire intervenir un grand nombre de données et donc de véhiculer une plus grande quantité d'informations que les autres estimateurs. Puisque nous voulions prendre en compte le comportement des agents, nous avons décidé de faire appel à la théorie de l'agrégation.

Quant à l'objectif de la deuxième partie, il consistait à proposer un modèle de gestion de portefeuille prenant en compte la dynamique de la volatilité. En particulier, ayant pris connaissance des faits stylisés de la volatilité, nous voulions intégrer dans ce modèle le fait que d'une part la volatilité évolue au cours du temps et d'autre part qu'elle présente le phénomène de mémoire longue. Pour ce faire nous avons décidé d'utiliser le mouvement brownien fractionnaire et la théorie du calcul stochastique qui s'y rapporte.

Pour commencer nous n'avons pas détaillé le sujet plus que cela, aussi bien concernant la première partie que la deuxième, car à l'origine, nous ne connaissions pas exactement les difficultés que nous allions rencontrer et la façon de les surmonter.

A.3.2 Conduite du projet

Le projet s'est bien passé grâce à la qualité de mon encadrement qui m'a accordé une grande autonomie tout en étant présent lorsque j'en avais besoin. De ce point de vue-là, je n'ai pas été déçu car je savais avec qui je m'engageais pour avoir côtoyé mes encadrants tout au long de mon année de DEA. Au contraire, j'ai apprécié leurs remarques et leurs observations toujours très justes qui ont eu le mérite de me recentrer sur le sujet lorsque je me dispersais ou au contraire de me ré-aiguiller lorsque j'étais dans une impasse, surtout en début de thèse où l'on manque de repère et de vision globale du sujet.

Toutefois, ce projet a duré plus longtemps que prévu en raison des difficultés scientifiques rencontrées mais également du séjour à l'ETH. J'ai rencontré un certain nombre de difficultés dans le domaine de la programmation dans la première partie de la thèse, notamment dues à des problèmes de convergence. Dans la deuxième partie, les difficultés provenaient des outils mathématiques complexes qu'il a fallu assimiler, par exemple la théorie du calcul stochastique adapté au mouvement brownien fractionnaire, ou encore sa dérivée de Malliavin.

A.3.3 Retombées scientifiques

Les retombées scientifiques de cette thèse sont nombreuses. Dans une première partie, nous avons proposé une interprétation de la mémoire longue de la volatilité en termes de comportements d'agents grâce à un modèle reliant la mémoire longue et les autres paramètres du modèle au comportement des agents. Nous montrons que la mémoire longue va de pair avec l'hétérogénéité des investisseurs. En outre, notre modèle nous a permis d'observer le comportement des investisseurs en fonction de la phase du marché. En particulier, nous constatons un changement d'attitude entre la phase précédant l'éclatement d'une bulle spéculative et celle qui la suit.

Ces travaux ont donné lieu à un article intitulé "Heterogeneous expectations and long range correlation of the volatility of asset returns", présenté dans plusieurs conférences internationnales telles que

- la 30ème conférence internationale "Forecasting Financial Markets" (Aix-en-Provence, 2006),
- le congrès international de l'association française de finance (AFFI) (Poitiers, 2006),
- le 3ème "European EIF Job Market in Finance and Accounting" (Paris, 2006),

et soumis à la revue *Quantitative Finance*. D'autre part, cette première partie ouvre de nombreuses perspectives, notamment celle consistant à étudier la performance en terme de prédiction de notre modèle en le comparant à d'autres modèles classiques à mémoire courte et à mémoire longue. Par ailleurs, nous avions fondé notre modèle sur l'une des deux hypothèses de l'origine de la mémoire longue. Dans le futur, nous pourrions nous intéresser à l'autre

hypothèse.

Dans la deuxième partie, nous avons pris en compte la mémoire longue dans la gestion de portefeuille. A notre connaissance, il n'existe pas de modèle traduisant explicitement ce phénomène. La principale conséquence est le changement de cadre du système économique : alors qu'il est markovien dans les modèles sans mémoire longue, il devient non-markovien. Autrement dit, le risque passé rentre en ligne de compte dans la stratégie d'allocation optimale. Il en résulte que la théorie classique par programmation dynamique et équation d'Hamilton-Jacobi-Bellman est inutilisable. Nous avons appliqué la théorie du calcul stochastique adapté au mouvement brownien fractionnaire et avons proposé une nouvelle méthode de résolution fondé sur un algorithme récursif faisant appel à la technique de Monte Carlo. Grâce à ce modèle, nous montrons l'importance de la prise en compte de la mémoire longue dans la gestion de portefeuille et mettons en garde l'investisseur contre les erreurs de spécification du modèle.

Suite aux résultats de cette deuxième partie, un article est en cours de rédaction. Une fois encore, les perspectives d'un tel modèle sont nombreuses. Premièrement, nous pourrions élargir la gamme de critères de gestion traités par le modèle. Nous pourrions aussi améliorer le modèle en faisant l'hypothèse d'un rendement de l'actif risqué variable au cours du temps. Enfin, nous pourrions généraliser notre modèle à n actifs risqués et déduire un modèle d'équilibre des actifs financiers tenant compte de la mémoire longue.

A.4 Compétences acquises et enseignements personnels

La thèse a constitué une tranche de vie durant laquelle j'ai pu progresser dans de nombreux domaines. Sur le plan scientifique d'abord, la thèse m'a permis d'appliquer les connaissances que j'avais acquises au cours de mes études de mathématiques appliquées et notamment au cours du master recherche de Sciences Actuarielle et Financière. Par ailleurs, j'ai eu l'occasion d'enrichir mes connaissances lorsque le sujet de thèse me conduisait à accéder à des concepts nouveaux et nécessitait l'assimilation de techniques mathématiques que je n'avais pas ou très peu abordées lors de mes études antérieures.

En outre, dans le domaine de la finance, j'ai à présent une bonne connaissance du phénomène de mémoire longue, de la volatilité, ses estimateurs et ses modèles ainsi que de la modélisation en gestion de portefeuille.

Sur le plan purement technique, j'ai acquis de bonnes bases en informatique. J'ai principalement utilisé le logiciel de calcul matriciel Matlab pour écrire mes programmes. Occasionnellement, j'ai également fait appel à Maple qui lui est un logiciel de calcul formel.

J'ai également appris à gérer mon temps de travail ainsi que le travail en équipe. Ceci est nécessaire à la bonne conduite de tout projet.

Par ailleurs, j'ai saisi l'opportunité d'exposer le résultat de mes recherches lors de séminaires et conférences afin d'obtenir des points de vue différents. Ces expériences ont été

bénéfiques à double titre :

- premièrement, j'ai obtenu des avis de chercheurs d'horizons différents, certains étaient spécialistes du sujet que je leur présentais et leur avis avait donc valeur d'expert, alors que d'autres exerçaient dans des domaines différents, plus économiques et moins mathématiques et dont les réflexions avaient une connotation plus pragmatique. Ces deux axes de réflexion ont positivement aiguïllé mes recherches.
- deuxièmement, mes participations à ces séminaires et conférences, pour certaines internationales, m'ont permis d'améliorer mes capacités de présentation orale.

Enfin, ayant eu la possibilité de passer un an dans la chaire "Entrepreneurial Risks" de l'ETH Zürich, pour poursuivre mes recherches, j'ai baigné dans un environnement scientifique multi-culturel où se côtoyaient et interagissaient des doctorants, docteurs et professeurs d'origines diverses telles que suisse, allemande, française, chinoise, américaine, luxembourgeoise. J'ai ainsi pu reparler anglais et à plus petite échelle allemand.

A.5 Conclusion

J'estime que ma thèse s'est bien déroulée et cela résulte en grande partie de la bonne entente qui a régné entre mes encadrants et moi-même. Je n'aurais sans doute pas atteint les mêmes objectifs si cela avait été différent. La dimension humaine est d'autant plus importante pour mener à bien un projet que le nombre de personnes interagissant est petit. Un autre facteur de réussite a été le choix du sujet, innovant et stimulant. En particulier la première partie car, de mon point de vue, les mathématiques doivent être un outil permettant d'apporter des réponses à des questions concrètes. En reliant les paramètres du modèle de volatilité au comportement des agents et fournissant une interprétation de la mémoire longue, mon objectif a été atteint.

Cette expérience m'a conforté dans l'idée que je me faisais avant la thèse de l'orientation que je voulais donner à ma carrière professionnelle, à savoir poursuivre dans le privé. Premièrement, en dépit du fait que la thèse se soit bien passée, j'ai travaillé la plupart du temps seul ou en petit groupe. Or, j'aime par dessus tout travailler en équipe et interagir avec un plus grand nombre de personnes. Hélas, j'ai constaté que la recherche académique est en règle générale plus une activité solitaire. Deuxièmement, comme je l'ai déjà évoqué, à mes yeux le principal intérêt des mathématiques est d'apporter des réponses à des problèmes concrets. Or, quoi de plus concret que les problèmes rencontrés au sein des banques et des assurances ?

Bibliographie

Abramowitz, M., et I. Stegun (1965) : *Handbook of Mathematical Functions*. Dover.

Akian, M., A. Sulem, et M. Taksar (2001) : "Dynamic optimization of long term growth rate for a portfolio with transaction costs and logarithm utility," *Mathematical Finance*, 11 (2), 153–188.

Alexander, G., et A. Baptista (2002) : "Economic implications of using a mean-VaR model for portfolio selection : A comparison with mean-variance analysis," *Journal of Economic Dynamics and Control*, 26, 1159–1193.

Alfarano, S., F. Wagner, et T. Lux (2005) : "Estimation of Agent-Based Models : The Case of an Asymmetric Herding Model," *Computational Economics*, 26, 19–49.

Allen, F., et D. Gale (1992) : "Stock-price manipulation," *The Review of Financial Studies*, 5 (3), 503–529.

Ané, T., et H. Geman (2000) : "Order flow, transaction clock, and normality of asset returns," *Journal of Finance*, 55, 2259–2284.

Andersen, T., T. Bollerslev, F. Diebold, et H. Ebens (2001) : "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61, 43–76.

Andersen, T., T. Bollerslev, F. Diebold, et P. Labys (2001) : "The distribution of exchange rate volatility," *Journal of the American Statistical Association*, 96, 42–55.

Areal, N., et S. Taylor (2002) : "The realized volatility of FTSE-100 future prices," *Journal of Future Markets*, 22, 627–648.

Atkinson, C., et S. Mokkhavaisa (2003) : "Intertemporal portfolio optimization with small transaction costs and stochastic variance," *Applied Mathematical Finance*, 10, 267–302.

Atkinson, C., S. Pliska, et P. Wilmott (1997) : "Portfolio management with transaction costs," *Proceedings : Math., Phys. and Eng. Sciences*, 453, 551–562.

Bachelier, L. (1900) : *Théorie de la spéculation*. Gauthier-Villars, Paris.

Bai, J. (1997) : "Estimating Multiple Breaks One at a Time," *Econometric Theory*, 13, 315–352.

Bai, J., et P. Perron (1998) : "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66, 47–78.

Baillie, R. (1996) : "Long memory process and fractional integration in econometrics," *Journal of Econometrics*, 73 (1), 5–59.

Baillie, R., T. Bollerslev, et H. Mikkelsen (1996) : "Fractionally integrated generalized autoregressive conditional heteroscedasticity," *Journal of Econometrics*, 74 (1), 3–30.

Baker, H., G. Farrelly, et R. Edelman (1985) : "A Survey of Management Views on Dividend Policy," *Financial Management*, 14, 78–84.

Baker, H., et G. Powell (1999) : "How Corporate Managers View Dividend Policy," *Quarterly Journal of Business and Economics*, 38, 17–35.

Baker, H., G. Powell, et E. Veit (2002) : "Revisiting Managerial Perspectives on Dividend Policy," *Journal of Economics and Finance*, 26 (3), 267–283.

Banerjee, A. (1992) : "A simple model of herd behavior," *Quarterly Journal of Economics*, 107 (3), 797–817.

——— (1993) : "The economics of rumours," *Review of Economic Studies*, 60, 309–327.

Banerjee, A., et G. Urga (2005) : "Modelling structural breaks, long memory and stock market volatility : an overview," *Journal of Econometrics*, 129, 1–34.

Barberis, N., et R. Thaler (2003) : "A survey of behavioral finance," in *Handbook of the Economics of Finance*, ed. par G. M. Constantinides, M. Harris, et R. M. Stulz, vol. I. Elsevier Science.

Barndorff-Nielsen, O., et N. Shephard (2001) : "Non-Gaussian Ornstein-Ulhenbeck-based Models and Some of Their Uses in Financial Economics," *J. Royal Stat. Soc. B*, 63, 167–241.

——— (2002a) : "Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, 253–280.

——— (2002b) : "Estimating Quadratic Variation Using Realized Volatility," *Journal of Applied Econometrics*, 17, 457–477.

Basak, S., et A. Shapiro (2001) : "Value-at-risk-based risk management : optimal policies and asset prices," *Review of Financial Studies*, 14, 371–405.

Bechu, T., E. Bertrand, et J. Nebenzahl (2008) : *L'analyse technique : Théories et méthodes (6è étidion)*. Economica.

Bell, D. (2007) : *The Malliavin Calculus*. Dover.

Benartzi, S., R. Michaely, et R. Thaler (1997) : "Do Changes in Dividend Signal the Future or the Past ?," *Journal of Finance*, 52, 1007–1034.

Benth, F., K. Karlsen, et K. Reikvam (2003) : "Merton's portfolio optimization problem in a Black and Scholes market with non-Gaussian stochastic volatility of Ornstein-Uhlenbeck type," *Mathematical Finance*, 13 (2), 215–244.

Beran, J. (1994) : *Statistics for Long-Memory Processes*. Chapman & Hall.

Bhattacharya, S. (1979) : "Imperfect Information, Dividend Policy, and "the Bird in the Hand Fallacy"," *Bell Journal of Economics*, 10, 259–270.

Biagini, F., Y. Hu, B. Oksendal, et A. Sulem (2002) : "A stochastic maximum principle for processes driven by fractional Brownian motion," *Stochastic Processes and their applications*, 100, 233–253.

Bingham, N., C. Goldie, et J. Teugels (1989) : *Regular variations*. Cambridge University Press.

Bjerring, J., J. Lakonishok, et T. Vermaelen (1983) : "Stock prices and financial analysts' recommendations," *The Journal of Finance*, 38 (1), 187–204.

Black, F. (1976) : "Studies of stock price volatility changes," in *Proceedings of the 1976 Meetings of the Business and Statistical Section*, pp. 177–181.

Black, F., et M. Scholes (1973) : "The pricing of options and corporate liabilities," *Journal of Political Economy*, 83, 637–654.

Bollen, B., et B. Inder (2002) : "Estimating daily volatility in financial markets utilizing intraday data," *Journal of Empirical Finance*, 9, 551–562.

Bollerslev, T. (1986) : "Generalized AutoRegressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31, 307–327.

Bos, C., P. Franses, et M. Ooms (2002) : "Inflation, forecast intervals and long memory regression models," *International Journal of Forecasting*, 18, 243–264.

Box, G., et G. Jenkins (1970) : *Time Series Analysis : Forecasting and control*. Holden-Day, San Fransisco.

Brady, N. (1988) : "Report of the Presidential Task Force on Market Mechanisms," U.S. Government Printing Office.

Breeden, D. (1979) : "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics*, 7, 265–296.

Brennan, M., et y. Xia (2002) : "Dynamic Asset Allocation under Inflation," *Journal of Finance*, 57, 1201–1238.

Brock, W., et C. Hommes (1997) : "A rational route to randomness," *Econometrica*, 65 (5), 1059–1095.

Brockwell, P., et R. Davis (1990) : *Times Series : Theory and Methods*. Springer-Verlag, second edition.

Calvet, L., et E. Common (2003) : "Behavioral Heterogeneity and the income effect," *The Review of Economics and Statistics*, 85 (3), 653–669.

Campbell, J., A. Lo, et C. MacKinlay (1997) : *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

Campbell, J., et L. Viceira (1999) : "Consumption and portfolio decisions when expected returns are time varying," *Quarterly Journal of Economics*, 114, 433–495.

Chacko, G., et L. Viceira (2005) : "Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets," *Review of Financial Studies*, 18, 1369–1402.

Chaudhuri, K., et Y. Wu (2003) : "Random walk versus breaking trend in stock prices : Evidence from emerging markets," *Journal of Banking & Finance*, 27, 575–592.

Chen, W., et R. Deo (2006) : "Estimation of mis-specified long memory models," *Journal of Econometrics*, 134 (1), 257–281.

Cheung, Y., et F. Diebold (1994) : "On maximum likelihood estimation of the differencing parameter of fractionally integrated noise with unknown mean," *Journal of Econometrics*, 62, 301–316.

Cheung, Y., et K. Lai (1995) : "A search for long memory in international stock market returns," *Journal of International Money and Finance*, 14, 597–615.

Chou, R. (1988) : "Volatility persistence and stock valuations : some empirical evidence using GARCH," *Journal of Applied Econometrics*, 3, 279–294.

Chow, K., K. Denning, S. Ferris, et G. Moronha (1995) : "Long-term and short-term price memory in the stock market," *Economics Letters*, 49, 287–293.

Christie, A. (1982) : "The stochastic behavior of common stock variances : value, leverage and interest rate effects," *Journal of Financial Economics*, 10, 407–432.

Clark, P. (1973) : "A subordinated stochastic process model with finite variance for speculative prices," *Econometrica*, 41, 135–155.

Cont, R. (2001) : "Empirical properties of asset returns : stylized facts and statistical issues," *Quantitative Finance*, 1, 223–236.

Cont, R., et J. Bouchaud (2000) : "Herd behaviour and aggregate fluctuations in financial markets," *Macroeconomic Dynamics*, 4, 170–195.

Corcos, A., J.-P. Eckmann, A. Malaspinas, Y. Malevergne, et D. Sornette (2002) : "Imitation and contrarian behavior : hyperbolic bubbles, crashes and chaos," *Quantitative Finance*, 2, 264–281.

Cover, T. (1991) : "Universal Portfolios," *Mathematical Finance*, 1 (1), 1–29.

Cox, J., et C.-F. Huang (1989) : "Optimal consumption and portfolio policies when asset prices follow a diffusion process," *Journal of Economic Theory*, 49, 33–83.

——— (1991) : "A variational problem arising in financial economics," *Journal of Mathematical Economics*, 20, 465–487.

Cox, J., J. Ingersoll, et S. Ross (1985) : "The Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–408.

Dacoragna, M., U. Müller, R. Nagler, R. Olsen, et O. Pictet (1993) : "A geographical model for the daily and weekly seasonal volatility in the foreign exchane market," *Journal of International Money and Finance*, 12, 413–438.

Dacorogna, M., R. Gençay, U. Müller, R. Olsen, et O. Pictet (2001) : *An introduction to high frequency finance*. Academic Press, San Diego, CA.

Dacorogna, M., U. Müller, R. Olsen, et O. Pictet (2001) : "Defining efficiency in heterogeneous markets," *Quantitative Finance*, 1 (2), 198–201.

Dai, W., et C. Heyde (1996) : "Itøformula with respect to fractional Brownian motion and its application," *Journal of Applied Mathematics and Stochastic Analysis*, 9, 439–448.

Davidson, J., et P. Sibbersten (2005) : "Generating schemes for long memory processes : regimes, aggregation and linearity," *Journal of Econometrics*, 128 (2), 253–282.

Davies, P., et M. Canes (1978) : "Stock prices and the publication of second-half information," *Journal of Business*, 51, 43–56.

Davis, M., et A. Norman (1990) : "Portfolio selection with transaction costs," *Mathematics of Operations Research*, 15, 676–713.

de Peretti, C. (2003) : "Computational Econometrics for the Long Memory in Financial Time Series," Ph.D. thesis, Université de la méditerranée, Marseille, Unpublished Doctoral Dissertation.

DeAngelo, H., L. DeAngelo, et S. Skinner (1992) : "Dividends and losses," *Journal of Finance*, 47, 1837–1863.

Dellacherie, C., et P. Meyer (1982) : *Probabilities and potentials B*. North-Holland, Amsterdam.

Deo, R., C. Hurvich, et Y. Li (2006) : "Forecasting realized volatility using a long-memory stochastic volatility model : estimation, prediction and seasonal adjustment," *Journal of Econometrics*, 131, 29–58.

Di Matteo, T., T. Aste, et M. Dacorogna (2005) : "Long-term memories of developed and emerging markets : Using the scaling analysis to characterize their stage of development," *Journal of Banking and Finance*, 29, 827–851.

Diebold, F., et A. Inoue (2001) : "Long memory and regime switching," *Journal of Econometrics*, 105, 131–159.

Diebold, F., et G. Rudebusch (1989) : "Long memory and persistence in aggregate output," *Journal of Monetary Economics*, 24, 189–209.

——— (1991a) : "Is consumption too smooth ? Long memory and the Deaton paradox," *Review of Economics and statistics*, 73, 1–9.

Ding, Z., C. Granger, et R. Engle (1993) : "A long memory property of stock market returns and a new model," *Journal of Empirical Finance*, 1, 83–106.

Dittman, I., et C. Granger (2002) : "Properties of nonlinear transformations of fractionally integrated processes," *Journal of Econometrics*, 110, 113–133.

Dodd, R., et P. Mills (2008) : "Crise hypothécaire : effet de contagion," *Finances & Développement*, 45 (2), 14–18.

Dorsey, R., et W. Mayer (1995) : "Genetic Algorithms for Estimation Problems with Multiple Optima, Non-Differentiability, and other Irregular Features," *Journal of Business and Economic Statistics*, 13, 53–66.

Duffie, D., et L. Epstein (1992) : "Stochastic differential utility," *Econometrica*, 60 (2), 353–394.

Duncan, T., Y. Hu, et B. Pasik-Duncan (2000) : "Stochastic calculus for fractional Brownian motion. I. Theory," *SIAM Journal on Control and Optimization*, 38, 582–612.

Eagles, J. (1994) : "The relationship between mood and daily hours of sunlight in rapid cycling bipolar illness," *Biological Psychiatry*, 36, 422–424.

Ellis, C., et P. Wilson (2004) : "Another look at the forecast performance of ARFIMA models," *International Review of Financial Analysis*, 13, 63–81.

Engle, R. (1982) : "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, 50 (4), 987–1007.

Engle, R., et A. Smith (1999) : "Stochastic permanent breaks," *Review of Economics and statistics*, 81, 553–574.

Epstein, L., et S. Zin (1989) : "Substitution, risk aversion and the temporal behavior of consumption and asset returns : a theoretical framework," *Econometrica*, 57, 937–969.

Fama, E. (1965) : "The behavior of stock-market prices," *Journal of Business*, 38, 34–105.

Fama, E., L. Fisher, M. Jensen, et R. Roll (1969) : "The Adjustment of Stock Prices to New Information," *International Economic Review*, 10 (1), 1–21.

Fama, E., et K. French (1988) : "Permanent and temporary components of stock prices," *Journal of Political Economy*, 96, 246–273.

Fouque, J.-P., G. Papanicolaou, et K. Sircar (2000) : *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, Cambridge.

Froot, K., D. Scharfstein, et J. Stein (1992) : "Herd on the street : Informational inefficiencies in a market with short-term speculation," *Journal of Finance*, 47 (4), 1461–1484.

Gabaix, X., P. Gopikrishnan, V. Plerou, et H. Stanley (2006) : "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics*, 121 (2), 461–504.

Gadea, M., M. Sabaté, et J. Serrano (2004) : "Structural breaks and their trace in the memory inflation rate series in the long-run," *Journal of International Financial Markets, Institutions & Money*, 14, 117–134.

Garman, M., et M. Klass (1980) : "On the estimation of security price volatilities from historical data," *Journal of Business*, 53, 67–78.

Geweke, J., et S. Porter-Hudak (1983) : "The estimation and application of long-memory time series model," *Journal of Time Series Analysis*, 4, 221–238.

Ghoulmie, F., R. Cont, et J.-P. Nadal (2005) : "Heterogeneity and feedback in an agent-based market model," *Journal of Physics : Condensed Matter*, 17, S1259–S1268.

Golberg, D. (1989) : *Genetic Algorithm in Search, Optimization and Learning*. Addison-Wesley.

Gonçalves, E., et C. Gouriéroux (1988) : "Agrégation de processus autorégressifs d'ordre 1," *Annales d'Economies et de Statistique*, 12, 127–149.

Gouriéroux, C., et J. Jasiak (1998) : "Truncated Maximum Likelihood, Goodness of Fit Tests and Tail Analysis," Working paper.

——— (2001) : "Memory and infrequent breaks," *Economics Letters*, 70, 29–41.

Granger, C. (1980) : "Long memory relationships and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227–238.

Granger, C., et N. Hyung (2004) : "Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns," *Journal of Empirical Finance*, 11, 399–421.

Granger, C., et R. Joyeux (1980) : "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis*, 1, 15–30.

Granger, C., et M. Morris (1976) : "Times series modeling ant interpretation," *Journal of the Royal Statistical Society A*, 38, 246–257.

Granger, C., et P. Newbold (1977) : *Forecasting economic time series*. Academic Press, New York.

Grauer, R., et N. Hakansson (1993) : "On the use of mean-variance and quadratic approximations in implementing dynamic investment strategies A comparison of the returns and investment policies," *Manag. Sci.*, 39, 856–871.

Hakansson, N. (1971) : "Capital growth and the mean-variance approach to portfolio management," *J. Financial Quant. Anal.*, 6, 517–557.

Hamilton, J. (1989) : "A new approach to the economic analysis of nonstationary time series and the business cycle," *Econometrica*, 57, 357–384.

Hansen, P., et A. Lunde (2002) : "Volatility estimation using high frequency data with partial availability," Discussion Paper, Brown University Working Paper.

——— (2005) : "A forecast comparison of volatility models : does anything beat a GARCH(1,1) ?," *Journal of Applied Econometrics*, 20 (7), 873–889.

Harvey, A. (1993) : "Long memory in stochastic volatility," Working Paper, London School of Economics.

Hassler, U., et J. Wolters (1995) : "Long memory in inflation rates : international evidence," *Journal of Business and Economic Statistics*, 13, 37–45.

He, H., et N. Pearson (1991) : "Consumption and portfolio policies with incomplete markets and short-sale constraints : The infinite dimensional case," *Journal of Economic Theory*, 54 (2), 259–304.

Heston, S. (1993) : "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

Hirshleifer, D., et T. Shumway (2003) : "Good day sunshine : stock returns and the weather," *The Journal of Finance*, 58 (3), 1009–1032.

Hog, E., et P. Frederiksen (2006) : "The fractional Orstein-Uhlenbeck process : term structure theory and application," Working paper.

Hommes, C. (2008) : "Interacting agents in finance," in *The New Palgrave Dictionary of Economics*, ed. par S. Durlauf, et L. Blume. Palgrave Macmillan, Préface.

Hosking, J. (1981) : "Fractional differencing," *Biometrika*, 68, 165–176.

——— (1984) : "Modelling Persistence in Hydrological Time Series Using Fractional Differencing," *Water Resources Research*, 20, 1898–1908.

——— (1996) : "Asymptotic distribution of the sample mean, autocovariances, and autocorrelations of long-memory time series," *Journal of Econometrics*, 73, 261–284.

Hu, Y., et B. Oksendal (1999) : "Fractional white noise calculus and applications to finance," *Preprint, Department of Mathematics, University of Oslo*, 10.

Hurst, H. (1951) : "Long Term Storage Capacity of Reservoirs," *Transaction of the American Society of Civil Engineers*, 116, 770–799.

——— (1957) : "A suggested statistical model of some time series which occur in nature," *Nature*, 180, 494–494.

Hyung, N., et P. Franses (2001) : "Structural breaks and long-memory in US inflation rates : do they matter for forecasting ?," Working Paper, Erasmus University.

Ingersoll, J. J. (1987) : *Theory of financial decision making*. Rowman & Littlefield, Savage, Maryland.

John, K., et J. Williams (1985) : "Dividends, Dilution and Taxes : A signaling Equilibrium," *Journal of Finance*, 40, 1053–1070.

Jones, C., G. Kaul, et M. Lipson (1994) : "Transactions, Volume and Volatility," *Review of Financial Studies*, 7, 631–651.

Jouini, E., P.-F. Koehl, et N. Touzi (1999) : "Optimal Investment with Taxes : An Optimal Control Problem with Endogeneous Delay," *Nonlinear Analysis*, 37, 31–56.

Karatzas, I., J. Lehoczky, et S. Shreve (1987) : "Optimal portfolio and consumption decisions for a "small investor" on a finite horizon," *SIAM Journal on Control and Optimization*, 25, 1557–1586.

Karatzas, I., et S. Shreve (1991) : *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, second edn.

Kearney, C., et K. Daly (1998) : "The causes of stock market volatility in Australia," *Applied Financial Economics*, 8, 597–605.

Kim, M., C. Nelson, et R. Startz (1991) : "Mean reversion of stock prices : A reappraisal of the empirical evidence," *Review of Economic Studies*, 58, 515–529.

Kirman, A. (1992) : "Whom or What Does the Representative Individual Represent ?," *Journal of Economic Perspectives*, 6 (2), 117–136.

——— (1993) : "Ants, rationality and recruitment," *Quarterly Journal of Economics*, 108, 135–156.

Kolmogorov, A. (1940) : "Wienersche Spiralen und einige andere interessante Kurven in Hilbertschen Raum," *Comptes Rendus (Doklady) Acad. Sci. USSR (N.S)*, 26, 115–118.

Koopman, S., B. Jungbacker, et E. Hol (2005) : "Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements," *Journal of Empirical Finance*, 12, 445–475.

Krugman, P. (2007) : "Workouts, Not Bailouts," *The New York Times*, 17/08/2007.

Kumar, P., et B.-S. Lee (2001) : "Discrete dividend policy with permanent earnings," *Financial Management*, 30, 55–76.

Kumar, P., et D. Seppi (1992) : "Futures manipulation with 'cash settlement'," *Journal of Finance*, 47 (4), 1485–1502.

Kyle, A. (1985) : "Continuous Auction and Insider Trading," *Econometrica*, 53 (6), 1315–1336.

Lagarias, J., J. Reeds, M. Wright, et P. Wright (1998) : "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," *SIAM Journal on Optimization*, 9 (1), 112–147.

Latané, H., et R. Rendleman (1976) : "Standard Deviations of Stock Price Ratios Implied in Option Prices," *Journal of Finance*, 31, 369–381.

Le Courtois, O. (2003) : "Modelling Stock Returns with Levy Process," *Banque & Marchés*, 66.

Leland, H. (1999) : "Optimal Portfolio Management in the Presence of Transactions Costs and Capital Gains Taxes," Working paper No. 290, Research Program in Finance, IBER.

Li, D., et W.-L. Ng (2000) : "Optimal dynamic portfolio selection : Multi-period mean-variance formulation," *Math. Finance*, 10 (3), 387–406.

Lim, A., et X. Zhou (2002) : "Dynamic mean-variance portfolio selection with no-shorting constraints," *Math. Oper. Res.*, 27, 101–120.

Lin, S. (1995) : "Stochastic analysis of fractional Brownian motions," *Stochastics and Stochastics Reports*, 55, 121–140.

Lintner, J. (1956) : "Distribution of Incomes of Corporations Among Dividends, Retained Earnings and Taxes," *American Economic Review*, 46, 97–113.

Liu, J. (2007) : "Portfolio selection in stochastic environments," *The Review of Financial Studies*, 20 (1), 1–39.

Liu, M. (2000) : "Modeling long memory in stock market volatility," *Journal of Econometrics*, 99, 139–171.

Liu, P. (1990) : "Stock price reactions to The Wall Street Journal's securities recommendations," *Journal of Financial and Quantitative Analysis*, 25 (3), 399–410.

Lo, A. (1991) : "Long term memory in stock market prices," *Econometrica*, 59, 1279–1313.

Lo, A., et A. MacKinlay (1988) : "Stock market prices do not follow random walks : Evidence from a simple specification test," *Review of Financial Studies*, 1, 41–66.

Lobato, I., et N. Savin (1998) : "Real and spurious long-memory properties of stock-market data," *Journal of Business and Economic Statistics*, 16, 261–268.

Lobato, I., et C. Velasco (2004) : "A simple test of normality for time series," *Econometric Theory*, 20, 671–689.

Lévy, P. (1948) : *Processus stochastiques et mouvement brownien*. Gauthier-Villars, Paris.

Magill, M., et G. Constantinides (1976) : "Portfolio selection with transactions costs," *The Journal of Economic Theory*, 13, 245–263.

Malevergne, Y., et D. Sornette (2005) : *Extreme Financial Risks : From dependence to risk management*. Springer.

Malliavin, P., et A. Thalmaier (2005) : *Stochastic Calculus of Variations in Mathematical Finance*. Springer.

Man, K. (2003) : "Lon memory time series and short term forecast," *International Journal of Forecasting*, 19, 477–491.

Mandelbrot, B. (1963) : "The variation of certain speculative prices," *Journal of Business*, 36, 394–419.

——— (1972) : "Statistical Methodology for Non-Periodic Cycles : From the covariance to R/S analysis," *Annals of Economic and Social Measurement*, 1, 259–299.

——— (1975) : "Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31, 271–285.

Mandelbrot, B., et M. Taqqu (1979) : "Robust R/S Analysis of Long-Run Serial Correlation," *Bulletin of the International Statistical Institute*, 48 (2), 59–104.

Mandelbrot, B., et J. Van Ness (1968) : "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, 10, 422–437.

Mandelbrot, B., et J. Wallis (1968) : "Noah, Joseph and operational hydrology," *Water Resources Research*, 4, 909–918.

Mankiw, N. (2000) : "The Savers-Spenders Theory of Fiscal Policy," *American Economic Review*, 90 (2), 120–125.

Markowitz, M. (1952) : "Portfolio Selection," *The Journal of Finance*, 7, 77–91.

Martens, M. (2002) : "Measuring and forecasting S&P500 index-futures volatility using high frequency data," *Journal of Future Markets*, 22, 497–518.

McLeod, A., et K. Hipel (1978) : "Preservation of the rescaled adjusted range, 1. A reassessment of the Hurst phenomenon," *Water Resources Research*, 14, 159–178.

Merton, R. (1971) : "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3, 373–413.

Michaelsen, J. (1961) : "The Determinants of Dividend Policies : A Theoretical and Empirical Study," Ph.D. thesis, Graduate School of Business, University of Chicago, Unpublished Doctoral Dissertation.

Miller, M., et K. Rock (1985) : "Dividend Policy under Asymmetric Information," *Journal of Finance*, 40, 1031–1051.

Morton, A., et S. Pliska (1995) : "Optimal portfolio management with fixed transaction costs," *Mathematical Finance*, 5, 337–356.

Muthuraman, K., et S. Kumar (2006) : "Multidimensional portfolio optimization with proportional transaction costs," *Mathematical Finance*, 16 (2), 301–335.

Nelder, J., et R. Mead (1965) : "A Simplex Method for Function Minimization," *Computer Journal*, 7, 308–313.

Newey, W., et D. McFadden (1994) : "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, ed. par R. F. Engle, et D. L. McFadden, vol. IV. Elsevier Science.

Nualart, D. (2006) : *The Malliavin Calculus and Related Topics*. Springer, second edn.

Oksendal, B. (1997) : "Introduction to Malliavin Calculus with applications to economics," Working paper, Department of Mathematics, University of Oslo.

——— (2004) : "Fractional Brownian motion in finance," Working paper, Department of Mathematics, University of Oslo.

Orléan, A. (1995) : "Bayesian interactions and collective dynamics of opinion : Herd behaviour and mimetic contagion," *Journal of Economic Behavior and Organization*, 28, 257–274.

——— (1999) : *Le pouvoir de la finance*. Odile Jacob, Paris.

Orléan, A., D. Bourghelle, O. Brandouy, et R. Gillet (2005) : *Représentations collectives et croyances sur les marchés financiers*. Economica, Paris.

Osborne, M. (1959) : "Brownian Motion in the Stock Market," *Operations Research*, 7, 145–173.

Pagan, A., et A. Ullah (1999) : *Nonparametric Econometrics*. Cambridge University Press.

Pari, R. (1987) : "Wall Street week recommendations : Yes or no ?," *Journal of Portfolio Management*, 14, 74–76.

Parkinson, M. (1980) : "The extreme value method for estimating the variance of the rate of return," *Journal of Business*, 53, 61–65.

Peters, E. (1994) : *Fractal Market Analysis*. Wiley, New York.

Pliska, S. (1986) : "A stochastic calculus model of continuous trading : Optimal portfolios," *Math. Oper. Res.*, 11, 371–384.

——— (1997) : *Introduction to Mathematical Finance*. Blackwell Publishers, Oxford.

Poncet, P., et A. Lioui (2001) : "Mean-variance Efficiency of the Market Portfolio and Futures Trading," *Journal of Futures Markets*, 21 (4), 329–346.

Portait, R., I. Bajeux-Besnainou, et J. Jordan (2001) : "An Asset Allocation Puzzle : Comment," *American Economic Review*, 91 (4), 1170–1179.

Poterba, M., et L. Summers (1988) : "Mean reversion in stock prices : Evidence and implications," *Journal of Financial Economics*, 22, 27–59.

Price, K., et R. Storn (1997) : "Differential Evolution," *Dr. Doob's Journal*, April, 18–24.

Quittard-Pinon, F. (2003) : *Marchés des capitaux et théorie financière*. Economica, Paris.

Resnick, S., et G. Samorodnitsky (1997) : "Performance Decay in a Single Server Exponential Queuing Model with Long Range Dependence," *Operations Research*, 45 (2), 235–243.

Richardson, M. (1993) : "Temporary components of stock prices : A skeptic's view," *Journal of Business and Economic Statistics*, 11, 199–207.

Richardson, M., et J. Stock (1989) : "Drawing inference from statistics based on multiyear asset returns," *Journal of Financial Economics*, 25, 323–348.

Rind, B. (1996) : "Effects of beliefs about weather conditions on tipping," *Journal of Applied Social Psychology*, 26, 137–147.

Robinson, P. (1995) : "Log-periodogram regression of time series with long range dependence," *The Annals of Statistics*, 23, 1048–1072.

Roll, R. (1984) : "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *The Journal of Finance*, 39, 1027–1039.

Roy, A. (1952) : "Safety-First and the Holding of Assets," *Econometrica*, 20 (3), 431–449.

Sadorsky, P. (2003) : "The macroeconomic determinants of technology stock price volatility," *Review of Financial Economics*, 12, 191–205.

Samuelson, P. (1965) : "Proof That Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, Spring, 41–49.

——— (1986) : "Lifetime portfolio selection by dynamic stochastic programming," *The Review of Economics and Statistics*, 51, 371–382.

Schwert, G. (1989) : "Why does stock market volatility change over time ?," *Journal of Finance*, 44, 1115–1153.

Shiller, R. (1989) : *Market Volatility*. MIT Press.

——— (2000) : *Irrational Exuberance*. Princeton University Press, Princeton, NJ.

Simon, H. (1957) : *"A Behavioral Model of Rational Choice", in Models of Man, Social and Rational : Mathematical Essays on Rational Human Behavior in a Social Setting*. Wiley, New York.

Simon, P. (1991) : "Bounded Rationality and Organizational Learning," *Organization Science*, 2 (1), 125–134.

Sornette, D. (2003) : *Why Stock Markets Crash ?* Princeton University Press, Princeton, NJ.

Sornette, D., Y. Malevergne, et J. Muzi (2003) : "What causes crashes ?," *Risk*, 16 (2), 67–71.

Sowell, F. (1992) : "Maximum likelihood estimation of stationary univariate fractionally integrated time series models," *Journal of Econometrics*, 53, 165–188.

Taskar, M., M. Klass, et D. Assaf (1988) : "A Diffusion Model for Optimal Portfolio Selection in the Presence of Brokerage Fees," *Math. Oper. Res.*, 13, 277–294.

Taylor, S. (1986) : *Modelling Financial Time Series*. John Wiley, Chichester.

Teverovsky, V., M. Taqqu, et W. Willinger (1999) : "A critical look at Lo's modified R/S statistic," *Journal of Statistical Planning and Inference*, 80, 211–227.

Thaleb, N. (2001) : *Fooled by randomness*. Random House.

Tietjen, G., et D. Kripke (1994) : "Suicides in California (1968-1977) - Absence of seasonality in Los Angeles and Sacramento counties," *Psychiatric Research*, 53, 161–172.

Trahan, E., et P. Bolster (1995) : "The impact of Barron's recommendations on stock prices," *Quarterly Journal of Business and Economics*, 34 (4), 3–15.

Van der Vaart, A. (2000) : *Asymptotic Statistics*. Cambridge University Press.

Van Dijk, D., P. Franses, et R. Paap (2002) : "A nonlinear long-memory model, with an application to US unemployment," *Journal of Econometrics*, 110, 135–165.

White, E. (1996) : *Stock market crashes and speculative manias*. Edward Elgar Publishing, Cheltenham, UK.

Wiener (1923) : "Differential space," *Journal of Mathematical Physics*, 2, 131–174.

Zhou, X., et D. Li (2000) : "Continuous time mean-variance portfolio selection : A stochastic LQ framework," *Appl. Math. Optim.*, 42, 19–33.

Zhou, X., et G. Yin (2003) : "Markowitz's mean-variance portfolio selection with regime switching : A continuous time model," *SIAM J. Contr. Optim.*, 42, 1466–1482.