HAL
open science

Investisseurs et Marchés Financiers : du comportement des agents à la formation de prix d'équilibre

Mamadou Konté

To cite this version:

Mamadou Konté. Investisseurs et Marchés Financiers : du comportement des agents à la formation de prix d'équilibre. Economies et finances. Université Panthéon-Sorbonne - Paris I, 2010. Français. ⟨NNT : 2 0 1 0 P A 0 1 0 0 2 5⟩. ⟨tel-00618863⟩

HAL Id: tel-00618863

https://theses.hal.science/tel-00618863v1

Submitted on 5 Sep 2011

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.

Autorisation HAL

HAL Authorization

Pôle Banque-Finance École doctorale C.E.S (Centre d'Economie de la Sorbonne.)

Investisseurs et Marchés Financiers : Du comportement des agents à la formation de prix d'équilibre.

Numéro d'identification bibliothèque : 2 0 1 0 P A 0 1 0 0 2 5

THÈSE

présentée et soutenue publiquement le 28 Juin 2010

pour l'obtention du

Doctorat de l'université Panthéon Sorbonne – Paris 1

(spécialité Sciences Economiques)

par

Mamadou Konté.

Composition du jury

Président : **Dominique Guégan**, Professeure à l' université Panthéon Sorbonne.

Rapporteurs : **Hélène Raymond**, Professeure à l'Université de Paris Ouest Nanterre.

Laurent Augier, Maître de Conférence à l'Université de la Rochelle.

Directeur : **Thierry Chauveau**, Professeur à l' université Panthéon Sorbonne.

Mis en page avec la classe thloria.

Remerciements

Je voudrais, en premier lieu, exprimer toute ma gratitude à Thierry Chauveau, mon directeur de Thèse, qui a bien voulu m'accepter sans avoir une bourse d'allocation ministérielle. Ce qui n'est pas toujours évident. Cette rencontre a été possible grâce à Dominique Guégan que je tiens également à remercier. Au début, on s'est intéressé à la loi du prix unique et de sa relation avec les mesures martingales signées. Même si cela n' a pas abouti, cette voie nous a mené vers le débat qui existe entre la finance néo-classique et comportementale. Dès lors, le sujet de la thèse s'est plus penché sur la formation des prix d'équilibre. Je remercie mon directeur de la liberté et de la confiance qu'il a su m' accorder pour que je puisse aller au bout de mes idées.

Une thèse, avant tout, c'est une formation. J'ai beaucoup appris des séminaires de l'axe finance co-organisés par Thierry Chauveau et Dominique Guégan. C'est pourquoi, je voudrais remercier tous les membres de cet axe qui ont participé de loin ou de près à la bonne animation de cet événement mensuel si important pour les doctorants. J'ai eu des notions de modèles à agents à travers ces séminaires et que j'ai eu à consacrer dans la deuxième partie de ma thèse.

D'autre part, je voudrais remercier tous mes enseignants aussi bien en informatique qu'en mathématiques de l'université Gaston Berger (Sénégal) en particulier Aliou Diop de m' avoir beaucoup soutenu dans les formalités de bourse. Ils ont su me donner le goût des mathématiques appliquées m' ayant permis de s'investir dans une thèse.

Je n'oublie pas mes collègues doctorants de la Salle 419 puis 513 après déménagement. On a passé de bons moments chaleureux même si je suis de nature introverti. Je remercie particulièrement Ghislain pour les discussions fructueuses qu'on a eu ensemble au cours de ces dernières années.

1- ***In memoriam***

- *À mon défunt père Mamadou Konté que je n'ai pas eu la chance de connaître.*
- *À ma mère adorée Ramatoulaye Sow dont je suis indéfiniment redevable.*
- *À mes grands parents.*

2- À toute ma famille.

Table des matières

Introduction générale et Résultats. **ix**

Première partie :
Finance Néo-classique
&
Comportementale.

Chapitre 1
Finance Néo-classique. **1**

1.1 Rationalisation des préférences. 1
1.2 Le critère d'arbitrage. 6
1.3 Efficience de marché. 12

Chapitre 2
Finance Comportementale.

2.1 Modèles quasi-normatifs. 18
2.2 Psychologie. 22
2.2.1 Etudes expérimentales/empiriques. 24
2.2.2 Modèle de Kahneman et Tversky. 28
2.3 Limite du critère d'arbitrage. 31
2.3.1 Exemples de Modèle. 32
2.3.2 Etudes empiriques. 35

Chapitre 3
Quel paradigme choisir ?

3.1 Tests basés sur la marche aléatoire. 40
3.1.1 Tests d'auto-corrélation. 40
3.1.2 Tests sur le ratio de variance. 45
3.1.3 Tests de changements de signe (*Run tests*). 49

3.2 Tests basés sur une sélection de portefeuilles. 50
3.3 Discussion. 55

Deuxième Partie :
Modèles
A base
D'agents virtuels.

Chapitre 4
Structure des modèles à agents.

4.1 Description des agents. 71
4.2 Mécanisme de marché. 72
4.3 Apprentissage. 74
4.3.1 Stratégies fixes. 75
4.3.2 Stratégies évolutionnistes. 75
4.4 Critères de validation. 78

Chapitre 5
Tentative de réconciliation des deux paradigmes.

5.1 Description et prix d'équilibre. 85
5.2 Extension et Commentaires. 99

Troisième Partie :
Applications.

Chapitre 6
Stratégies d'investissement.

6.1 Description de la méthodologie 103
6.2 Résultats 108

Chapitre 7
Prédiction de la volatilité.

7.1 Introduction 119
7.2 Quelques modèles GARCH asymétriques. 120
7.2.1 La classe SUGARCH. 120
7.2.2 Modèles de Comparaison. 125
7.3 Données et Méthodologie. 127
7.4 Résultats empiriques. 131

Chapitre 8
Conclusion et discussion

Bibliographie **143**

Introduction générale et Résultats.

Un des secteurs de première importance, dans le domaine l'économie, est celui des marchés financiers. Ils constituent généralement à travers leurs indices (CAC 40, Dow Jones, etc.) un bon indicateur de la santé d'une économie. Dans ces marchés, des personnes se rencontrent pour acheter ou vendre des titres financiers tels que des actions, des obligations ou des produits dérivés. Le prix de ces titres varie de façon aléatoire selon la loi de l'offre et de la demande. Pour bien mener une étude sur ces marchés, il est donc utile de s'intéresser aux principaux acteurs à savoir les investisseurs. On peut se demander entre autres comment les agents font leurs décisions dans un cadre d'incertitude. Ensuite, après avoir décrit ou axiomatiser leur comportement, obtient-on au niveau agrégé des prix d'équilibre qui sont conformes aux données financières. On revient essentiellement sur ces deux questions qui ont été beaucoup étudiées dans la littérature. Actuellement, on distingue trois mouvances : la finance néo-classique ou finance traditionnelle, la finance comportementale et enfin la finance artificielle. La première partie de ce document est consacrée aux deux premières approches. Dans la finance néo-classique, on a essentiellement les points suivants :

1. Rationalité des investisseurs.
   C'est le point fondamental de cette théorie qui stipule que les décisions des agents respectent la norme. On entend par norme toute décision vérifiant les axiomes de Von Neumann Morgersten [111] ou de Savage [103]. D'autre part, si de nouvelles informations arrivent, les agents sont capables de les intégrer de façon correcte. En d'autres termes, ils savent tous utiliser la formule de Bayes.

2. Irrationalité d'investisseurs avec des actions non corrélées.
   Demander une rationalité totale pour tous les investisseurs est une hypothèse assez forte pour une théorie. C'est pourquoi il est supposé que certains investisseurs sont irrationnels mais avec des actions non corrélées. Dans ce cas, on obtient par agrégation, que l'effet des agents rationnels.

3. Présence d'arbitragistes.
   Le point 2 est motivé par le fait que les erreurs humaines sont de nature multiformes et donc par agrégation on s'attend à ce que celles-ci se compensent. Au cas où les actions des agents irrationnels deviendraient corrélées, alors les arbitragistes interviennent pour corriger les anomalies.

Avec ces 3 hypothèses et sous l'hypothèse de marchés sans friction, on obtient un marché efficient au sens informationnel qui stipule qu'on ne peut utiliser l'information disponible pour prédire les rentabilités ou les prix futurs. Un autre résultat qui en découle, sous certaines conditions, est le fait que les prix des actifs reflètent leurs valeurs fondamentales c'est-à-dire l'espérance actualisée de leurs dividendes futurs.

La révolution dans cette théorie a été le point 3 qui permet de dire qu'il y a absence d'opportunité d'arbitrage dans un marché concurrentiel sans friction. Cette condition permet d'obtenir des prix d'équilibre en demandant simplement aux agents de vouloir plus que moins.

Les implications de cette théorie d'efficience de marché ont été beaucoup testées sous ses trois formes très répandues à savoir l'efficience au sens faible, au sens semi-fort ou au sens fort. Dans l'efficience au sens faible, on stipule que l'on ne peut utiliser l'information passée (historique de prix, de taux d'intérêt, de volatilité, de volumes de transactions etc.) pour avoir un excès de rentabilité. DeBondt et Thaler [16] ont montré une violation de cette efficience en utilisant les rentabilités passées. Ils ont classé des actifs en fonction de leur performance passée (gagnants, perdants). Ils montrent que les gagnants dans le passé récent tendent à devenir perdants à long terme et réciproquement les perdants tendent à être gagnants entraînant ainsi un effet renversement. Toujours dans l'efficience au sens faible, Jegadeesh et Titman [59] ont montré quant à eux un effet *momentum* qui stipule toujours en considérant la classification précédente que les gagnants vont conserver leurs positions à court terme et pareillement pour les perdants.
Dans la théorie de l'efficience au sens semi-fort, on stipule qu'on ne peut utiliser l'information publique pour battre le marché. Cette information contient à la fois l'information passée et les informations liées aux entreprises. Banz [9] trouve une violation de cette efficience en considérant la caractéristique capitalisation des entreprises. Il note que la prime de risque associée aux petites firmes est très élevée par rapport à leur facteur béta (β). Dans cette même lancée, Fama et French [38] classifient les actifs en fonction de leur capitalisation. Ils remarquent sur une période d'une année, les petites firmes ont une rentabilité supérieure par rapport aux grandes firmes.
D'autres études basées sur les études d'événements corroborent cette présence de prévisibilité. Par exemple, les firmes ayant reçues une surprise positive par rapport à l'événement dividende se comportent mieux par rapport à celles ayant reçues une surprise négative dans le court terme. Des facteurs tels que le ratio valeur comptable sur valeur de marché (B/M), le ratio prix sur bénéfice par action (PER) peuvent être utilisés aussi pour prédire les rentabilités futures des actifs mettant en cause l'efficience au sens semi-fort.

Dès lors, il fallait une alternative à la finance néo-classique pour expliquer les prévisibilités et les excès de volatilité constatés durant la décennie octante. La finance comportementale entre en jeu avec les points suivants :

1. Rationalité des investisseurs.
   Il est supposé la présence d'agents rationnels comme dans la théorie de la finance néo-classique.

2. Irrationalité d'investisseurs avec des actions corrélées.
   Contrairement au point 2) énuméré dans la partie finance néo-classique, les agents irrationnels ont ici des comportements corrélés. En faisant leurs choix devant des problèmes complexes, ils se servent de raisonnements assez simplifiés qui sont souvent identiques (heuristiques). Le domaine de la psychologie est utilisée pour étayer ces faits.

3. Arbitrage Limité.
   La corrélation des actions des agents irrationnels que l'on appellera dès fois bruiteurs fait que les prix des actifs peuvent dévier de leurs valeurs fondamentales. Dans ce cas, le point

3) de la finance traditionnelle stipule que les arbitragistes vont intervenir pour profiter des imperfections et rendre le marché efficient. La finance comportementale répond par le fait que les opportunités d'arbitrage sont risquées et coûteuses en pratique de sorte que les bruiteurs peuvent bel et bien persister dans le marché.

Pour cette théorie, les marchés ne sont pas efficients. Des techniques chartistes ou fondamentalistes peuvent permettre d'avoir un excès de rentabilité. En effet les agents irrationnels créent les anomalies avec la corrélation de leurs actions et cette déviation persiste due au fait qu'il y a une motivation limitée pour les arbitragistes d'exploiter ces erreurs.
Aujourd'hui, cette approche est bien représentée en termes de publications ou de livres. Son avantage vient de sa méthodologie qui est plutôt descriptive que normative. On essaie de comprendre le comportement des agents à travers leurs choix par exemple sur des loteries. Ensuite, les résultats obtenus sont utilisés ultérieurement pour expliquer des puzzles.

Par ailleurs, il faut noter que les phénomènes de prévisibilité, d'excès de volatilité n'entraînent pas automatiquement une présence d'anomalies. Des explications rationnelles sont aussi possibles (prise supplémentaire de risque, présence d'hétérogénéité, etc.). Dès lors, il devient difficile de trancher sur la question de l'efficience. Souvent, cette difficulté est reliée d'abord aux problèmes méthodologiques. Par exemple, pour le cas lié à la prévisibilité des cours ou d'excès de volatilité, les résultats obtenus peuvent disparaître si on passe de données (intra)journalières à des données mensuelles voire annuelles (dû à l'effet agrégation). La difficulté peut provenir aussi de l'hypothèse jointe au sens où on doit spécifier un modèle d'équilibre pour les prix (CAPM, CCAPM, etc.). Si l'hypothèse jointe est rejetée, il peut s'agir d'une inefficience de marché ou simplement d'une mauvaise spécification de modèle.
La lecture qui est faite ici de cette difficulté est que la frontière qui sépare les deux branches est très fine et on peut basculer facilement dans l'un des deux camps. Précisément, notre propos principal sera de montrer que dans la formation des prix d'équilibre, on retrouve à la fois les arguments de la finance néo-classique et comportementale. C'est ce qui explique la difficulté des tests de l'efficience et ceci indépendamment des problèmes méthodologiques et/ou de l'hypothèse jointe. En d'autres termes, les deux théories ont tous une part de vérité et il faut donc chercher un moyen de les réconcilier. Cette idée de réconciliation n'est pas nouvelle. Lo [79] a proposé ce qu'il a appelé *Adaptive Market Hypothesis* qui stipule que la propriété de marché efficient n'est pas une propriété binaire prenant les valeurs 0 ou 1 de façon définitive mais plutôt pouvait varier dans le temps entre les différents marchés et selon l'évolution des conditions macro-économiques.
Une autre approche possible et qui est adoptée ici est de dire que les prix actualisés des actifs ou leurs rentabilités suivent une martingale qu'en moyenne.
La motivation en est que sous l'hypothèse d'agents neutre au risque, il est connu que le processus de rentabilités des titres est une martingale sous la théorie de l'efficience de marché (rationalité des agents, marché sans friction). Cependant, en intégrant le facteur aversion, on peut alors s'en écarter et des corrélations peuvent apparaître, voir LeRoy [76]. Si on suppose que les corrélations qui sont positives et négatives se révèlent de façon équitable, on aura la propriété de martingale en moyenne. Cette assertion est vue comme un argument en faveur de la finance néo-classique ou du moins ne la contredit pas. Par exemple, Leroy [76] fait la remarque suivante " *In the absence of prior information about the sign of the correlation coefficient relating successive rates of return, it is clearly reasonable to conclude that the empirical evidence in favor of the martingale is at least not inconsistent with theory*".
Fama [37] va également dans le même sens, par rapport aux anomalies liées aux rentabilités à

long terme des titres, en disant que "*But in an efficient market, apparent under-reaction will be about as frequent as over-reaction. If anomalies split randomly between under- and over-reaction, they are consistent with market effciency.* "

Nous montrons aussi que cette assertion ne discrédite pas non plus les arguments de la finance comportementale voir Konté [69]. C'est ce qui explique la difficulté du débat qui va au delà de l'hypothèse jointe et des méthodologies utilisées. Pour illustrer ce point, on considère un modèle de marché avec une violation de la loi du prix unique. Formellement, le marché dispose de trois actifs. Les deux premiers notés A et B sont risqués et identiques et le troisième noté C est sans risque. Pour créer une anomalie, on suppose que le marché est segmenté. Il existe des investisseurs de type A et B qui investissent dans les actifs A, C respectivement B et C. Ensuite, on met des hypothèses différentes sur les investisseurs de type A et B de sorte que les prix des actifs A et B puissent être différents. Par exemple, on peut supposer que les A-investisseurs sont rationnels tandis que les B-investisseurs sont divisés en deux groupes : agents rationnels et irrationnels. Comme conséquence de ces suppositions, l'actif A reflète toujours sa valeur fondamentale alors que l'actif B peut dévier de celle-ci. Le cadre du marché est supposé continu. On note respectivement par $(X_t)_{t\geq 0}$ et $(Y_t)_{t\geq 0}$ le processus de prix des actifs A et B. La violation de la loi du prix unique est matérialisée par l'existence d'un intervalle $I = [t_0, t_1]$ avec $t_0 < t_1$ tel que $X_s \neq Y_s$, $\forall s \in I$. Vu que les actifs sont identiques, on suppose que l'hypothèse suivante est satisfaite :

$$P(Y_t \geq X_t) = P(Y_t \leq X_t). \tag{1}$$

Puisque X_t représente la valeur fondamentale de l'actif A à l'instant t, l'équation (1) nous dit que les sous ou surévaluations de l'actif B arrivent avec une même probabilité. Le choix du processus X ne joue pas un rôle important et ne sert que de référence pour modéliser la valeur fondamentale de l'actif A (marché composé que de rationnels). Cependant, le choix du processus Y joue un rôle majeur vu qu'il va modéliser la présence d'anomalies. Il sera composé d'une dynamique similaire à celle de X (effet des agents rationnels) à laquelle on ajoute un terme additionnel (effet des agents irrationnels). Pour garder un cadre général, on suppose que le processus (X_t), prix d'équilibre dans le marché des rationnels suit un modèle à volatilité stochastique de type Heston [55]. Dans ce cadre, on choisit pour Y la même dynamique que celle de X à laquelle on ajoute une partie saut. Précisément, on a :

$$dX_t = (r + \eta V_t)X_t dt + \sqrt{V_t} X_t dB_t^1 \ .$$

$$dY_t = (r + \eta V_t)Y_t dt + \sqrt{V_t} Y_t dB_t^1 + Z_t Y_{t-} dN_t \ , \tag{2}$$

où

$$dV_t = (\alpha - \beta V_t)dt + \sigma\sqrt{V_t} dB_t^2 \ .$$

Ici, (N_t) est un processus de poisson et Z_t une variable aléatoire symétrique. L'hypothèse de symétrie permet de ne pas modifier le drift de la diffusion de Y puisque $E(Z_t) = 0$ et permet de gérer la contrainte (1).

On montre alors que dans un tel marché descriptif, les arguments de la finance néo-classique et ceux de la finance comportementale sont valides par rapport à l'interprétation des anomalies. L'argument chance ou arrivée aléatoire des erreurs est obtenu grâce à l'hypothèse de symétrie de la variable Z_t. On retrouve les raisons fournies par LeRoy et Fama.
Pour la finance comportementale, l'effet des agents irrationnels est déjà pris en compte dans la partie saut de la diffusion Y. Cette dernière est responsable de la survenue des erreurs précisément de la violation de la loi du prix unique. Il reste donc à vérifier le deuxième argument de la finance comportementale à savoir que l'action des arbitragistes est limitée. On a regardé quelle est la stratégie optimale pour exploiter les inefficiences. Le résultat obtenu est que si la variable Z_t est centrée, alors la stratégie optimale consiste à ne détenir que l'actif A même si les deux actifs risqués ont des prix différents. En d'autres termes, la stratégie consistant à acheter l'actif le moins cher et à vendre l'actif le plus cher donne une utilité moindre. Par conséquent, les arbitragistes n'interviennent pas pour profiter des imperfections (limites du critère d'arbitrage). On retrouve donc les arguments des deux camps sous l'hypothèse que $E(Z_t) = 0$. Dans le cas contraire, c'est-à-dire Z_t non centrée, on montre aussi qu'on perd les arguments des deux paradigmes et le modèle descriptif ne tient plus car les arbitragistes interviennent dans le marché modifiant ainsi les prix d'équilibre supposés au départ. Une façon de réconcilier donc la finance néo-classique et comportementale est de considérer des dynamiques similaires à Y. [1]

Dans la suite, on s'intéresse qu'aux modèles discrets. L'objectif est de pouvoir introduire une description au niveau micro du comportement des agents que ne permet pas le cadre continu. Cette description devrait intégrer à la fois les arguments des deux paradigmes dans l'optique d'une réconciliation comme fait dans le cadre continu avec le processus descriptif Y. On s'est penché sur les modèles à agents qui sont plus flexibles à faire cette tâche. Ils sont basés sur les hypothèses suivantes :

1. Hétérogénéité des investisseurs.
   Dans les deux premières mouvances (finance néo-classique et comportementale), il a été supposé la présence d'agents rationnels. Ces derniers maîtrisaient la dynamique de l'économie même avec la présence d'agents irrationnels. Ce qui implique que leur connaissance intégrait celle des agents bruiteurs. Ici, cette hypothèse est relâchée. Tous les investisseurs sont hétérogènes et on retombe sur un raisonnement similaire au concours de beauté de Keynes.

1. Les modèles de diffusion avec sauts ont été introduits depuis les années 80 pour modéliser les prix d'actifs. Ici, on a fait ressortir le rôle qu'ils peuvent jouer dans la réconciliation des deux paradigmes dans le cas où la partie saut est d'espérance nulle.

2. Rationalité limitée.
   Une conséquence de l'hétérogénéité énumérée ci-dessus est que les agents ne peuvent plus avoir une rationalité illimitée. Chaque agent essaie de déterminer les paramètres de l'économie à travers un phénomène d'apprentissage ou en copiant (effet communication) les stratégies des agents voisins.

Cette approche a pris de l'intérêt ces dernières années. L'idée est au lieu de donner des caractéristiques figées aux agents, par exemple des investisseurs rationnels et irrationnels qui le resteront à vie, on crée une dynamique d'interaction entre les agents. Ces types de modèles sont résolus souvent en utilisant une approche numérique ou parfois combinent aspect numérique et analytique. Plusieurs types de modèles d'agents ont été développés de nos jours. Une première catégorie assez générale utilise des méthodes évolutionnistes. C'est le cas, par exemple, des auteurs Arthur et al. [6]. Ces derniers utilisent des algorithmes génétiques pour décrire l'évolution des croyances des individus dans le temps. Chaque investisseur a un ensemble de stratégies pour faire leurs prédictions. Une seule sera choisie à chaque instant dépendant de sa performance passée.
L'apprentissage évolutionniste est flexible au sens où après initialisation de la population des stratégies, on laisse les choses aller d'elles-mêmes. Il utilise, cependant, plusieurs paramètres en entrée de sorte qu'il est difficile après d'identifier les points importants qui ont généré les faits stylisés. Brock et Hommes [17] ont développé un concept moins complexe que le précédent et qui offre des possibilités de calculs analytiques et donc facilitant l'interprétation des résultats. Dans leur modèle, chaque investisseur dispose de H stratégies publiques partagées par tous. Contrairement à l'approche de l'algorithme génétique, les stratégies sont fixes et ne bougent pas dans le temps. Le terme en jeu est n_h, la fraction de personnes utilisant la stratégie $h \in 1, \cdots, H$. L'idée est que plus une stratégie h donne de meilleures rentabilités dans le passé plus n_h sera grand. On parle alors d'apprentissage adaptatif. Dans les deux modèles précédents, les agents étaient indépendants. Leur apprentissage se faisait dans un coin en utilisant que les données passées. D'autres auteurs ont introduit des modèles avec présence de communication modélisée par une structure de graphe. Cela permet de relier des personnes ou groupes de personnes qui vont s'échanger des informations dans leurs décisions.
Par ailleurs, ces modèles ont également investi la question de l'efficience. Par exemple, Arthur et al. [6] montrent qu'avec des croyances hétérogènes, on peut converger vers la théorie de l'efficience si les agents revoient de façon lente leurs stratégies de prévision (ici des modèles linéaires). Par contre, si la fréquence d'adaptation (mise à jour) des stratégies est élevée, la convergence vers cette croyance commune ne se produit plus et des stratégies chartistes peuvent être profitables. On retombe sur les arguments de la finance comportementale.
Chen et Yeh [20] utilisent les principes de la programmation génétique pour voir si la stratégie Achat et Conservation est optimale. Ils trouvent que si l'échantillon de données est assez faible, la stratégie acheter et ne rien faire n'est pas optimale. Par contre, si on augmente la taille de l'échantillon, cette stratégie devient prédominante signifiant que le marché est efficient.
Ces résultats, obtenus avec des algorithmes poussés, laissent présager également notre opinion à savoir que les rentabilités des titres ne suivent une martingale qu'en moyenne. Pour y apporter un éclairage, il est utile d'avoir un modèle à agents qui puisse permettre la détermination de façon explicite de la dynamique des prix. Un moyen d'y arriver est de décrire un modèle minimal en termes de complexité mais réaliste en termes d'hypothèses. Ce modèle pourra être validé en regardant s'il arrive à reproduire les principales propriétés des séries financières à savoir un faible pouvoir prédictif des prix passés, un effet de regroupement de volatilité (clustering), un excès de kurtosis, une décroissance persistante des estimations de la fonction d'auto-corrélation de la

forme $|r_t|^q$, $q > 0$ où r_t est la rentabilité de l'actif à l'instant t.
On traite de cette question dans le papier Konté [68]. La description du modèle suit Ghoulmie, Cont et Nadal [43]. On considère donc que les agents font leurs décisions dans leur coin respectif sans communication et que les données sont en hautes fréquences. L'intervalle entre deux instants est mesuré, par exemple, en secondes ou minutes. Ce qui fait que l'on peut négliger l'effet du taux sans risque et des dividendes. Les prix bougent seulement sous l'influence du flux d'informations disons $(I_t)_{t\geq 0}$ qui arrive dans le marché. On supposera que ce flux $(I_t)_{t\geq 0}$ est une séquence indépendante et identiquement distribuée (i.i.d) de loi normales de moyenne 0 et de variance σ_I^2. Chaque investisseur d'index k, évalue cette information fondamentale I_t pour en tirer un nouveau signal R_t^k qui varie d'un agent à l'autre dû à l'hétérogénéité. Précisément, ce signal perçu de I_t est de la forme

$$R_t^k = I_t + \epsilon_t^k, \quad \epsilon_t^k = \nu_t^k \cdot p_t^N. \tag{3}$$

où

$$p_t^N \;=\; \sum_{i=1}^{N} \omega_i |r_{t-i}| \;,\; r_{t-i} = \log S_{t-i} - \log S_{t-i-1}, \quad \omega_i \geq 0, \quad \sum_{i=1}^{N} \omega_i = 1. \tag{4}$$

Le terme ϵ_t^k représente le bruit idiosyncratique de l'agent k. Il est interprété comme l'erreur faite par l'agent en essayant d'évaluer l'information fondamentale I_t. Il est supposé que ν_t^k suit une loi normale $\mathcal{N}(0, \sigma_k^2)$ et que l'erreur commise est plus grande dans les périodes agitées que stables. Cela motive la pondération de ν_t^k par p_t^N qui représente une approximation de la volatilité courante basée sur la taille des rentabilités passées des actifs. Plus celle-ci est élevée plus les agents extrapolent l'information I_t. Le biais commis en évaluant I_t doit être par contre moins important que l'information I_t elle-même. C'est pourquoi, on impose que le paramètre σ_k soit très petit par rapport à σ_I.

Dans notre modèle, il n'y a pas de communication comme mentionné ci-dessus et les termes ν_t^k et ν_t^l sont supposés indépendants pour deux investisseurs différents (erreurs idiosyncratiques). Nous supposons aussi que le flux d'information (I_t) est mutuellement indépendant du bruit idiosyncratique (ϵ_t^k).

On montre que, si les décisions des agents sont proportionnelles à leur signal perçu R_t^k et si la fonction d'impact sur les prix est log-linéaire,[2] alors la dynamique des prix est donnée par :

$$\log S_t \;=\; \sum_{l=1}^{p} \log S_{t-l} \; X_t^l \quad + I_t \tag{5}$$

2. La fonction d'impact est log-linéaire si elle est de la forme $\log S_t - \log S_{t-1} = \frac{E_t}{\lambda}$ avec E_t représentant l'excès de demande et $\lambda > 0$ une constante positive de normalisation.

où $p \in \mathbb{N}^*$, (X_t^l) est tel que $E(X_t^1) = 1, \quad E(X_t^l) = 0 \quad \forall\ l \geq 2$. Nous avons aussi, $\forall l_1, l_2 \geq 1$ $cov(X_t^{l_1}, I_t) = 0$, et $cov(X_t^{l_1}, X_t^{l_2}) > 0$.

L'équation (5) montre que la dynamique des log des prix appartient à la classe des modèles auto-régressifs à coefficients aléatoires que l'on va noter RCA(p). [3] Les résultats de la simulation issus de ce modèle montrent que les principaux faits stylisés des séries financières sont reproduits. Une autre chose intéressante que l'on montre est la dynamique obtenue se résume à celle d'un processus de type ARCH(1) (voir Engle [35]) dans le cas simple où $N = 1$, $\omega_1 = 1$ correspondant au cas $p_t^1 = |r_{t-1}|$.

La description du modèle fait ressortir les principaux arguments des modèles à agents. D'abord, l'hétérogénéité est prise en compte vu que chacun agent fait une erreur idiosyncratique qui lui est propre (ϵ_t^k). La seconde est l'interaction entre agents fondamentalistes et bruiteurs. Ici, les agents fondamentalistes à l'instant t sont ceux qui ont un R_t^k de même signe que l'information fondamentale I_t et les bruiteurs l'ensemble complémentaire à savoir ceux tels que $R_t^k.I_t < 0$. Notons qu'un fondamentaliste à l'instant t peut devenir bruiteur à un instant futur et vice-versa montrant ainsi un phénomène d'interaction qui permet d'avoir à tout moment des acheteurs et des vendeurs. D'autre part, le fait de ne pas utiliser des comportements à seuils, comme dans Ghoulmie et al. [43], amène à déterminer la dynamique des prix.
Cette dynamique des prix trouvée confirme notre assertion vu que les prix sont assez similaires à ceux du processus Y en temps continu. En effet, les prix d'équilibre fournis ne peuvent être expliqués ni par la théorie de l'efficience ni par celle de la finance comportementale, à elles seules, mais plutôt par les deux combinées ensemble. En effet, les décisions des agents suivent principalement le signe de l'information I_t en faisant en même temps des extrapolations de sorte qu'il est difficile de dire si le marché est efficient ou pas. Notons que les rentabilités des actifs n'ont une racine unitaire qu'en moyenne $E(X_t^1) = 1, \quad E(X_t^l) = 0$ pour $l \geq 2$, voir (5) expliquant ainsi la complexité du débat comme dans notre illustration faite précédemment en temps continu.

Enfin, la dynamique RCA(p) n'est pas nouvelle pour modéliser la dynamique des prix. Plusieurs tests ont été proposés pour discerner les modèles auto-régressifs à coefficients constants à ceux aléatoires. On peut voir par exemple Nagakura [88] et les références qui y sont citées. La contribution est relative ici au lien établi entre certains modèles à agents avec phénomène d'interaction (ici entre deux groupes) et les modèles auto-régressifs à coefficients stochastiques. Dans la thèse de S. Alfarano [2], l'auteur établit plutôt un lien avec les modèles à volatilité stochastique même s'il existe une certaine similitude entre ces deux dynamiques.

En dernière partie, nous proposons deux applications. La première consiste à proposer des stratégies d'investissement basés sur les modèles RCA(p). L'idée est de voir si on peut générer une performance qui dépasse la stratégie Achat et Conservation. On ne s'intéresse qu'au modèle le plus simple à savoir RCA(1) donné par

$$\log S_t = \log S_{t-1}\Big(1 + b_t\Big) + I_t, \tag{6}$$

3. Dans ce document, on utilisera toujours l'acronyme anglais pour abréviation (Random Coefficient Autoregressive) de sorte à être en phase avec la littérature. Par exemple, GARCH est utilisé pour Generalized AutoRegressive Conditional Heteroskedasticity

où $E(I_t) = E(b_t) = 0$, $cov(b_t, I_t) = 0$.

Nous proposons la procédure d'estimation locale suivante qui capte l'esprit des modèles à agents (feedback). L' équation (6) peut être récrite en fonction de la rentabilité r_t définie par :

$$r_t = Y_t - Y_{t-1} = Y_{t-1}b_t + I_t, \quad \text{où } Y_t = \log(S_t). \tag{7}$$

Pour estimer b_t qui est une variable aléatoire, on fait la simplification suivante. Pour tout t, la variable b_t suit une loi binomiale prenant deux valeurs α et $-\alpha$ avec une même probabilité $p = 0,5$ de sorte à respecter l'hypothèse $E(b_t) = 0$. Dans ce cas, α^2 représente sa variance inconditionnelle. Sa variance conditionnelle est supposée stochastique comme dans les modèles ARCH/GARCH. Ce qui veut dire que l'écart type α va donc bouger dans le temps en fonction de l'état du marché et on le notera α_t. Ce paramètre peut être estimé à l'instant t, en considérant par exemple, les n dernières données pour définir la régression suivante : $K = c\mathbf{1} + L\alpha + I$ avec

$$K = (r_{t-n+1}, \cdots, r_t)', \quad et \quad L = (Y_{t-n}u_1, \cdots, Y_{t-1}u_n)' \text{ où } \forall k,\ u_k \in \{-1, 1\}.$$

Le terme $\mathbf{1}$ représente le vecteur colonne de $\mathbb{R}^n$ constitué que de 1 et u' signifie la transposée de u.

Par rapport à l'équation (7), on a ajouté une variable constante c dans la régression comme il est souvent d'usage dans la méthode des moindres carrées. Le terme I représente le vecteur colonne des résidus.

Au total, on aura 2^n régressions à faire dont chacune utilisant les n dernières données pour fournir une estimation $\hat{\alpha}_t$. Pour rendre le problème faisable, on a proposé d'abord une technique consistant à choisir n modèles parmi les 2^n dans le cas où l'entier n est tel que $n/2$, $n/12$ ou $n/20$ est une puissance de 2. Dans ce cas, on a utilisé une base de Hadamard pour faire cette réduction. Ensuite, parmi ces n modèles linéaires choisis, on en prendra le meilleur (celui qui minimise la variance des résidus) pour prédire la direction des prix futurs de la façon suivante. Par exemple, si on est en T, Y_{T+1} est prédit en utilisant la formule

$$\hat{Y}_{T+1} = \hat{c} + Y_T * (1 + \hat{\alpha_T} * \hat{u}_n), \tag{8}$$

où $(\hat{c}, \hat{\alpha_T}, [\hat{u_1}, \cdots, \hat{u_n}])$ est le terme minimisant la distance $\|I\|^2 = \|K - c\mathbf{1} - L\alpha\|^2$ avec $K = (r_{T-n+1}, \cdots, r_T)'$ et $L = (Y_{T-n}u_1, \cdots, Y_{T-1}u_n)'$.

Nous définissions, dans cette étude, une famille de stratégies paramétrée par deux entiers (m, n) qui sont tels que $m < n$ et satisfaisant de plus aux contraintes nécessaires pour l'existence d'une base de Hadamard. Pour la construction de la stratégie (m, n), on part des deux meilleurs modèles issus des régressions linéaires précédentes utilisant m et n données respectivement. Si ces deux meilleurs modèles linéaires prévoient tous une hausse (baisse) alors l'actif est vendu (acheté). En d'autres termes, on prend le contre-sens de la tendance détectée (stratégie contrarienne) car on mise sur un retour à la moyenne sur le court terme du fait que les rentabilités

financières suivent presque une marche aléatoire (voir la deuxième partie de ce document pour plus d'informations). En faisant varier les paramètres m et n tout en respectant les contraintes mentionnées ci-dessus, on peut générer plusieurs stratégies que l'on comparera à celle préconisée par la théorie de l'efficience (Achat et Conservation) consistant à ne détenir que l'actif. Pour plus d'exhaustivité dans la partie comparaison, j'ai inclus les stratégies basées sur les techniques de type moyennes mobiles. L'évaluation de performance est faite sur deux indices. Le premier étant bien entendu le CAC 40 et le second est l'indice S&P 500 qui revient souvent dans les études liées à l'efficience. Les résultats obtenus confirment que l'on peut dominer tantôt la stratégie Achat et Conservation mais le faire de façon persistante reste difficile.

La deuxième application consiste à étendre le modèle GARCH standard en utilisant le concept toujours de moyenne. En gros, on a proposé un modèle asymétrique qui se comporte comme un modèle GARCH 'en moyenne'. On y arrive en rendant les coefficients du modèle GARCH stochastiques tout en restant proche du modèle initial. Précisément, on montre contrairement aux autres modèles asymétriques qu'on a ici une même variance inconditionnelle que le GARCH standard mais avec un moment d'ordre quatre plus élevé. On a comparé ensuite ce modèle par rapport à d'autres modèles compétitifs proposés dans la littérature (EGARCH, GJR-CARCH). Nous montrons que cette nouvelle extension du modèle GARCH donne des résultats satisfaisants pour des des questions de prédiction de la volatilité du marché aussi bien pour des questions liées au management de risque.

Ce document est divisée en trois parties. L'objectif principal est de proposer une théorie qui utilise à la fois les arguments de la finance néo-classique et comportementale. C'est pourquoi, on a jugé utile de revenir sur les arguments de ces deux paradigmes dans la première partie. Dans la deuxième, on revient sur la difficulté à tester la théorie de l'efficience en se basant sur des études empiriques. Cette difficulté est ensuite reliée au fait que dans la formation des prix d'équilibre, on retrouve à la fois les arguments des deux théories. Comme conséquence, on propose un modèle qui utilise les deux types d'arguments puis la dynamique des prix au niveau macro-économique est trouvée. Sous certaines hypothèses, elle appartient à la classe RCA(p). Enfin, la troisième partie est dédiée aux applications. La première utilise le modèle économétrique RCA(1) pour créer des stratégies d'investissement. La seconde étend le modèle GARCH en intégrant une asymétrie entre les chocs positifs et négatifs sur les volatilités futures tout en modifiant la structure initiale du modèle GARCH de façon minimale.

Première partie :
Finance Néo-classique
&
Comportementale.

Cette première partie est composée de deux chapitres. Dans le premier, nous rappelons l'approche de la Finance Néo-classique. L'hypothèse de rationalité est traitée en premier donnant lieu à la théorie de l'espérance d'utilité. Les modèles obtenus sont dits additifs et se fondent principalement sur les axiomes de Von Neuman et Morgenstern [111] et plus généralement de Savage [103]. Ensuite, on revient sur le deuxième pilier de la Finance néo-classique à savoir le critère d'arbitrage qui stipule que dans un marché compétitif et sans friction, on ne peut gagner de l'argent sans prendre de risque. Enfin, on terminera en montrant le lien entre ces hypothèses et la théorie de l'efficience de marché au sens informationnel.

Dans le deuxième chapitre, on traite la partie finance comportementale. Son premier fondement est la Psychologie dont les travaux de Kahneman et Tversky [64] ont beaucoup contribué à comprendre les biais cognitifs des agents. Ces biais permettent la déviation des prix d'actifs de leurs valeurs fondamentales. Pour que ces erreurs d'évaluation puissent perdurer, il faudrait que les arbitragistes ne soient pas trop actifs. Qu'est ce qui peut expliquer un tel fait puisque ces derniers, de nature, préfèrent avoir plus que moins. C'est de là qu'apparaît le deuxième pilier pour dire que les arbitrages sont risqués et coûteux en pratique diminuant ainsi la motivation des arbitragistes à intervenir dans le marché. La combinaison de ces arguments montre que les marchés ne sont pas efficients.

Chapitre 1

Finance Néo-classique.

La théorie de la finance néo-classique a beaucoup évolué pour atteindre une phase de maturité vers les années 1980. Même si elle est aujourd'hui critiquée, elle demeure une théorie prisée car fondée sur des axiomes et produisant des règles de recommandations uniformes. Ses résultats proviennent essentiellement de deux points. Le premier est le fait que les agents disposent d'une fonction d'utilité qu'ils maximisent. Cette partie a été le fruit des théorèmes fondamentaux de Von Neuman Morgerstern [111] et de Savage [103] que nous rappelons dans la première section.

Le second point est la notion de concurrence pure et parfaite. Dans un tel cadre, on peut supposer que l'on ne peut gagner de l'argent sans prendre de risque. Si un tel fait existait, il serait immédiatement exploité par les agents. Ce point lié à la notion d'arbitrage, a révolutionné la finance moderne en permettant d'unifier dans un seul cadre le CAPM, le CCAPM, l'APT et la théorie de l'équilibre général d'Arrow et de Debreu à travers le noyau d'évaluation, voir le livre de Cochrane [23]. Ce critère est étudié dans la deuxième section.

Enfin, les deux points précédents produisent la théorie de l'efficience de marché au sens informationnel. Celle-ci stipule que les prix des actifs reflètent de façon correcte et instantanée toute l'information disponible. Nous traitons ce point dans la troisième section.

1.1 Rationalisation des préférences.

Dans notre vie quotidienne, on est souvent appelé à faire des choix qui se déroulent dans un cadre de risque ou d'incertitude modélisé par un espace probabilisé $(\Omega, \mathcal{F}, P)$. Ici, Ω représente l'ensemble des états de la nature, $\mathcal{F}$ une algèbre de sous ensembles de Ω et P une probabilité sur $\mathcal{F}$. Un acte ou une décision est une variable aléatoire de Ω à valeurs dans un ensemble $\mathcal{C}$ appelé ensemble de conséquences muni d'une algèbre $\mathcal{G}$. Si la probabilité P est connue(probabilité objective) ainsi que les conséquences $\mathcal{C}$, on dit que l'on est dans une situation de risque. C'est le cas, par exemple, du jeu de lancer de dés équilibrés. Dans le cas contraire, on est dans une situation d'incertitude (P ou $\mathcal{C}$ non connus).
Une préférence $\succeq$ est définie comme une relation binaire sur l'ensemble des décisions possibles noté $\mathcal{X}$. En d'autres termes, $\mathcal{X}$ représente l'ensemble des variables aléatoires définies de Ω à valeurs dans $\mathcal{C}$. Dans l'approche de la finance néo-classique, il est supposé que les agents suivent certaines règles (normes) dans leur décision. La première est le postulat du pré-ordre total.

Axiome 1.1.1 *pré-ordre total.*

Une structure de préférence $(\mathcal{X}, \succeq)$ est un pré-ordre total si on a les propriétés suivantes.

Pour tout $x, y, z \in \mathcal{X}$,

- $x \succeq x$ *(Réflexibilité).*
- $x \succeq y$ *et* $y \succeq z \Rightarrow x \succeq z$ *(transitivité)*
- $x \succeq y$ *ou* $y \succeq x$ *(complétude).*

Ce postulat est admis presque dans tous les modèles financiers. Dans la pratique, faire un choix dans l'ensemble $\mathcal{X}$ en utilisant $\succeq$ est difficile même dans le cas où Ω est de cardinal fini. Il est plus facile de comparer des réels que des vecteurs. C'est pourquoi, on essaie de transférer la relation de préférence $\succeq$ sur $\mathcal{X}$ en une autre notée $\geq$ sur $\mathbb{R}$. L'espace des probabilités va jouer un rôle important dans ce processus. Notons que toute variable aléatoire X sur $(\Omega, \mathcal{F}, P)$ induit une probabilité P_X appelée probabilité image sur l'espace $(\mathcal{C}, \mathcal{G})$. Cette probabilité est définie par

$$P_X(A) = P(X \in A) = P\big\{\omega \in \Omega, X(w) \in A\big\}.$$

D'autre part, deux probabilités images P_X et P_Y sont égales si et seulement si $X = Y$, $P-$ p.s. Ce qui veut dire que la relation de préférence $\succeq$ définie sur $\mathcal{X}$ induit une autre relation de préférence sur l'ensemble $\mathcal{L}$ des probabilités définie sur $\mathcal{G}$. On notera cette relation aussi $\succeq$ et le lien est donné par

$$X \succeq Y \Longleftrightarrow P_X \succeq P_Y.$$

Il est facile de voir que l'ensemble $\mathcal{L}$ est convexe et $\forall \alpha \in [0,1]$, P, $Q \in \mathcal{L}$, on peut définir la probabilité $R = \alpha P + (1-\alpha)Q \in \mathcal{L}$ par

$$R(A) = \alpha P(A) + (1-\alpha)Q(A), \quad A \in \mathcal{G}.$$

La probabilité R est interprétée comme une composition de probabilités. Elle est égale à P avec une probabilité α et égale à Q avec une probabilité $1-\alpha$.

La relation $\succeq$ sur $\mathcal{L}$ hérite des propriétés de $\succeq$ sur $\mathcal{X}$ et donc du pré-ordre total. Pour avoir une représentation numérique de $\succeq$ sur $\mathcal{L}$, on a besoin de deux nouveaux axiomes pour renchérir ce qu'on entend par décision respectant la norme à savoir la notion de continuité et d'indépendance.

Axiome 1.1.2 *Continuité.*

La relation de préférence $\succeq$ est continue au sens de Jensen [60], si pour tout $P, Q, R \in \mathcal{L}$ tel que $P \succ Q \succ R$, il existe $\alpha, \beta \in]0,1[$ tel que :

- $\alpha P + (1-\alpha)R \succ Q$
- $Q \succ \beta P + (1-\beta)R$.

Une autre façon de définir la continuité faisant intervenir la topologie de $\mathcal{L}$ est de dire que pour tout $Q \in \mathcal{L}$, les ensembles $\{P, P \succ Q\}$ et $\{P, Q \succ P\}$ sont ouverts où $\succ$ est la relation au sens strict de $\succeq$.

Axiome 1.1.3 *Indépendance.*

Pour tout $P, Q, R \in \mathcal{L}$, et $\alpha \in]0,1]$,

$$P \succeq Q \iff \alpha P + (1-\alpha)R \succeq \alpha Q + (1-\alpha)R.$$

Les deux nouveaux axiomes sont intuitifs. L'axiome de continuité dit que si on a trois probabilités reliées par une préférence stricte $P \succ Q \succ R$, on peut combiner la plus préférée et la moins préférée à travers $\alpha \in]0,1]$ de sorte que la probabilité résultante $\alpha P + (1-\alpha)R$ soit strictement préférée à la probabilité du milieu (ici Q). De même, un mélange peut se faire à travers le coefficient $\beta \in [0,1[$ de sorte que Q soit strictement préférée à la probabilité résultante $\beta P + (1-\beta)R$. La deuxième formulation de la continuité stipule que si on préfère une probabilité P à Q de façon stricte, on peut alors trouver une modification minime P' de P qui soit préférée strictement à Q.

L'axiome d'Indépendance stipule que si on préfère une loterie P à Q, notre choix reste intact si on mélange similairement ces probabilités avec une autre R. En effet, deux choix sont possibles. Soit, on est dans l'événement de probabilité $(1-\alpha)$, auquel cas on est indifférent aux deux probabilités proposées soit on est dans l'événement complémentaire auquel cas on préfère la loterie P à Q. On peut énoncer le théorème suivant établi pour Ω fini.

Théorème 1.1.1 *Von Neumann Morgenstern.*

Soit $\mathcal{L}_0$ l'ensemble des probabilités à support fini sur $\mathcal{C}$ muni de la relation de préférence $\succeq$. Les deux propositions sont équivalentes.

1. *$(\mathcal{L}_0, \succeq)$ satisfait aux axiomes de pré-ordre total, de continuité et d'indépendance.*

2. *Il existe une fonction d'utilité U représentant la relation de préférence $\succeq$ i.e $P \succeq Q \;\Leftrightarrow\; U(P) \geq U(Q)$ où*

$$U(P) = \sum_{i=1}^{n} p_i u(x_i)$$

avec u une application de $\mathcal{C} :\longrightarrow R$, croissante définie par une transformation affine croissante près.

Différentes preuves de ce théorèmes existent, voir par exemple Mas-Colell et al. [86]. Des généralisations ont été apportées en ajoutant des axiomes supplémentaires tels que la dominance où la monotonicité. Ces axiomes permettent d'étendre les résultats de Von Neumann Morgenstern dans des espaces plus généraux (support continu) tout en conservant l'existence d'une probabilité objective sur $(\Omega, \mathcal{F})$ comme hypothèse de départ. Une autre généralisation est de supposer que celle-ci n'existe pas (situation d'incertitude). Savage [103] a alors proposé six autres axiomes en plus de l'axiome de pré-ordre partiel qui permettent à la fois d'extraire une probabilité et une fonction d'utilité en même temps. Nous rappelons ci-dessous ses axiomes en introduisant quelques définitions.

On définit $X_A^Y(\omega)$ par $X_A^Y(\omega) = X(\omega)1_{A^c} + Y(\omega)1_A$. C'est donc une variable aléatoire égale à X sur le complémentaire de A et égale à Y sur l'événement A.

Une relation de préférence conditionnelle à un événement A, notée $\succeq_A$ est définie par

$$X \succeq_A Y \text{ si } \forall X',\ Y' \in \mathcal{X} \text{ telles que } X \ =_A\ X',\ \ Y \ =_A\ Y' \text{ alors } X' \succeq Y'$$

où $Z \ =_A\ Z'$ signifie $\forall \omega \in A,\ Z(\omega) = Z'(\omega)$.

Nous pouvons alors énoncer les postulats de Savage.

- (P_1) La relation de préférence $\succeq$ est un pré-ordre total.

- (P_2) $\forall X,\ Y,\ Z,\ Z' \ \in \mathcal{X}$, $\forall\ A \subset \Omega$, on a :

$$X_{A^c}^Z \succeq Y_{A^c}^Z \quad ssi \quad X_{A^c}^{Z'} \succeq Y_{A^c}^{Z'}.$$

- (P_3) $\forall\ X,\ Y,\ Z \in \mathcal{X}$, $\forall A \subset \Omega$ non nul, on a

$$Y \succeq Z \quad ssi \quad X_A^Y \succeq X_A^Z.$$

$A \subset \Omega$ est dit nul si $\forall X, Y, Z \in \mathcal{X}$, on a $X^Z_{A^c} \sim Y^Z_{A^c}$.

- (P_4) $\forall A, B \subset \Omega$, et $\forall X,\ Y,\ Z,\ W \in \mathcal{X}$ avec $X \succ Y$, $Z \succ W$,

$$Y^X_A \succeq Y^X_B \quad ssi \quad W^Z_A \succeq W^Z_B.$$

- (P_5) Il existe $X, Y \in \mathcal{X}$ tel que $X \ \succ \ Y$.

- (P_6) $\forall X, Y, Z \in \mathcal{X}$ avec $X \succ Y$, il existe une partition finie $A_1, \ \cdots \ , A_n$ de Ω telle que $\forall\ i \leq n$

$$X^Z_{A_i} \succ Y \ \ et \ \ X \succ Y^Z_{A_i}.$$

- (P_7) On considère $X,\ Y \in \mathcal{X}$ et $A \subset \Omega$. Si $\forall \omega \in A,\ X \succeq_A Y(\omega)$ alors $X \succeq_A Y$ et si $\forall \omega \in A,\ Y \succeq_A X(\omega)$ alors $Y \succeq_A X$.

Théorème 1.1.2 *(Savage) La relation de préférence $\succeq$ vérifie les axiomes de P_1 à P_7 ssi il existe une distribution de probabilité finiment additive P sur $(\Omega, \mathcal{P}(\Omega))$ et une fonction d'utilité bornée $u : \mathcal{C} \to \mathbb{R}$ unique à une transformation affine croissante près telle que*

$$\forall X, Y \in \mathcal{X}, \quad X \succeq Y \quad ssi \quad \int_\Omega u(X(\omega))dP(\omega) \geq \int_\Omega u(Y(\omega))dP(\omega).$$

Bien entendu, ce théorème est bien plus difficile à démontrer. On infère à la fois l'utilité et la probabilité P à partir des seules préférences sur les actes. Aucune topologie n'est supposée sur $(\mathcal{X}, \succeq)$ pour pouvoir utiliser des théorèmes de séparation. On peut voir les notes de cours de Gilboa[44] pour les commentaires de ces différents postulats.

Cette approche de modèle d'espérance subjective d'utilité est le fondement de la finance Néo-classique. Elle se résume en ces trois points pour un agent rationnel :

- Toute situation d'incertitude peut se ramener à une situation de risque. L'agent sera toujours capable de définir la probabilité P.

- Devant un problème de décision, l'agent maximise son espérance d'utilité sous l'unique probabilité P qui matérialise sa croyance et qui englobe toute l'information disponible. Dans le domaine de la finance, cela se traduit par le problème suivant :

$$\underset{\theta}{\text{Max}} \quad E_t^P\{U(X_T^\theta)\} \quad , \tag{1.1}$$

où X_T^θ est la valeur future d'un portefeuille ou d'une consommation pour une stratégie θ suivie. La fonction d'utilité est choisie comme strictement concave pour tenir compte de l'effet aversion au risque des agents.

– Quand un flot d'informations arrive, chaque agent fait une mise à jour correcte de sa croyance P en utilisant la loi de Bayes. La notation $E_t^P(\cdot)$ dans (1.1) est interprétée comme $E^P(\cdot|I_t)$ où I_t est l'information disponible à l'instant t.

Notons le fait qu'une fonction d'utilité soit utilisée dans la théorie de décision remonte depuis le 18eme siècle avec Daniel Bernoulli [12]. Il résolut le paradoxe de St Petersburg qui consistait à un jeu où l'on gagnait des sommes 2^n, $n \in \mathbb{N}$ avec des probabilités respectives de $\frac{1}{2^n}$, $n \in \mathbb{N}$. Bien que l'espérance de gain soit infinie, on n'est prêt qu'à miser une somme finie pour participer au jeu. Il leva ce paradoxe en disant que les individus n'utilisent pas les gains réels mais plutôt des valeurs subjectives ou utilités et c'est l'espérance de ces valeurs subjectives qui donnait le prix du jeu. L'apport de Von Neumann et Morgenstern ou de Savage est le fait qu'ils n'aient pas considéré cette fonction d'utilité comme une donnée. Plutôt, ils la construisent à partir d'axiomes.

Enfin, bien que l'hypothèse de rationalité joue un rôle très important, il est utile d'introduire la notion de concurrence dans les marchés financiers. En effet, supposer qu'il n'existe que des agents rationnels dans un marché est assez contraignant. Plusieurs études expérimentales ont montré que les agents ne respectaient pas tous les axiomes de von Neumann et Morgenstern ou de Savage. On peut donc s'attendre à des inefficiences sur la formation des prix d'équilibre. La concurrence permettra de les faire disparaître sous l'hypothèse de marché sans friction. Pour ces imperfections, la finance néo-classique en a défini plusieurs types. Elles sont relatives au fait de gagner de l'argent sans prendre de risque (opportunité d'arbitrage). Dans la section suivante, on rappelle les différentes formulations de ce critère d'arbitrage et les résultats qui y découlent.

1.2 Le critère d'arbitrage.

Dans la finance néo-classique, il est supposé une concurrence pure et parfaite. C'est une hypothèse qui assez acceptable dans le cadre des marchés financiers. L'information circule de façon libre avec le développement des nouvelles technologies. L'accès est devenu facile pour tous les agents professionnels vu que les informations s'affichent en temps réel. Pour les investisseurs individuels, un décalage peut apparaître. Ils reçoivent l'information à travers les agences de presse, les journaux, la télévision, etc. Leur effet est ici négligé vu qu'ils sont de petits porteurs. On peut supposer aussi que l'information est presque gratuite. Pour les grand porteurs, il existe des coûts liés à l'information. Souvent ils sont abonnés à des services spécialisés dans la diffusion de l'information tels que Reuter, Bloomberg, etc. Ces coûts sont cependant assez négligeables par rapport à la taille de leurs portefeuilles gérés. La seule hypothèse faite et moins réaliste est l'absence de frictions. Dans toutes les places boursières, il existe des coûts de transactions non négligeables avec des taux variant selon les marchés, selon le type d'investisseur (petit ou gros

porteur) entre autres. On conservera malgré tout cette hypothèse de marché sans friction. Dans une telle situation, il est naturel de supposer que l'on ne peut gagner de l'argent sans prendre de risque. Un premier critère que l'on va définir et rattaché à ce point est la loi du prix unique. C'est un concept économique qui dit que deux produits identiques doivent avoir le même prix. Pour le formaliser, on décrit d'abord le cadre du marché financier. Il est supposé l'existence d'un actif sans risque S^0 et d actifs risqués où $d = 1$ dans la suite pour simplifier les notations. L'incertitude est modélisée par un espace probabilisé filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. Dans le cas où le temps est discret, on note les processus de prix par $S_t = (S_t^0,\ S_t^1)$ où $t = 0, 1, \ldots, T$. Ici, T est supposé fini et $(\mathcal{F}_t)$ représente une suite croissante de tribus modélisant l'arrivée du flux d'information et le fait que les agents ne perdent pas toute information acquise. Généralement, la filtration est prise comme la tribu engendrée par les processus de prix i.e $\mathcal{F}_t = \sigma\{(S_u^0, S_u^1)\}, u \leq t)$.
L'actif sans risque S^0 donne un taux d'intérêt constant r. On introduit la nouvelle variable $\tilde{S}_t = \frac{1}{S_t^0} S_n = (1, \tilde{S}_t^1)$ avec $\tilde{S}_t^1 = \frac{S_t^1}{S_t^0}$. Cela consiste à supposer que l'actif sans risque donne un taux d'intérêt nul $r = 0$.

A partir de ces prix, les investisseurs construisent des stratégies qui sont une suite $\phi = (\phi_t^0, \phi_t^1)_{t \in \{0,1,\cdots,T-1,T\}}$ de variables aléatoires à valeurs dans $\mathbb{R}^2$ où ϕ_t^k représente le nombre de parts de l'actif k dans le portefeuille à la date t. Les stratégies sont supposées prévisibles c'est-à-dire que ϕ_t est $\mathcal{F}_{t-1}$-mesurable ou dit d'une autre manière que les agents vont faire leurs décisions à l'instant t en utilisant l'ensemble d'information $\mathcal{F}_{t-1}$. Par ailleurs, il est supposé que durant la période $[0, T]$, l'investisseur ne retire ni n'injecte de l'argent dans son portefeuille i.e $\phi_t \cdot \tilde{S}_t = \phi_{t+1} \cdot \tilde{S}_t$ pour $0 \leq t \leq T - 1$. On dit alors alors que la stratégie est autofinancée.
Dans ce cas

$$V_t(\phi) - V_{t-1}(\phi) = \phi_t \cdot \tilde{S}_t \ - \ \phi_{t-1} \cdot \tilde{S}_{t-1} = \phi_t \cdot (\tilde{S}_t - \tilde{S}_{t-1}) = \phi_t^1 \cdot (\tilde{S}_t^1 - \tilde{S}_{t-1}^1).$$

On note que la variation du portefeuille ne dépend que de celle de l'actif risqué.[4] Ce qui veut dire que l'on peut remplacer la notation $\phi = (\phi^0, \phi^1)$ par ϕ^1 et on a

$$V_t(\phi^1) - V_{t-1}(\phi^1) = \phi_t^1 \cdot (\tilde{S}_t^1 - \tilde{S}_{t-1}^1).$$

Par sommation, l'équation ci-dessus est équivalente à

$$V_T(\phi^1) = V_0(\phi^1) + \sum_{t=1}^{T} \phi_t^1.(\tilde{S}_t^1 - \tilde{S}_{t-1}^1) = V_0(\phi^1) + \sum_{t=1}^{T} \phi_t^1.\Delta_t \tilde{S}^1. \tag{1.2}$$

On définit une stratégie ϕ^1 comme admissible si elle est autofinancée et si elle vérifie pour tout t, $V_t(\phi^1) \geq c$ où c est un réel. Cette dernière condition est toujours satisfaite dans notre cas où les transactions se font en des instants discrets avec $T < \infty$. Cependant, dans le cas continu ou pour un horizon infini, cette hypothèse peut être violée.

4. Cela vient du fait que $\Delta_t \tilde{S}^0 = \tilde{S}_t^0 - \tilde{S}_{t-1}^0 = 1 - 1 = 0$, pour tout t.

Enfin, on définit l'ensemble K_0 des variables aléatoires $\mathcal{F}_T$-mesurables qui sont atteignables par des stratégies admissibles de richesse initiale nulle c'est-à-dire

$$K_0 \;=\; \{X \in L^0(\Omega, \mathcal{F}_T),\ tel\ que\ \ X = \sum_{t=1}^{T} \phi_t^1.\Delta_t \tilde{S},\ \ \phi^1\ \ admissible.\} \tag{1.3}$$

$$K_0 \;=\; \left\{ \sum_{t=1}^{T} \phi_t^1.(\tilde{S}_t - \tilde{S}_{t-1}),\ \ \phi_t^1 \in L^0(\Omega, \mathcal{F}_{t-1}). \right\}$$

où $L^0(\Omega, \mathcal{F}_{t-1})$ signifie l'ensemble des variables aléatoires $\mathcal{F}_{t-1}$ - mesurable.

La loi du prix unique est formalisée par la définition suivante, voir Courtault et al. [26]

Définition 1.2.1 *(La loi du prix unique) La notion de la loi du prix unique est définie à $t = 0$, par la relation*

$$K_0 \cap L^0(\Omega, \mathcal{F}_0) = \{0\}. \tag{1.4}$$

L'équation ci-dessus peut être récrite comme suit

$$V_0(\phi^1) + \sum_{t=1}^{T} \phi_{t-1}^1.(\tilde{S}_t - \tilde{S}_{t-1}) = 0 \implies V_0(\phi^1) = 0$$

ou encore

$$V_0(\phi^1) + \sum_{t=1}^{T} \phi_{t-1}^1.(\tilde{S}_t - \tilde{S}_{t-1}) = V_0(\phi^2) + \sum_{t=1}^{T} \phi_{t-1}^2.(\tilde{S}_t - \tilde{S}_{t-1}) \implies V_0(\phi^1) = V_0(\phi^2).$$

Cette dernière équation traduit le fait que si deux variables ont même pay-off alors elles doivent avoir un prix identique même si les stratégies permettant de les répliquer sont différentes.

Avant de donner le théorème lié à la loi unique, on rappelle la définition d'une martingale.

Définition 1.2.2 *Martingale.*
On dit qu'un processus (S_t) est une P-martingale par rapport à une filtration $(\mathcal{F}_t)$ si on a :

1. *Pour tout t, S_t est P- intégrable.*
2. *Pour tout u, t tel que $u < t$, $\quad E_P(S_t|F_u) = S_u$.*

On a le théorème suivant, voir Courtault et al. [26] :

Théorème 1.2.1 *Les conditions suivantes sont équivalentes,*
a) $K_0 \cap L^0(\Omega, \mathcal{F}_0) = \{0\}$.

b) Il existe une martingale bornée (Z_t) *satisfaisant* $E_P(Z_T) = 1$, $Z_0 > 0$ *et tel que le processus* $(Z_t S_t)$ *soit aussi une* P*-martingale par rapport à* $(\mathcal{F}_t)$.

Notons que le processus (Z_t) n'est positif de façon certaine qu'à l'instant $t = 0$. Il peut prendre des valeurs négatives pour les instants futurs donnant lieu à des mesures signées.
Gaussel [42] a défini la notion de marché faiblement arbitré qui est similaire à celui de la loi unique et qui donne aussi l'existence d'une mesure signée. Son formalisme est plutôt proche de celle de Harrison et Kreps [54].
L'autre propriété que les auteurs Courtault et al.[26] montrent est l'existence d'un portefeuille de marché si on se restreint aux variables de carré intégrables dans K_0.

Bien que ce concept de loi du prix unique soit très utilisé dans le domaine de l'économie, il faudra des notions plus fortes pour décrire un marché compétitif. En effet, comme on vient de le voir ci-dessus la mesure martingale obtenue peut être signée. Ce qui signifie que l'hypothèse consistant à dire que les agents préfèrent obtenir plus que moins et admise presque dans tous les modèles financiers peut donc être violée.

Pour garantir l'existence et la positivité du processus Z, Kreps [70] remplace la condition de la loi du prix unique par celle d'absence d'arbitrage définie de la façon suivante :

$$K_0 \cap L^0_+(\Omega, \mathcal{F}_T) = \{0\}. \tag{1.5}$$

L'équation (1.5) peut être traduite par il n'existe pas :

- ϕ admissible.
- $V_0(\phi) = 0$, $V_T(\phi) \geq 0$ et $P(V_T(\phi) > 0) > 0$.

Dalang et al. [28] étendent le résultat de Kreps [70] en démontrant pour la première fois l'équivalence entre absence d'opportunité d'arbitrage et existence d'une mesure martingale dans le cadre discret, Ω infini et T fini. Cependant, pour un cadre continu ou pour un horizon de temps infini, il faudra des notions plus fortes pour conserver ce théorème fondamental.
Notamment pour le cadre continu, Kreps [70] introduit la notion de No Free Lunch (NFL). Ce concept utilise des approximations d'opportunités d'arbitrage à travers une certaine topologie. D'abord, l'ensemble K_0 défini dans l'équation (1.3) est remplacé par K qui représente son équivalent en temps continu.

$$K = \left\{ \int_0^\infty H_s dX_s, \quad H \quad admissible \right\} \tag{1.6}$$

où le processus (X_t) est une semi-martingale. [5]

5. X est de la forme $X_t = M_t + A_t$ où (M_t) est une martingale locale et (A_t) un processus à variation bornée.

Une stratégie H est dite admissible si

- pour t, $\int_0^t H_s dX_s \ \geq c$, $c \in \mathbb{R}$
- H est prévisible et X - intégrable.

Le deuxième point important que Kreps a introduit est les cônes de la forme $C_1 = (K - L^p_+) \cap L^p$ ou $C_2 = (K - L^0_+) \cap L^\infty$. On dira qu'il y a No Free Lunch (NFL) si la condition suivante est satisfaite

$$\bar{C_1} \cap L^p_+ = \{0\}. \tag{1.7}$$

où $\bar{C}$ est la fermeture de C par rapport à la norme $\| \ \|_p$ pour $1 \leq p < \infty$ et par rapport à la topologie faible étoile pour $p = \infty$.

Mesurer une approximation par rapport à la topologie faible n'a pas trop d'intérêt économique. Cela a motivé Dalbaen et Schachermayer [32] à utiliser la norme max. Ils définissent un nouveau critère qu'ils appellent No Free Lunch With Vanishing Risk(NFLVR) par

$$\bar{C_2} \cap L^\infty_+ = \{0\}. \tag{1.8}$$

où $\bar{C}$ est la fermeture de C par rapport à la norme maximum $\| \ \|_\infty$.

Le théorème qui montre l'équivalence entre absence d'opportunité d'arbitrage et existence au moins d'une mesure martingale positive (probabilité risque neutre) est connu sous le label de premier théorème fondamental de la finance. On peut se référer à l'ouvrage de Dalbaen et Schachermayer [33] pour plus de détails sur l'évolution de ce théorème allant de Ω de cardinal fini au cas infini et pour un horizon T fini comme infini aussi.

L'avantage de cette approche d'évaluation par arbitrage est que l'on n'a pas besoin de spécifier un type de fonction d'utilité (logarithme, exponentiel, quadratique, etc.) pour les agents afin de déterminer les prix d'équilibre des autres actifs appartenant à K ou plus généralement à $x + K$ où $x \in \mathbb{R}$. Un autre point important est que l'on n'a pas besoin de connaître la richesse ou la disposition de biens initiale des agents. En gros, le théorème établit les équivalences suivantes :

- Absence d'opportunités d'arbitrage (NFLVR).
- Existence d'un opérateur de prix positif (probabilité risque neutre).
- Existence d'un optimum pour un agent préférant plus que moins (viabilité).

Remarquons que ce principe d'arbitrage a permis d'établir la formule de Black et Scholes [13]. D'abord les hypothèses émises sont telles à obtenir un marché complet. Ce cadre de complétude stipule que le pay-off de tout actif peut être obtenu par une stratégie dynamique comme celle de l'équation (1.6) translatée à une constante près c. Cette constante représentera le prix de l'actif vu que les éléments de K ont un prix nul sous la probabilité risque neutre. La notion de complétude est fondamentale dans cette approche puisqu'il permet de déterminer de façon unique le prix des actifs. Dans le cas d'un marché incomplet, le critère d'arbitrage favorise toujours la détermination des prix d'actifs redondants de façon unique. Cependant, pour ceux qui ne le sont pas, seul un intervalle de prix viables est proposé.
Une autre application du critère d'arbitrage est la théorie d'évaluation par arbitrage de Ross [97], [98] (Asset Pricing Theory). Les actifs de base sont supposés être des facteurs qui affectent les rentabilités de tous les titres. On n'a pas une redondance au sens strict du terme. Pour autant, on peut à partir des rentabilités de ces facteurs de base en déduire celle des autres actifs. L'idée consiste à faire une régression des rentabilités des tires sur les facteurs de base. Le nombre de ces facteurs doit être de sorte que les résidus issus de la régression des différents titres ne soient pas corrélés (bruit idiosyncratique). De là, on détermine la prime de risque de tout titre en utilisant le critère d'arbitrage consistant à dire qu'un portefeuille de prix initial nul et ne représentant aucun risque (variance nulle) ne peut procurer un rendement espéré positif. Notons que le critère d'arbitrage s'applique ici sur les moments des actifs et non sur les pay-offs des actifs. Formellement le choix de ces facteurs se fait par des méthodes d'analyse en composante principale (ACP) et on perd du coup l'aspect interprétation économique des facteurs macro-économiques. On pourra adopter aussi l'approche des facteurs explicites comme dans le CAPM où par contre on perd en termes de pouvoir explicatif par rapport à l'APT.

On termine cette section en notant que dans l'évaluation par le critère d'arbitrage le travail qui est fait est le suivant (voir Harrison et Kreps [54]). On part d'un marché initial décrit comme suit

- $(\Omega, \mathcal{F}, P)$ pour l'incertitude du marché.

- T pour l'horizon de temps des agents.

- $X = L^p(\Omega, \mathcal{F}, P)$ pour l'ensemble des plans de consommations des agents ou l'ensemble des actifs financiers disponibles.

- (M, π) pour les actifs de base qui sont évalués à travers l'opérateur de prix π. Il est supposé que M est un sous espace vectoriel de X.

- K pour l'ensemble des actifs dans le marché augmentant le bien être des agents. Il est supposé être un cône pour matérialiser la propriété de non satiété des agents. Précisément $k \in K$ si $\forall x \in X, \forall \lambda > 0 \;\; x + \lambda k \succ x$.

Ensuite, l'opérateur de prix π est étendu du sous-espace vectoriel M à l'espace entier X sous certaines conditions de viabilité. La viabilité stipule que tout agent arrive à trouver un portefeuille maximisant ses préférences tenant compte de ses contraintes budgétaires. Par exemple, dans le cadre du modèle de Black et Scholes, l'opérateur π donne des prix pour le sous-jacent de la forme

$$dS_t \;=\; \mu S_t dt \;+\; \sigma S_t dW_t.$$

De là, on utilise l'opérateur étendu ϕ pour évaluer des actifs généraux de la forme $f(S_t)$. Bien entendu, pour le bon fonctionnement de l'économie, le marché de base (M, π) ne doit pas être seulement viable mais doit refléter aussi les fondamentaux. Sinon, l'opérateur ϕ prolongeant π ne pourra plus redirectionner les fonds vers les firmes les plus productives. C'est pourquoi une notion plus forte que les individus préfèrent plus que moins est nécessaire dans ce cas. Précisément, il est supposé que les investisseurs savent aussi traiter toute l'information disponible de façon correcte. C'est la théorie de l'efficience de marché au sens informationnel qui combine à la fois rationalité des agents et absence de frictions.

1.3 Efficience de marché.

La notion d'efficience informationnelle a été introduite par Fama [36]. Plusieurs types de définitions existent actuellement. Pour celle de Fama, elle stipule que les cours des actifs reflètent pleinement et de façon correcte toute l'information disponible. Par conséquent, on ne peut avoir un excès de rentabilité en utilisant l'information disponible. Le terme pleinement est lié à la structure du marché qui doit être parfaite. Comme on l'a mentionné précédemment, l'information doit être gratuite et circuler de façon libre. Cela permet aux agents de la disposer dès son apparition et sans délai. Le terme correct réfère à la rationalité des investisseurs. Ils doivent être capables, après avoir reçu l'information, d'en faire un bon traitement. Dans le critère d'arbitrage, on ne s'est pas soucié de la formation des prix d'actifs de base fournis par l'opérateur de prix π. Pour que la théorie soit complète, il faudra préciser comment ces prix de base se forment aussi. Plutôt que d'entrer dans la théorie de l'équilibre général, on va plutôt énoncer un concept qui stipule que plus les agents mettent des moyens pour obtenir de bonnes prévisions sur les rentabilités futures plus la variation des prix de ces actifs de base devient aléatoire. Ce concept est l'aboutissement de plusieurs années de recherche dont l'un des précurseurs a été Louis Bachelier [7]. Il a introduit et utilisé le mouvement brownien pour modéliser la dynamique des cours boursiers. Comme conséquence, la fluctuation des cours boursiers suit une marche aléatoire qui est définie comme suit :

Définition 1.3.1 *On dit qu'un processus (S_t) suit une marche aléatoire si*

$$S_{t+1} = S_t + \epsilon_{t+1}$$

où les variables aléatoires $\epsilon_t,\ t \in \mathbb{N}$ sont indépendantes, centrées et identiquement distribuées.

En termes statistiques, la notion de marche aléatoire implique que toute fonction dépendante de la variation des prix doit être indépendante de son passé. Or dans les séries financières, il a été remarqué que les variances des rentabilités étaient corrélées. Une forte volatilité a tendance à entraîner une forte volatilité et de même les petites volatilités ont tendance à se succéder. C'est pourquoi vers les années 1960, la notion de marche aléatoire a été remplacée par celle de

martingale voir définition 1.2.2 qui est un concept moins contraignant et dont l'un des initiateurs a été Samuelson [101]. Une autre définition liée à la notion de martingale est celle de jeu équitable.

Définition 1.3.2 *Jeu équitable.*
On dit qu'un processus (S_t) est un jeu équitable si

1. *Pour tout t, S_t est P- intégrable.*
2. *Pour tout u, t tel que $u < t$,* $\quad E_P(S_t|F_u) = 0$.

Dans un cadre discret, on peut remarquer qu'un processus (S_t) est une martingale si et seulement si sa différence $(S_t - S_{t-1})$ est un jeu équitable. Ce concept de jeu équitable sera utilisée par Fama [36] pour formaliser la définition de l'efficience.

Définition 1.3.3 *(Efficience de marché) Un marché est efficient par rapport à un ensemble d'information $\mathcal{F}_t$ si pour tout actif j, $\epsilon^j_{t+1} = r^j_{t+1} - E(r^j_{t+1}|\mathcal{F}_t)$ entraîne que $E(\epsilon^j_{t+1}|\mathcal{F}_t) = 0$ où r^j_{t+1} est la rentabilité du titre j à l'instant $t+1$.*

La définition stipule que si tous les agents forment des prévisions rationnelles par rapport aux rentabilités futures alors l'excès de rentabilité de tout actif j $\epsilon^j_{t+1} = r^j_{t+1} - E(r^j_{t+1}|\mathcal{F}_t)$ a une espérance nulle. De plus (ϵ^j_t) est un jeu équitable.[6] Les paramètres r^j_{t+1}, ϵ^j_{t+1} sont aléatoires à l'instant t et celui de $E(r^j_{t+1}|\mathcal{F}_t)$ connu et déterminé à partir d'un modèle de prévision.

Pour tester l'hypothèse d'efficience de marché, on doit préciser deux choses. La première est le contenu de $\mathcal{F}_t$. Trois ensembles d'information ont été distingués :

1. $\mathcal{F}_t$ contient les prix passés des actifs, l'historique des dividendes, des taux d'intérêt, des volumes de transactions, des prix d'options entre autres. Dans ce cas, on parle de test d'efficience au sens faible. Si le marché est efficient par rapport à cet ensemble d'information alors les techniques graphiques (chartistes) ne peuvent permettre d'avoir un excès de gain persistant. On peut réaliser un gain mais simplement par chance.

2. $\mathcal{F}_t$ contient en plus de l'historique citée ci-dessus, des données spécifiques aux entreprises. On parle alors de tests d'efficience au sens semi-fort. A ce niveau, l'efficience signifie que les stratégies fondamentalistes ne permettent pas elles aussi de battre le marché. Souvent, les tests utilisent la méthodologie de l'étude d'événements. Cette technique permet de mesurer la vitesse d'ajustement des prix face à un événement donné. Plusieurs types d'événements ont été étudiés dans la littérature tels que les divisions d'actions, les distributions de dividendes, les offres publiques d'achats entre autres.

6. Il a été noté que cette formulation est toujours vraie de par la propriété de l'espérance conditionnelle.

3. Enfin, si $\mathcal{F}_t$ contient à la fois l'information publique et privée, l'efficience de marché est alors testée au sens fort. L'efficience à ce niveau signifie qu'un agent informé ne peut tirer profit de son avantage par rapport à un agent non informé. L'idée sous-jacente est qu'un initié en faisant ses transactions communique en même temps son information au marché et perd de façon automatique son avantage. Par ailleurs, si son information privée a été déjà anticipée par les analystes financiers, il ne pourra pas en tirer profit par rapport au reste des agents. On compare souvent la performance des gérants des fonds supposés avoir plus d'information à celle du portefeuille de marché.

Le deuxième point important est le choix d'un modèle d'équilibre pour prédire la valeur de $E(r^j_{t+1}|\mathcal{F}_t)$. De façon générale, cette rentabilité est évaluée en utilisant un portefeuille de référence qui est choisi à travers une maximisation d'une fonction d'utilité. Dans les années 1960, la fonction d'utilité quadratique a été utilisée donnant lieu au modèle du CAPM avec le portefeuille de marché comme portefeuille de référence. Dans les années 1990, la fonction d'utilité logarithmique a pris de l'intérêt et le portefeuille de référence appelé numéraire a la caractéristique d'être optimale en termes de croissance sur le long terme. Dans les deux cas, on aboutit à un modèle avec séparation de fonds. Pour plus de détails, on peut se référer à Bajeux-Besnainou et Portait [8]. L'approche moderne part du premier théorème fondamental de la finance qui stipule que l'absence d'opportunité d'arbitrage est équivalente à l'existence d'un noyau d'évaluation ξ_{t+1} tel que

$$S^j_t \;=\; E(S^j_{t+1}\xi_{t+1}|\mathcal{F}_t), \qquad \xi_{t+1} \;>\; 0. \tag{1.9}$$

Souvent dans les études empiriques, on préfère travailler sur les rentabilités des titres. Elles ont la caractéristique de ne pas avoir de tendance et d'être proches d'un bruit blanc. Si on note la rentabilité du titre j, $r_{j(t+1)} = \frac{S^j_{t+1}-S^j_t}{S^j_t}$ [7], alors l'équation précédente devient :

$$E_t((1+r^j_{t+1})\xi_{t+1}) \;\stackrel{\triangle}{=}\; E_P((1+r^j_{t+1})\xi_{t+1}|\mathcal{F}_t) \;=\; 1. \tag{1.10}$$

En notant r_t la rentabilité déterministe de l'actif sans risque [8] sur $[t, t+1]$, on a $E_t(\xi_{t+1}) = \frac{1}{1+r_t}$. La relation (1.10) peut alors être récrite comme suit

$$E_t(r_{j(t+1)}) - r_t = -cov\left(r_{j(t+1)}, \xi_{t+1}(1+r_t)\middle|\mathcal{F}_t\right) = -cov(r_{j(t+1)}, L_{t+1}|\mathcal{F}_t),$$

7. Il est supposé que les dividendes sont inclus dans le pay-off de S^j_{t+1}.
8. L'hypothèse de la présence d'un actif sans risque peut être levée. Il suffit dans ce cas de considérer un actif non corrélé au noyau d'évaluation ξ (actif ou portefeuille zéro béta).

où on a posé $L_{t+1} = \xi_{t+1}(1 + r_t)$.

Si on considère la variable aléatoire L_{t+1} comme un actif réplicable et si on note $R_{L_{t+1}}$ la rentabilité du portefeuille sur $[t, t+1]$, alors l'équation précédente devient

$$E_t(r_{j(t+1)} - r_t) = -cov(r_{j(t+1)}, L_t(1 + R_{L_{t+1}})|\mathcal{F}_t) = -L_t cov(r_{j(t+1)}, R_{L_{t+1}}|\mathcal{F}_t). \tag{1.11}$$

En particulier si l'actif j est égal à l'actif de référence L, on a

$$E_t(R_{L_{t+1}} - r_t) = -L_t V(R_{L_{t+1}}|\mathcal{F}_t) \quad i.e \quad -L_t = \frac{E_t(R_{L_{t+1}} - r_t)}{V(R_{L_{t+1}}|\mathcal{F}_t)}.$$

Remplaçant $-L_t$ par sa valeur obtenue dans l'équation (1.11), on obtient que

$$E_t(r_{j(t+1)} - r_t) = \frac{cov(r_{j(t+1)}, R_{L_{t+1}}|\mathcal{F}_t)}{V(R_{L_{t+1}}|\mathcal{F}_t)}\Big(E_t(R_{L_{t+1}}) - r_t)\Big) = \beta_t^{j,L} E_t(R_{L_{t+1}} - r_t). \tag{1.12}$$

L'équation (1.12) nous dit que l'excès de rentabilité conditionnelle de tout actif actif j par rapport au taux sans risque (prime de risque) est proportionnelle à celui du portefeuille de référence. Le facteur de proportionnalité est mesuré par $\beta_t^{j,L}$.

Des tests empiriques ont montré des violations liées à la prédictibilité des rentabilités (1.12). Cela a motivé le développement d'un nouveau courant vers la décennie nonante appelé Finance Comportementale que l'on traite dans le chapitre suivant. Nous précisons que du point de vue des partisans de la finance néo-classique, ces faits ne sont pas automatiquement liés à des anomalies car des explications rationnelles pouvant être données. En effet, dans le test de l'équation (1.12), aucune contrainte n'est faite sur le noyau d'évaluation ξ_{t+1} dans la théorie de l'arbitrage qui a été rappelée dans la section précédente. Cependant pour des raisons empiriques, une forme doit être spécifiée pour ξ basée sur la théorie de maximisation d'utilité des agents. Cela veut dire que si l'équation (1.12) est rejetée, le marché est inefficient ou simplement que le noyau d'évaluation est mal spécifié par un mauvais choix de la fonction d'utilité par exemple. Le chapitre 3 de cette première partie revient sur la complexité de cette question d'efficience de marché aussi bien sur le plan empirique que théorique.

Chapitre 2

Finance Comportementale.

La finance comportementale est un courant bien implémenté, aujourd'hui, dans la littérature financière. Son approche consiste à relâcher l'hypothèse de rationalité des agents basée sur les axiomes de Savage ou de Von Neumann Morgenstern. On tombe alors sur la théorie des modèles non additifs qui est traitée en deux sections. Dans la première, on rappelle les modèles dits quasi additifs au sens où ils se fondent sur des axiomes moins contraignants tendant par contre à normaliser toujours les décisions. Dans la deuxième, on revient sur la Psychologie qui est beaucoup plus utilisée dans la finance comportementale et qui n'a aucune intention de rattacher le comportement des individus à de la rationalité. On accordera donc plus d'importance à cette partie. D'abord, on s'intéresse à sa partie expérimentale qui montre principalement que les agents utilisent des règles de décisions simples devant des problèmes complexes. Ces règles simplificatrices, souvent communes aux personnes, sont appelées heuristiques. Comme conséquence de leur simplicité, elles génèrent des biais par rapport aux décisions rationnelles où les individus savent résoudre tout problème complexe. En deuxième lieu, on rappelle la théorie du prospect développée par Kahneman et Tversky [64] qui essaie de formaliser ces biais au niveau des préférences des agents. C'est une théorie qui a joué un rôle fondamental dans le développement de la finance comportementale. Barberis et al. [11], Daniel et al. [30] montrent qu'on peut utiliser certaines heuristiques pour expliquer les déviations des prix d'actifs de leur valeur fondamentale. Si de telles imperfections arrivent dans le marché, pour la finance néo-classique, les arbitragistes entrent en jeu pour profiter de ces erreurs rendant ainsi le marché efficient. Pour que l'argument de la finance comportementale tienne, il faut donc la présence de facteurs qui puissent limiter l'action de ces arbitragistes. La dernière section est consacrée à ce fait connu sous le label de *Limits to arbitrage*. Dans cette section, on rappelle comme dans le cas de la psychologie, un modèle important en guise d'illustration. C'est celui développé par De Long et al. [81]. Les auteurs ont proposé ce qu'on appelle *noise trader risk* qui est lié au risque que le comportement des agents irrationnels s'accentue dans le court terme. On énumérera, ensuite, quelques autres facteurs proposés dans la littérature pouvant expliquer l'effet limité des arbitragistes. On terminera, enfin, par mentionner quelques études empiriques qui ont investi cette question d'arbitrage limité. Il a été noté, par exemple, que plus il y a de l'incertitude face à la valeur fondamentale d'un actif, plus il est difficile de prévoir les distributions futures d'un actif et moins les arbitragistes sont enclins à intervenir pour profiter des imperfections.

2.1 Modèles quasi-normatifs.

Dans la première partie, on a rappelé les axiomes de Von Neumann et Morgenstern ou plus généralement de Savage qui permettent d'avoir une représentation numérique des relations de préférences. Malheureusement, certains de ces axiomes sont contraignants et ne sont pas respectés dans le choix des individus. La littérature montrant ces évidences est assez vaste que ce soit dans le cadre du risque ou de l'incertitude. On se contente ici d'un seul exemple (paradoxe) en guise d'illustration dans un cadre de risque. La probabilité P est donc connue ainsi que les conséquences $\mathcal{C}$. Dans ce cas, une loterie est matérialisée par un vecteur L dans $\mathbb{R}^{2n}$ où $L = (x_1, p_1; \cdots; x_n, p_n)$. Cette notation stipule qu'il y a n états de la nature et la loterie donne dans l'état $1 \leq i \leq n$ une valeur (conséquence) x_i avec une probabilité p_i. L'illustration est l'effet de la conséquence commune. Ce type de paradoxe a été pour la première fois illustré par Allais [4]. Dans cette expérience, on demande aux individus de choisir entre les couples de loteries suivantes (1M=1 million d'euros par exemple.)

$$L_1 = (1\ M;\ 1.0) \qquad \textit{contre} \qquad L_2 = (5M, 0.10;\ 1M, 0.89;\ 0, 0.01)$$

$$L_3 = (5M, 0.10;\ 0, 0.90) \qquad \textit{contre} \qquad L_4 = (1M, 0.11;\ 0, 0.89)$$

Les participants, dans la grande majorité, choisissent la loterie L_1 contre L_2 et la loterie L_3 contre L_4. Ces choix sont en contradiction avec l'axiome d'indépendance. En effet le choix de la loterie L_1 sur L_2 entraîne que $u(1M) > 0.10u(5M) + 0.89u(1M) + 0.01u(0)$ alors que le deuxième choix des investisseurs entraîne que $0.10u(5M) + 0.90u(0) > 0.11u(1M) + 0.89u(0)$. On peut récrire ces inégalités de la façon suivante :

$$0.10u(1M) + 0.89u(1M) + 0.01u(1M) > 0.10u(5M) + 0.89u(1M) + 0.01u(0).$$
$$0.10u(5M) + 0.89u(0) + 0.01u(0) > 0.10u(1M) + 0.89u(0) + 0.01u(1M).$$

On remarque que les loteries L_1 et L_4 donne les mêmes utilités (0.10 u(1 M)+ 0.01 u(1M)) si l'événement de probabilité 0.89 ne se réalise pas. Il en est de même pour les loteries L_2 et L_3 (0.10 u(5M) + 0.01 u(0)). D'autre part, sur l'événement complémentaire les loteries L_3 et L_4 donnent la même utilité (0.89 u(0)). L'axiome d'indépendance (voir **Axiome** 1.1.3) entraîne alors que si L_1 est préférée à L_2 alors L_4 doit être préférée à L_3. Choisir L_3 par rapport à L_4 après avoir choisi L_1 sur L_2 est une violation donc de l'axiome d'indépendance.

Ce paradoxe parmi tant d'autres ont motivé beaucoup de chercheurs à revoir l'axiome d'indépendance tout en conservant les axiomes de pré-ordre total et de continuité. Cela permet d'obtenir une fonction V représentant la relation de préférence $\succeq$ avec des propriétés moins contraignantes. Dans la théorie de l'espérance d'utilité, les courbes d'indifférence des agents sont linéaires et parallèles. Le seul paramètre sur lequel on peut jouer est la pente des courbes d'indifférence (aversion au risque des agents). Considérons des loteries de la forme $L = (x_1, p_1; x_2, 1 - p_1 - p_3; x_3, p_3)$ où $x_3 \succ x_2 \succ x_1$. Celles-ci peuvent se représenter dans un plan (triangle de Marschak-Machina) où

n'apparaît que x_1 et x_3 avec leurs probabilités respectives p_1 et p_3 variant sur [0,1]. Par convention, l'axe vertical représente la meilleure conséquence x_3 tandis que la mauvaise x_1 se retrouve au niveau de l'axe horizontal.

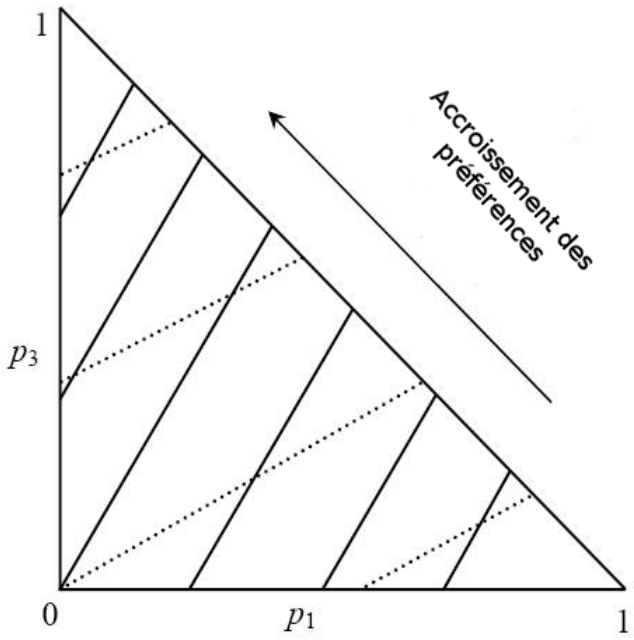

FIGURE 2.1 – Triangle de Marschak-Mashina pour le cas des modèles additifs.

Les courbes d'indifférence dans la figure 2.1 sont définies par la relation

$$\left\{(p_1, p_3) \in [0,1]^2, \quad V(L) = u(x_1)p_1 \ + \ u(x_2)(1 - p_1 - p_3) \ + \ u(x_3)p_3 = \ c \ = constante\right\}.$$

En augmentant la valeur c, on fait une translation de la courbe d'indifférence de la droite vers la gauche dans la direction de la flèche. Plus la pente des courbes d'indifférence est élevée (ligne pleine dans le schéma), plus la personne a un risque d'aversion élevé. En effet, une petite augmentation p_1 de la probabilité d'apparition de la mauvaise conséquence est compensée par un accroissement assez important de la probabilité d'apparition de la meilleure conséquence p_3.

Une façon de comprendre le choix des individus est, dans un premier temps, de permettre la linéarité des courbes d'indifférence tout en relâchant l'hypothèse de parallélisme. En d'autres termes, l'aversion au risque des agents augmente selon l'accroissement de leur richesse. Pour conserver d'abord la linéarité, on utilise au moins l'axiome suivant.

Axiome 2.1.1 *Betweeness.*

Soient deux loteries $Q,\ R$ telles que $Q \succ R$. Alors pour tout $0 < p < 1$, on a

$$Q \ \succ \ (Q,\ p;\ R,\ (1-p)) \ \succ R$$

L'axiome stipule que si une probabilité Q est strictement préférée à R alors tout mélange de ces deux loteries sera en termes de préférences compris entre Q et R. Bien entendu, si on est indifférent entre les loteries Q et R, alors on le sera de même pour toutes les loteries sur le segment $[R, Q]$ grâce à l'axiome de continuité. Ce fait explique ainsi la linéarité des courbes d'indifférences. Dans ce qui suit, on relâche la contrainte de parallélisme en affaiblissant l'axiome d'indépendance.

Axiome 2.1.2 *Indépendance au sens faible.*
Soient trois loteries $Q,\ R,\ S$ telles que $Q \succ R$. Alors pour tout p_q, il existe p_r tel qu'on ait

$$(Q,\ p_q;\ S,\ 1-p_q) \ \succ (R,\ p_r;\ S,\ 1-p_r)\ .$$

Avec cet axiome, on arrive à expliquer, par exemple, le paradoxe d'Allais présenté précédemment.
Une autre alternative est aussi de relâcher la propriété de linéarité des courbes d'indifférence (axiome Betweeness). Dans ce cas, les préférences des agents sont convexes ou concaves. Un agent peut être indifférent entre deux loteries L_1 et L_2 et préférer de façon stricte la loterie $\alpha L_1 + (1-\alpha)L_2$ pour un $\alpha \in]0, 1[$. L'inverse peut se produire aussi avec L_1 et L_2 préférées à la loterie composée $\alpha L_1 + (1-\alpha)L_2$.
Toutes ces approches relâchant les conditions de linéarité et/ ou de parallélisme des courbes d'indifférence aboutissent à des fonctions d'utilités V de la forme

$$V(L) \quad = \quad V(x_1, p_1; \cdots; x_n, p_n) \quad \text{avec V une fonction non linéaire.}$$

Par exemple, on peut avoir une fonction de la forme :

$$V(L) = \frac{\sum_{i=1}^{n} p_i.u(x_i)}{\sum_{i=1}^{n} p_i v(x_i)},$$

où u et v sont des fonctions donnant des poids non nuls aux conséquences.

On est, alors, dans le cadre de la théorie de l'utilité pondérée qui englobe celle de l'espérance d'utilité. Cette dernière est obtenue en prenant v comme la fonction constante égale à 1. Pour des fonctions plus réalistes sur v, cette théorie permet de résoudre le paradoxe d'Allais.

La théorie subjective de l'utilité pondérée, est obtenue quant à elle, à travers la fonction V définie par

$$V(L) = \frac{\sum_{i=1}^{n} \pi(p_i).u(x_i)}{\sum_{i=1}^{n} \pi(p_i)} \ .$$

Notons que l'on ne travaille plus directement sur les probabilités objectives p_i mais plutôt sur des probabilités subjectives dénotées ici par $\pi(p_i)$, $i = 1, \cdots, n$.

La théorie de l'utilité dépendant du rang est une autre formulation qui a été proposée dans la littérature basée aussi sur une distorsion de la probabilité objective comme dans celle de la théorie subjective de l'utilité. La transformation s'effectue cependant sur les probabilités cumulées. Comme conséquence, les problèmes de violation de la condition de dominance stochastique du premier ordre sont évités. La fonction d'utilité V est de la forme :

$$V(L) = \sum_{i=1}^{n} u(x_i) \Big[\pi(\sum_{j=1}^{i} p_j) - \pi(\sum_{j=1}^{i-1} p_j) \Big].$$

Comme son nom l'indique, l'utilité est définie en tenant compte de la position de x_i par rapport aux autres conséquences (x_j), $j \neq i$. Une modification de x_i qui ne perturbe pas l'ordre des conséquences a peu d'impact sur l'utilité des personnes. Par contre, une petite modification même minime de x_i qui modifie le rang des (x_j) aura un impact majeur pour les agents.

L'axiome central dans cette théorie est celui de la chose sure comonotone. Il stipule que si deux loteries $P = (x_1, p_1; \cdots; x_n, p_n)$ et $Q = (y_1, p_1; \cdots; y_n, p_n)$ rangées dans l'ordre croissant ont une conséquence commune disons $x_k = y_k$ alors l'ordre de préférence de ces deux loteries reste inchangé si on remplace x_k par toute autre conséquence x ne perturbant pas l'ordre initial i.e $x_{k-1} < x < x_{k+1}$, $y_{k-1} < x < y_{k+1}$.

Enfin, d'autres approches sont plutôt proches de la psychologie. C'est le cas de la théorie de la déception ou du regret. Pour la théorie de la déception, l'idée est de dire que toute personne dans un environnement de risque a une valeur de référence (valeur attendue) et qu'il aura une déception plus ou moins grande si la valeur ex-post obtenue est inférieure à celle qui était attendue. Pour une axiomatisation de la théorie de la déception, on peut voir Chauveau et Nalpas [19] et les références qui y sont citées. La fonction V représentant les préférences est de la forme :

$$V(L) = \sum_{i=1}^{n} p_i \Big[u(x_i) + D(u(x_i) - \bar{U}) \Big].$$

Le terme $\bar{U}$ représente l'utilité espérée de l'individu et D est une fonction qui permet de pondérer les écarts entre les utilités $u(x_i)$ et $\bar{U}$.
Ce type de modèle étend aussi la théorie de l'utilité espérée qui est obtenue en prenant D comme

étant la fonction nulle. D'autre part, on peut donner des propriétés différentes à la fonction D tenant compte de la déception pour $u(x_i) \leq \bar{U}$ et de la satisfaction pour $u(x_i) \geq \bar{U}$.

Pour terminer, on note que Machina [82] a généralisé la notion d'utilité espérée en la définissant de façon locale. Aucune spécification de la fonction V n'est faite. On exige simplement que les courbes d'indifférence soient lisses. Ce qui permet d'avoir des représentations assez flexibles comme celles montrées dans la figure 2.2. L'idée est de remplacer $u(x_i) = \frac{\partial V(L)}{\partial p_i}$ dans la théorie de l'utilité espérée par $u(x_i, L) = \frac{\partial V(L)}{\partial p_i}$ mettant en relief le caractère local des courbes d'indifférence.

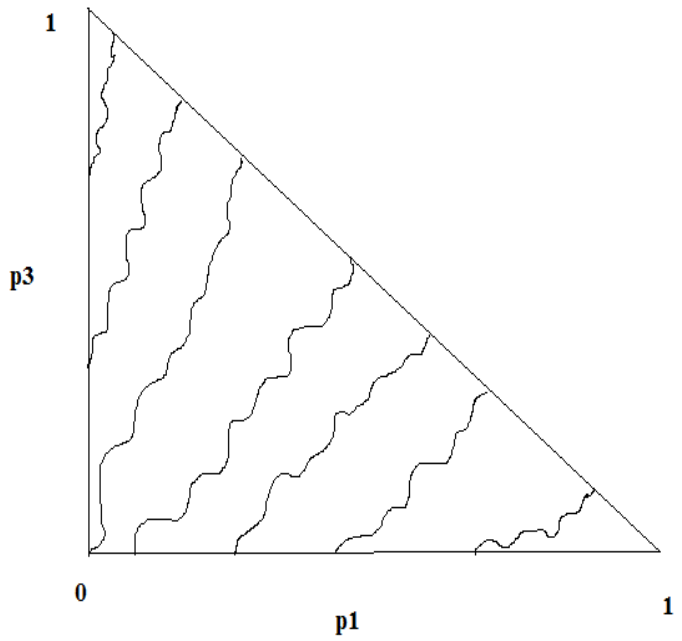

FIGURE 2.2 – Courbes d'indifférence dans le cas de la théorie généralisée de l'utilité espérée.

2.2 Psychologie.

Le domaine de la psychologie va au-delà de la quasi-normativité au sens où on ne tente pas de rationaliser les comportements des agents en donnant des axiomes. Formellement, il a été noté que les individus pouvaient se départir de l'axiome de la transitivité qui joue un rôle fondamental dans la finance standard. D'autre part, les décisions des agents étaient dépendantes de la manière dont les loteries sont formulées. Par exemple, remplacer un pourcentage α de vies sauvées par son équivalent de $(1 - \alpha)$ de vies perdues pouvaient modifier les préférences des agents, voir Tversky et Kahneman [109]. Une explication de ces résultats empiriques est que les agents utilisent des règles simplificatrices dans leurs décisions. Cette recherche de simplicité dans les décisions fait que contrairement à l'approche conventionnelle, des concepts de base même peuvent être violés. On peut citer entre autre l'invariance par rapport à la procédure utilisée ou l'invariance par

rapport à la description de probabilités comme mentionnée dans l'exemple ci-dessus.
La psychologie est un sujet assez vaste que nous avons choisi de traiter en la repositionnant par rapport à la théorie de la finance néo-classique. Dans cette dernière, les agents étaient supposés pouvoir traiter toute l'information disponible de façon correcte. Leur prédiction était optimale au sens où elle coïncide avec l'espérance conditionnelle. Précisément, la prédiction faite à l'instant t du prix du sous-jacent S_T avec $t < T$ est donnée par

$$S(t,T) = E_P(S_T|\mathcal{F}_t) \tag{2.1}$$

où $\mathcal{F}_t$ est l'information disponible à l'instant t.

Dans ce cas, il est facile de monter que le processus de prix $(S(0,T), S(1,T)), \cdots, S(T-1,T)$ est une P-martingale par rapport à $(\mathcal{F}_t)$.[9] C'était l'affirmation de Samuelson [101] montrant que si les prix (sous entendu prix actualisés) sont correctement anticipés, alors ils suivent une marche aléatoire (martingale).
Dans l'équation (2.1), quatre remarques peuvent être faites toujours dans cette théorie.

1. D'abord il est supposé que les individus s'accordent sur les différents scénarios possibles Ω, et sur leurs probabilités d'occurrence données par P.

2. La probabilité P qui est commune à tous les investisseurs est objective au sens où elle reflète réellement la probabilité d'apparition des différents scénarios sur Ω.

3. Dans un marché financier, contrairement aux loteries, le cadre est dynamique. La probabilité P n'est pas statique et doit être mise à jour au fil et à mesure que l'information arrive. Cette mise à jour doit être telle que le point 2 soit satisfait.

4. Dans la prédiction à l'instant t, n'intervient que la probabilité objective P, la variable aléatoire S_T et l'information disponible $\mathcal{F}_t$. Le sentiment des individus, leurs humeurs, ne sont pas pris en compte.

Les points 2, 3 énoncés ci-dessus préconisent que les individus ont une capacité cognitive illimitée. Le point 4 permet de comparer ce super Homme à une machine au sens où seuls les inputs fournis sont utilisés et rien d'autres de supplémentaire. Cependant il est connu que l'Homme n'est pas une machine. Il a des sentiments, des émotions, des humeurs. Par ailleurs, sa capacité cognitive est limitée. Dans des problèmes complexes, des règles de décisions simplificatrices sont utilisées (heuristiques). La première section se consacre aux parties expérimentales/empiriques qui ont illustré quelques erreurs communes que les agents ont tendance à faire et qui violent au moins l'un des points 2, 3 ou 4.

9. $X_t = E_P(Y|\mathcal{F}_t)$ est une des façons de construire une martingale à partir d'une variable aléatoire Y intégrale par rapport à P.

2.2.1 Etudes expérimentales/empiriques.

La psychologie est une science basée principalement sur des expériences pour comprendre la manière de penser ou d'agir des individus. Nous recensons ici quelques règles de décisions qui sont issues des études expérimentales.

La représentativité.
Cette une heuristique qui réfère à la tendance qu'ont les agents à extrapoler une ressemblance. Pour le montrer, on peut considérer un objet[10] nommé disons A. Ensuite, on considère deux types de descriptions. Le premier noté B est fait de sorte qu'on trouve peu de ressemblance avec l'objet A tandis que le deuxième noté C est décrit de sorte qu'on s'approche des caractéristiques de l'objet A. Enfin, on demande aux personnes de choisir l'affirmation la plus probable entre :

1. L'objet A appartient à B.

2. L'objet A appartient à $C \cap B$.

La majorité des participants choisissent la deuxième option plutôt que la première. Cette erreur est appelée erreur de conjonction. Même si l'objet A a plus de chance d'appartenir à C qu'à B, en prenant l'intersection $C \cap B$, cet événement devient moins probable que celui de B. La présence de C dans $C \cap B$ a suscité une amplification de cet événement chez les individus.

Ce biais se matérialise dans le domaine financier par une extrapolation d'informations. Par exemple, si on a une suite de six chiffres $0, 1, 0, 1, 0, 1$, on peut inférer qu'ils viennent de la réalisation de variables aléatoires indépendantes de Bernoulli de paramètres $0,5$. Or l'information disponible n'est pas si importante pour tenir ces propos. Deux erreurs peuvent être faites. La première consiste à se tromper sur la fréquence qui peut être différente de 0,5. L'autre erreur peut être même liée à la loi ; le prochain chiffre pouvant être 4 par exemple.
Comme conséquence de ce biais, les investisseurs vont mettre beaucoup de poids sur le passé récent pouvant expliquer ainsi les effets tendance (momentum) c'est-à-dire une hausse ou baisse successive de prix a tendance à se conserver dans le court terme. Une stratégie consistant donc à acheter des actifs performants et à vendre ceux non performants pour une période passée à court terme a de forte chance d'obtenir un gain de rentabilité. Cette évidence a été montrée par Jegadeesh et Titman [59]. Ici, le point 2 est violé, dû à une mise à jour incorrecte de l'information.

La surconfiance.
Elle consiste à ce que les agents tendent à surestimer leur capacité à faire de bons jugements. S'ils supposent que l'intervalle $[a,\ b]$ contient une valeur inconnue à 95%, en réalité l'intervalle ne la contient qu'avec une probabilité moindre. Les événements qu'ils considèrent comme certains peuvent ne pas se réaliser et inversement les événements supposés impossibles peuvent se réaliser. En d'autres termes, la probabilité subjective qu'ils mettent sur un événement A est différente de la probabilité objective (fréquence de l'événement A). Ici, le point 2 n'est pas satisfait.
Une application de ce biais dans les marchés financiers est que les agents vont réviser lentement leur croyance vu qu'ils sont convaincus de leur signal privé. Cela peut ainsi expliquer un excès de volume d'échanges si on est dans un milieu où les investisseurs sont hétérogènes. Dans le cadre

10. Le terme objet est utilisé au sens informatique ou il permet de représenter des éléments ayant des caractéristiques similaires. Il peut être une personne, une voiture, un animal etc.

traditionnel de la finance néo-classique, même si on part sur l'hypothèse d'hétérogénéité, il est supposé que les croyances vont s'uniformiser du fait que chaque agent va intégrer les signaux émis par les autres (pas de surconfiance). Au bout, on obtient un volume d'échanges faibles.
Pour illustration de ce biais, on réfère Odean [92], De Bondt et Thaler [15]. Les études sont basées essentiellement sur des questionnaires. Il a été noté que ce biais de surconfiance est assez robuste au sens où il touche plusieurs catégories de personnes telles que les agents de la santé, des ingénieurs, banquiers entre autres.

Le conservatisme.
C'est le fait que les agents ne donnent pas beaucoup d'importance aux informations qui ne collent pas à leur croyance et sur-évaluent celles qui confirment leurs opinions. C'est une heuristique disons inverse à celle de la représentativité. Ici, les agents modifient rarement leur croyance à moins qu'il y ait vraiment un signal fort allant dans le sens inverse de leur vision. Donc le point 2 est violé dû à une mise à jour incorrecte de l'information (point 3). Un autre biais qui est lié au conservatisme est celui de confirmation. Si des informations nous arrivent en contradiction avec nos croyances, on a deux choix. Le premier consiste à l'ignorer comme dans le conservatisme. La seconde est de l'accepter tout en cherchant à réduire la dissonance cognitive. Ce qui veut dire que l'on sera à la recherche d'informations futures tendant à confirmer notre opinion. Ce type de biais peut expliquer comme l'ont souligné Goetzman et Peles [47] le fait que les individus sont plus enclins à injecter de l'argent dans des fonds performants que de s'en retirer dans le cas de fonds perdants. Le conservatisme quant à lui peut se répercuter dans les marchés financiers par des sous- réactions (sous-pondération de l'information récente) par exemple face aux annonces des bénéfices d'entreprises.

L'ancrage.
Les agents en faisant des estimations commencent généralement par un point référence qui peut être aléatoire où un ajustement est ensuite fait. Pour illustration, Kahneman et Tversky [63] demandent aux participants d'estimer le nombre de pays africains membre de l'organisme des Nations unies (ONU). Ils leur demandent de faire leur évaluation par rapport à un nombre tiré de façon aléatoire entre 0 et 100. Les résultats de cette expérience montrent que le nombre tiré influe beaucoup sur l'estimation faite par les agents. Pour un premier groupe dont le numéro aléatoire tiré était 10, leur estimation médiane était de 25. Pour un second groupe dont le nombre tiré était 65, l'estimation médiane du groupe a donné 45. Dans les marchés financiers, plusieurs points peuvent être pris comme référence. Cela peut être le cours de clôture de la veille, du week-end passé, de l'année passée ou le prix auquel on a acheté notre action entre autres. C'est un point fondamental dans la finance comportementale. Le point 4 est ici violé au sens où l'évaluation fait intervenir autre chose que le pay-off S_T précisément le point de référence de l'individu.

L'Optimisme.
Les psychologues ont trouvé que les humains ont tendance à donner une probabilité inférieure à celle objective qu'un événement malheureux leur touche (cambriolages, accident, maladie, etc.). Ce fait traduit leur optimisme et explique d'une certaine manière pourquoi les agents sont prêts à jouer sur une loterie dont l'espérance de gain est inférieure au prix du billet. Même si la probabilité (objective) de gain du gros lot est très faible, il se donne une probabilité plus élevée d'y faire parti. C'est une heuristique qui met le point 2 en cause. La transformation de la probabilité objective est plus prononcée au niveau des queues (valeurs extrêmes). Ce biais peut se répercuter dans les marchés financiers en termes de surévaluation. En effet, les probabilités d'obtenir des résultats favorables sont augmentées et celles correspondantes aux cas défavorables réduites. Bien

entendu, à côté de ce biais d'optimisme, on peut associer celui du pessimisme où les agents font l'opération inverse.

Biais de disponibilité.
Ce biais a été mis en évidence par Kahneman et Tversky [63]. Considérons un espace probabilisé $(\Omega, \mathcal{F}, P)$ où le cardinal de Ω est fini. Si la capacité cognitive des agents était illimitée alors la probabilité de tout événement disons A est déterminée par

$$P(A) \;=\; \frac{card\ A}{card\ \Omega} \;=\; \frac{card\{\omega \in \Omega, \quad \omega \in A\}}{card\ \Omega}.$$

Il se trouve, cependant, que nos capacités ne le sont pas. La probabilité de l'événement A sera évaluée donc par la facilité de trouver des ω dans A. Plus les agents peuvent donner d'exemples qui satisfont A, plus cet événement aura une probabilité grande. Dans ce cas, on peut avoir $A \subset B$ et $P(A) > P(B)$ s'il est plus facile de trouver des exemples dans l'événement A que dans celui de B. Comme illustration, les auteurs ont posé à des sujets la question suivante :

Dans un échantillon aléatoire de textes en langue anglaise, est -il plus probable d'avoir K en première lettre du mot ou en troisième position (exclusion faite des trois lettres). ?

Les résultats montrent que prés de 70 % (105 sur 152) des personnes ont opté pour la première alternative alors qu'en réalité celle-ci est en termes de probabilité deux fois moindre. Des expériences similaires ont été faites aboutissant aux mêmes résultats. Ici, le point 2 n'est pas satisfait.

Biais de disposition.
Il est aussi appelé l'erreur du parieur. Ce biais est relié au point 3 au sens où les agents en faisant une mise à jour de leurs informations appliquent souvent un retour vers la moyenne à court terme. En considérant une suite de Bernoulli de paramètre connue 1/2, alors pour une séquence 01101111, les individus ont tendance à donner une probabilité plus élevée qu'au prochain lancer un 0 apparaisse. En réalité, cette probabilité est la même pour qu'un 1 apparaisse de nouveau à savoir 0,5 vu que que les tirages sont indépendants.
Dans les marchés financiers, ce biais peut entraîner des individus à suivre une stratégie contrarienne. Ils vendent les titres gagnants à court terme et achètent les titres perdants.

Biais du à des phénomènes extérieurs.
Il existe toute une catégorie de biais qui entre dans le dernier point 4 telle que les émotions ou les humeurs des individus. A premier point de vue, l'émotion ne peut aller avec un comportement rationnel. Des études récentes basées sur le domaine de la neuroscience ont montré cependant que tel n'est pas le cas. C'est une branche qui utilise des procédés tels que la tomographie par émission de positons pour étudier l'activité métabolique d'un organe. De ces études, il en ressort que les émotions permettent de prendre rapidement de bonnes décisions. Cette rapidité est due au fait que les individus partent directement sur l'essentiel en accordant moins d'importance aux autres détails. Ce fait a été illustré par le neurologue Damasio [29] à partir d'individus souffrant de pathologies du cerveau ayant affecté le facteur émotionnel. Pour un application de cette émotion dans les marchés financiers, on peut se référer à André Lo [79].

Un autre facteur, qui entre en jeu dans les décisions, est l'humeur des individus. Selon qu'ils soient dans un bon ou mauvais état, leur perception du risque change. Cette modification du risque apparaît soit par une déformation des probabilités objectives soit par une modification de l'ensemble des conséquences. Pour un individu de bonne humeur, il aura tendance à minimiser les risques et par conséquent, prendra des positions plus risquées par l'achat par exemple d'actions. Pour mesurer l'impact de cette humeur dans les marchés financiers, des facteurs d'approximation sont pris tels que l'effet soleil (jours ensoleillés), l'effet week-end, l'effet congés pour voir s'ils ont un impact sur la rentabilité future des actifs. Ici, les études sont empiriques plutôt qu'expérimentales. Hirshleifer et Shumway [56] regardent si l'effet soleil, les temps nuageux ou la pluie pouvaient être reliés aux rentabilités des actifs. Leur étude couvre plusieurs places financières bien réparties géographiquement. Les résultats obtenus montrent en général que le beau temps a une influence positive sur les rentabilités futures des marchés. L'explication donnée tient au fait que le temps ensoleillé met les gens de bonne humeur et donc plus enclins à devenir optimistes se matérialisant par des positions longues sur les marchés financiers. Une variation d'ailleurs consistante de performances (rentabilités annualisées) a été trouvé entre les beaux et mauvais temps pour beaucoup de places financières.

A côté de l'effet soleil, des études ont établi l'effet lune sur le comportement des individus, voir par exemple Yuan, Zheng et Zhu [114]. Les auteurs considèrent un champ vaste de pays bien répartis aussi bien géographiquement qu'industriellement. Ils montrent que la performance des actifs dans la phase nouvelle lune était supérieure à celle de la pleine lune. Une explication qui en ressort est que pendant les périodes de pleine lune, les individus ont tendance à s'adonner à des activités de violence, à boire beaucoup d'alcool pouvant aboutir à des crises de nerfs et donc à des rentabilités moindre sur les marchés financiers.

D'autres facteurs, indépendamment de la lumière, peuvent influer le comportement des agents. Il s'agit de l'effet week-end et congés. Souvent, la mauvaise humeur peut nous rattraper si on s'apprête à reprendre le travail le Lundi après un week-end de tout repos. L'impression inverse se produit si on sait que le week-end s'approche. Des études ont essayé de voir si cet effet week-end à un impact sur les comportements des individus. Comme pour l'effet lumière, il est difficile d'apporter une réponse au niveau micro et souvent la question est étudiée au niveau macroéconomique en regardant si les rentabilités sont influencées selon que l'on est en début ou fin de semaine. French [39] donne une réponse positive en montrant que les rentabilités des actifs étaient négatives le Lundi et positive les autres jours. Ce résultat était presque le même indépendamment des sous périodes considérées sur la bourse de New-York.
L'explication qu'il a donné ne se rapporte pas par contre au facteur humeur. Pour lui, les mau-

vaises informations sont souvent données après la fermeture du marché le vendredi. Cela permet aux investisseurs de digérer cette information et donc de diminuer l'impact de la baisse pour l'ouverture du marché le Lundi. Les bonnes nouvelles quant à elles apparaissent dans la semaine vu qu'une surévaluation de l'information est à mettre à leur avantage. C'est ce phénomène qui explique donc la variabilité des rentabilités en début et fin de semaine.
Le phénomène de week-end entre dans un cadre plus global de fin de cycle. Il a été découvert que les variabilités des rentabilités dépendaient des fins du mois et d'années. Enfin, il a été noté pour tous ces phénomènes de fin de cycle, ils sont plus accentués pour les petites capitalisations que pour les grandes. Pour des informations complémentaires, on peut se référer au livre de M. Mangot [83] qui fournit assez de détails sur la psychologie et son lien avec les marchés financiers.

Dans cette sous section, on a rappelé quelques études empiriques ou expérimentales qui ont tendance à rejeter l'hypothèse d'efficience de marché. Ces études montrent que les agents commettent des biais dans leur raisonnement. Dans la suite, on s'intéresse à quelques modèles qui ont intégré ces résultats issus de la psychologie pour une application soit au niveau des préférences des agents soit au niveau de la formation des prix.

2.2.2 Modèle de Kahneman et Tversky.

Les études expérimentales ont montré que les règles de décisions (heuristiques) des individus aboutissent souvent à des transformations de la probabilité objective P. Il est alors utile de favoriser les modèles basés sur les distorsions de probabilité de la forme $\phi(P)$ où ϕ est une fonction différente de la fonction identité. Une question naturelle est comment doit être la forme de la fonction ϕ de sorte à pouvoir expliquer le maximum de phénomènes. La littérature montre d'abord que les violations sont souvent obtenues au bord du triangle de Machina qui été représentée dans la figure 2.1. Quand on se trouve dans la partie centrale du triangle souvent les prédictions de la théorie de l'utilité espérée sont vérifiées. Par contre, lorsque des loteries présentant des probabilités extrêmes sont utilisés, on s'en écarte. D'autre part, la sensibilité des résultats est assez forte au niveau de ces probabilités extrêmes. Ce qui veut dire que la fonction choisie devrait permettre une flexibilité en termes de courbure en ces endroits. De plus, pour les petites probabilités p, il a été noté que la fonction ϕ qu'utilisent les sujets vérifie $\phi(p) > p$ (optimisme ou surconfiance des agents). Un consensus qui en est ressorti de ces remarques est une forme de S renversé pour la fonction ϕ dans l'intervalle $[0, 1]$. Tversky et Kahneman [110] proposent la fonction de pondération suivante :

$$\phi(p) = \frac{p^\delta}{\left(p^\delta + (1-p)^\delta\right)^{1/\delta}}, \qquad 0 < \delta < 1. \tag{2.2}$$

Pour $\delta \neq 1$, la fonction ϕ est non linéaire par rapport à p. Nous donnons quelques illustrations de cette fonction pour les valeurs suivantes $\delta = 0,4;\ 0,5;\ 0,6;\ 1$ dans la figure 2.3.

D'autres variantes pour représenter la forme inversée de S ont été proposées. Par exemple, Prelec [96] considère la fonction ϕ définie par

$$\phi(p) = \exp(-(-\log p)^\alpha), \qquad 0 < \alpha < 1.$$

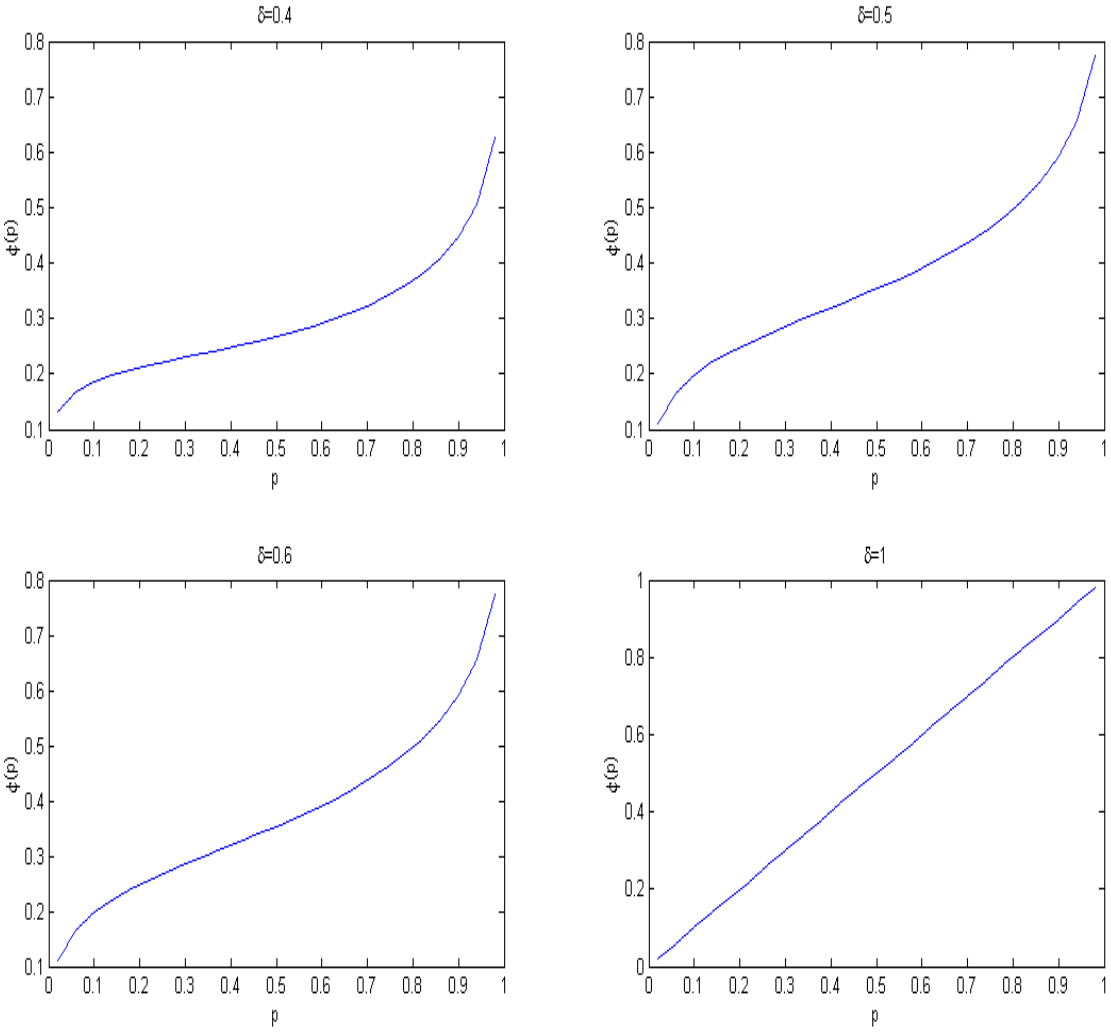

FIGURE 2.3 – Différentes valeurs de δ pour la fonction de pondération de Kahneman et Tversky.

Le point important qui différencie surtout la théorie du prospect des autres est l'ajout d'un point de référence. Selon que l'on est au dessus ou en dessous de ce point, la préférence des individus change. Les agents ne maximisent pas la valeur finale de leur portefeuille comme le suggère la théorie de l'utilité espérée mais plutôt les variations du portefeuille (gains, pertes) par rapport au point de référence. Les auteurs définissent la fonction V représentant les préférences des individus de la manière suivante :

$$V(L) = u(x_1)\phi(p_1) \ + \ \cdots \ + \ u(x_n)\phi(p_n). \tag{2.3}$$

où ϕ est de la forme (2.2) et u est concave pour des conséquences positives et convexe pour des conséquences négatives. Une illustration de la fonction u est donnée dans la figure 2.4.

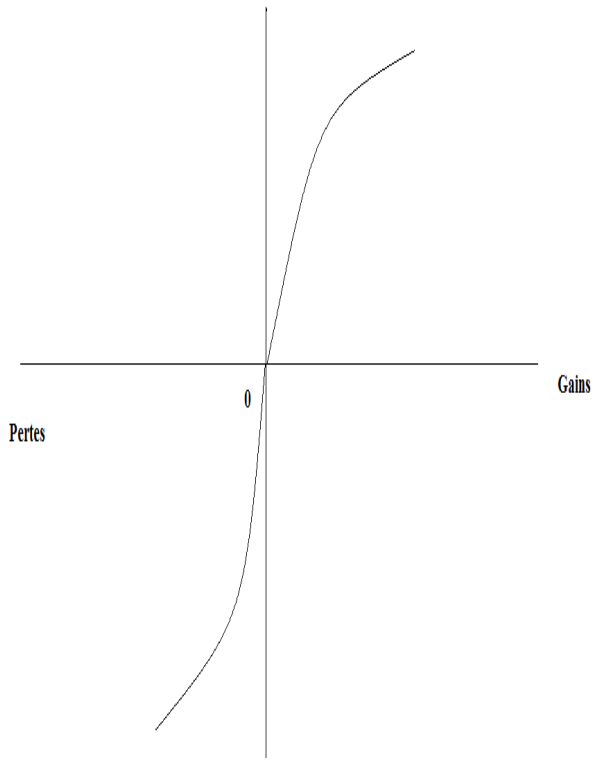

FIGURE 2.4 – Forme de la fonction d'utilité dans la théorie du prospect.

Remarquons que la fonction de distorsion ϕ peut entraîner une violation de la condition de monotonicité du fait que $\phi(p_i + p_j) \neq \phi(p_1) + \phi(p_2)$. Les auteurs se sont adaptés à la théorie de l'utilité dépendant du rang donnant lieu à la *Cumulative Prospect Theory*.
Ils proposent pour la fonction u représentée dans la figure 2.4 la forme suivante :

$$u(x) = \begin{cases} x^{\alpha} & \textit{pour } x \geq 0. \\ -\lambda(-x)^{\alpha} & \textit{pour } x < 0. \end{cases}$$

Par ailleurs, deux modèles importants ont été développés en 1998 pour expliquer ce qui semble être un contraste à la théorie de l'efficience de marché à savoir la surévaluation ou la sous évaluation des prix d'actifs. Ces modèles, bien qu'utilisant des heuristiques différentes, arrivent à expliquer simultanément ces deux anomalies. Le premier modèle est celui de Barberis, Shleifer et Vishny [11]. Il se fonde sur les heuristiques de conservatisme et de représentativité. Par exemple, le conservatisme permet une mise à jour lente des nouvelles informations qui arrivent sur le marché. Les individus qui disposent de ce biais ont souvent une croyance qui peut être vue comme un point de référence où ensuite un ajustement faible est fait. Souvent ce réajustement reste insuffisant et n'intègre pas correctement l'information réelle de sorte les prix dévient de leurs valeurs fondamentales. Pour le biais de la représentativité, on a l'opération inverse. Une petite information est considérée comme représentative (loi du petit nombre).

Le second modèle, qui arrive à expliquer aussi bien la sous évaluation et la surévaluation des prix d'actifs, est celui de Daniel, Hirshleifer et Subramanyam [30]. Dans le modèle précédent les auteurs avaient utilisés les biais de conservatisme et de représentativité. Ici, les auteurs se basent plutôt sur les biais de la sur-confiance et de l'auto-attribution. Le modèle suppose la présence d'agents informés (C) qui sont neutres au risque et des agents non informés qui sont averses au risque (U). Le modèle arrive à générer une corrélation positive au niveau de la variation des prix (effet momentum) à court terme dû à la sur-réaction des agents informés. Au fur et à mesure que l'information publique devient abondante, plus l'incertitude au niveau de la valeur fondamentale diminue. Ce fait entraîne une correction de la surévaluation à long terme se matérialisant par un effet revirement.

Supposons qu'un actif soit sous-évalué ou sur-évalué dû par exemple aux explications données par Daniel, Hirshleifer et Subrahmanyam ou par Barberis, Shleifer et Vishny. En présence d'agents rationnels, ils interviendront pour profiter des inefficiences et rendre donc le marché efficient. Par conséquent, la présence de comportements irraisonnés n'entraîne pas nécessairement des imperfections au niveau macro-économique. Pour qu'un tel fait soit possible, il faudra montrer que ces agents bruiteurs peuvent survivre longtemps dans un marché compétitif. C'est l'objet de la section suivante.

2.3 Limite du critère d'arbitrage.

Dans la partie psychologie, on a vu que les agents avaient tendance à faire les mêmes types d'erreurs. Cela se matérialise dans le cas des marchés financiers par une déviation des prix d'actifs de leurs valeurs fondamentales. Cependant, les expériences étaient faites dans un cadre statique différent de celui des marchés financiers qui offre un cadre dynamique. Dans un tel contexte, des imperfections peuvent-elles encore exister ? Pour la finance néo-classique, la réponse est non car dans un cadre de concurrence, seuls les meilleurs vont demeurer comme dans la théorie de Darwin. D'autre part, il y aura un phénomène d'apprentissage qui prendra place chez les agents irrationnels. Cet argument bien que convaincant n'est pas toujours suffisant. En réalité,

les opportunités d'arbitrage sont risquées et coûteuses. D'autre part, le processus d'extinction des agents irrationnels n'est pas automatique, car ils peuvent prendre plus de risque et avoir par chance plus de rentabilité par rapport aux agents rationnels. Ce fait leur assure une survie au moins pour le court terme. On illustre ces propos à travers De Long et al. [81] qui présentent un modèle parmi tant d'autres expliquant comment ces agents bruiteurs peuvent survivre à un environnement concurrentiel et dynamique.

2.3.1 Exemples de Modèle.

De Long et al. [81].

De Long, Shleifer, Summers, Waldmann [81] ont développé le concept de *Noise trader risk*. Il s'agit du risque associé à l'imprévisibilité du comportement des agents irrationnels à court terme. La description de leur modèle est la suivante. Ils considèrent un marché dans un cadre discret où à chaque instant t, il existe deux types de générations : les jeunes et les âgés. Chaque génération ne vit que sur deux périodes. Les personnes âgées à l'instant t disparaissent à l'instant $t+1$. De même les jeunes à l'instant t qui n'existaient pas à l'instant $t-1$ deviennent âgés en $t+1$. Le marché est constitué de deux types d'actifs donnant les mêmes dividendes disons r. L'un s est sans risque et sans aucune contrainte sur la quantité offerte et l'autre u est risqué avec une contrainte sur la quantité offerte disons normalisée à 1. On suppose que le prix de l'actif s pour " safe" est $p_t = 1$ pour tout t. La valeur fondamentale de u pour "unsafe" est aussi de 1 vu qu'ils ont les mêmes distributions de dividendes. A l'instant t, les jeunes choisissent à partir de leur richesse normalisée à 1 un portefeuille composé des deux actifs u et s. Ces actifs leur sont vendus par les personnes âgées qui consomment automatiquement l'argent tiré de cette vente. L'utilité de tous les investisseurs provient de cette consommation qu'ils auront à faire quand ils seront âgés. Tout jeune cherche donc à construire un portefeuille de sorte à maximiser sa future consommation. Leurs fonctions d'utilité sont de type exponentielle (CARA). Pour faire dévier les prix de l'actif u de sa valeur fondamentale 1, il est supposé la présence d'agents rationnels et irrationnels. Ces derniers commettent des erreurs sur la rentabilité associée à l'actif u modélisées par une variable aléatoire $\rho \sim N(\rho^*, \sigma_\rho^2)$. Quand ρ^* est positive, les agents irrationnels sont optimistes sur la bonne performance de l'actif u se matérialisant par une position plus élevée en actifs risqués par rapport aux agents rationnels. Ils peuvent aussi avoir une vision pessimiste en moyenne correspondant à $\rho^* < 0$ et donc à une baisse de leur exposition sur l'actif u. Le modèle décrit par les auteurs De Long et al. est résumé dans le tableau ci-dessous.

	Date t, génération jeune		Date $t+1$, génération âgée	
	Agents rationnels	Agents irrationnels	Agents rationnels	Agents irrationnels
Investissement	λ_t^a en actif u $1-\lambda_t^a$ en actif s	λ_t^n en actif u $1-\lambda_t^n$ en actif s	— —	— —
Consommation	— —	— —	$\lambda_t^a(p_{t+1}+r)$ $+(1-\lambda_t^a)(1+r)$	$\lambda_t^n(p_{t+1}+r)$ $+(1-\lambda_t^n)(1+r)$

La rentabilité supposée normale de l'actif risqué u combinée au fait que l'utilité de tous les types d'investisseurs est exponentielle entraîne que

$$\lambda_t^a = \frac{r \ + \ {}_t p_{t+1} \ - \ (1+r)p_t}{2\gamma. \ {}_t\sigma^2_{p_{t+1}}} \tag{2.4}$$

$$\lambda_t^n = \frac{r \ + \ {}_t p_{t+1} \ - \ (1+r)p_t}{2\gamma. \ {}_t\sigma^2_{p_{t+1}}} \ + \ \frac{\rho_t}{2\gamma.{}_t\sigma^2_{p_{t+1}}} \tag{2.5}$$

où ${}_t p_{t+1} = E_t(p_{t+1}), \quad {}_t\sigma^2_{p_{t+1}} = E_t(p_{t+1} - E_t(p_{t+1}))^2$ représentent respectivement l'espérance et la variance conditionnelle de l'actif risqué à l'instant t, γ représente le coefficient d'aversion au risque.

En utilisant le fait que les agents âgés vendent toutes leurs quantités en actif u normalisées à 1 à la jeune génération, on a $\lambda_t^a \ + \ \lambda_t^n = 1$. Cette égalité ajoutée aux définitions de λ_t^a et λ_t^n données respectivement dans les équations (2.4) et (2.5) entraîne que

$$p_t = \frac{1}{1+r} \ \ [r +_t p_{t+1} - 2\gamma \ {}_t\sigma^2_{p_{t+1}} + \mu\rho_t].$$

Ils montrent que les prix donnant des équilibres stables vérifient

$$p_t = \ 1 \ + \frac{\mu(\rho_t - \rho^*)}{1+r} \ + \ \frac{\mu\rho^*}{r} - \frac{2\gamma\mu^2\sigma_\rho^2}{r(1+r)^2}. \tag{2.6}$$

La valeur fondamentale de l'actif est ici représentée par 1 et le reste tient compte de l'effet des agents irrationnels. Le premier facteur $\frac{\mu(\rho_t-\rho^*)}{1+r}$ mesure l'impact d'une variation de l'opinion des agents irrationnels sur les prix de l'actif u. Quand une génération donnée a une croyance plus optimiste (pessimiste) par rapport à la croyance moyenne ρ^*, le prix de l'actif u est alors poussé à la hausse (baisse). Le facteur μ montre que cette déviation des prix est proportionnelle au nombre d'agents irrationnels présents dans le marché.
Le terme $\frac{\mu\rho^*}{r}$ montre l'effet de la croyance moyenne des individus irrationnels. Si ρ^* est nul alors il n'y a aucun impact sur les prix. Par contre, un optimisme ou un pessimisme en moyenne entraîne une hausse respectivement une baisse des prix.
Le dernier terme de l'équation (2.6) a un effet de compensation sur les prix du fait que les agents irrationnels ont ajouté un risque supplémentaire non lié aux fondamentaux. En effet, les individus rationnels ne sont pas prêts à accepter ce risque supplémentaire sur l'actif u sans une compensation matérialisée par une diminution de prix.
Bien qu'ils savent que le prix de l'actif ne reflètent pas sa valeur fondamentale, les agents rationnels ne sont pas motivés à jouer à fond sur cette inefficience à l'instant t. Un inconnu leur échappe à savoir le comportement de la nouvelle génération à l'instant $t+1$ qui leur achètera leurs actifs. Ce qui fait que leurs actions seront limitées d'autant qu'ils sont averses au risque et que leur horizon de temps est fini.

Les auteurs étudient en deuxième lieu la possibilité que les agents irrationnels puissent gagner plus que les agents rationnels. Ils montrent que l'excès de rentabilité espérée entre ces deux types d'investisseurs est de la forme

$$E(\Delta R_{n-a}) \;=\; E(R_n - R_a) \;=\; \rho^* - \frac{(1+r)^2(\rho^*)^2 \;+\; (1+r)^2\sigma_\rho^2}{2\gamma\mu\sigma_\rho^2} \tag{2.7}$$

L'équation (2.7) montre qu'une condition nécessaire pour que les agents fassent mieux en termes de rentabilité espérée est qu'ils soient en moyenne optimistes $\rho^* > 0$. Cet optimisme ne doit pas être trop excessif dû à la présence d'un terme en $(\rho^*)^2$ dans la partie négative. Par ailleurs un coefficient d'aversion γ élevé, une fraction d'agents irrationnels μ importante et une variance σ_ρ^2 élevée sont requis pour affaiblir le second terme négatif. Dans ce cas, les agents bruiteurs peuvent prendre le dessus. La performance s'explique alors par une prise de risque élevée accompagnée par de la chance. Bien entendu, cette performance ne peut continuer à long terme. Le fait de prendre plus de risque entraîne une forte probabilité de subir des pertes et donc de disparaître.

Une application naturelle de leur travail est l'explication de la violation de la loi du prix unique dans les marchés financiers. Un exemple d'illustration est celui des fonds fermés. Comme son nom l'indique, c'est un fond dont le nombre de parts reste fixe. Dans les normes, la valeur du fond à l'instant t, notée V_t, devrait être égale à la somme du nombre de parts de chaque titre multiplié par son prix respectif i.e $V_t = \sum_{i=1}^{n} N_i.P_t^i$. Sauf que dans la pratique, cette égalité est souvent violée et une décote assez importe peut se présenter. Dans ce cas, il suffirait de vendre le portefeuille pour acheter son équivalent moins cher à savoir les n titres. Si tout le monde s'y met, par la loi de l'offre et de la demande, cet écart aura tendance à diminuer jusqu'à ce que l'égalité $V_t = \sum_{i=1}^{n} N_i.P_t^i$ se rétablisse. Si ces décotes encore appelés discounts persistent c'est que les arbitragistes ne se sont pas mis pleinement dus à la présence de facteurs de risque. Par exemple, la mauvaise perception des titres individuels par rapport aux portefeuille global peut augmenter à court terme chez les agents irrationnels entraînant un discount plus élevé au lieu d'une correction.
Par ailleurs, des explications rationnelles ont été fournies pour cette décote. Entre autres, elle peut être due à un manque de liquidité des titres de sorte que celui qui dispose du portefeuille et décide de le liquider à l'instant t mettra une grande quantité de titres dans le marché entraînant une baisse des prix si le marché est illiquide. Une autre explication est due à des raisons fiscales. Le fait de vendre une forte quantité de titres entraînera une rentrée d'argent qui sera imposable. Cet argent allant au fisc représente en quelque sorte le discount constaté. Par ailleurs, des décotes ont été notées sur des actions individuelles où des problèmes de liquidité ou de fiscalité n'interviennent pas. Froot et Dabora [41] ont montré que Royal-Dutch et Shell Transport, deux sociétés qui ont fusionné en 1907 avec des proportions respectives de 60 % et 40% avait des ratios qui déviaient de leurs valeurs fondamentales de -35% à 15%. Cette déviation importante est donc plus à mettre sur le comportement non prévisible des agents irrationnels à court terme. La finance néo-classique pourra voir cet exemple comme un fait isolé mais qu'en moyenne on peut retenir que les prix des actifs reflètent leurs valeurs fondamentales.

Dans le cas d'une action n'ayant pas de substitut parfait alors un risque supplémentaire fait surface. Considérons par exemple deux actifs A et B d'un même secteur de sorte l'actif B ne réplique pas l'actif A. Si l'actif A est sous évalué par rapport à l'actif B alors le portefeuille

$[A, -B]$ [11] ne nous assure pas un gain systématique. En effet, ce type de portefeuille se protège contre toute information mauvaise qui toucherait les deux actifs. Cependant, si une mauvaise information touche l'actif A et pas l'actif B ou réciproquement une bonne information touche l'actif B et pas l'actif A, alors on obtient une perte. Ce risque dû au fait que les deux actifs ne sont pas des substituts est appelé **risque fondamental**.

On termine par rappeler d'autres facteurs qui peuvent limiter les opportunités d'arbitrage.

Risque de synchronisation.

Ce risque a été introduit par Abreu et Brunnermeier [1]. Il consiste à dire qu'un problème de coordination peut exister entre les arbitragistes. Si une anomalie apparaît, il faut un nombre important d'arbitragistes pour corriger l'inefficience. On encourt un risque si on est seul ou assez peu à agir, par exemple, en intervenant tôt. D'autre part, on ne doit pas retarder beaucoup notre action au risque de ne pas faire parti de ceux qui profiteront de la bulle avant éclatement. On est donc dans un dilemme qui ne favorise pas la synchronisation.

Coûts de transaction.

Généralement pour exploiter un arbitrage, on combine des achats et des ventes impliquant des coûts de transaction qui peuvent être élevés, voir Mitchell et al.[87]. Si un actif est sous-évalué on l'achète et on vend à découvert son substitut pour se protéger du risque fondamental. On ne s'aventure sur cette opération d'arbitrage que si on est sur que notre bilan sera positif après une prise en compte des coûts de transaction. Si le bilan est difficile à faire comme dans le marché des actions où la position peut changer d'un instant à l'autre, on peut opter pour l'abstention. Souvent, les individus préfèrent exploiter les opportunités d'arbitrage apparaissant dans les marchés obligataires où les flux sont presque certains. Les marchés de change sont aussi utilisés pour les arbitrages quand la banque centrale essaie de corriger un taux directeur assez élevé ou faible.

Séparation richesse-compétence.

Sleifer et Vishny [105] ont montré un autre risque lié à la séparation qui peut exister entre ceux qui disposent les fonds et ceux qui les gèrent. Un arbitragiste peut entamer une stratégie qui ne s'avère pas gagnante à court terme. S'il est incompris, les investisseurs peuvent le mettre dans le compte de l'incompétence et retirer donc leur agent. Il est alors obligé d'abandonner sa position même si son portefeuille peut à long terme générer un gain certain. Un autre point qu'ils soulignent est que l'horizon des arbitragistes est souvent court. Les gestionnaires de fonds sont classés sur une période généralement inférieure ou égale à une année. Si on n'est pas sûr de pouvoir exploiter l'opportunité d'arbitrage d'ici cette date, on préfère ne pas agir.

2.3.2 Etudes empiriques.

Les études empiriques ont pour objectif de mettre en valeur les différents faits mentionnés ci-dessus. Nous avons choisi deux exemples en guise d'illustration. Le premier est en rapport avec le

11. Ce portefeuille consiste à vendre a découvert B et à acheter l'actif A.

noise tarder risk et le risque de synchronisation. Dans la finance comportementale, on explique les anomalies par la présence d'agents irrationnels. Comme vu précédemment, celles-ci provenaient de faits psychologiques pouvant varier d'un modèle à un autre. Il est naturel de supposer que cette irrationalité est beaucoup plus importante sur des actifs possédant une incertitude grande sur leurs rentabilités futures. En effet, une incertitude par rapport à l'information accroît l'effet des agents irrationnels tout en limitant l'action des arbitragistes qui auront du mal à déterminer la valeur fondamentale de l'actif. Jiang et al. [61] montrent empiriquement ces faits en utilisant comme facteur psychologique la surconfiance pour expliquer l'accroissement des anomalies face à une incertitude élevée sur l'information. Leur argument se base entre autres sur le modèle de Daniel et al. [30]. Pour faire l'étude, on a besoin d'utiliser des facteurs d'approximation pour le facteur incertitude ou le facteur ambiguïté lié à la valeur fondamentale. Les auteurs utilisent :

- L'âge de la firme.
  Il est déterminé normalement en utilisant la date d'introduction en bourse. Pour l'étude empirique, le premier mois dès son son apparition au centre CRSP (The Center for Research in Security Prices) est pris comme origine.
  L'âge d'une firme mesure d'une certaine manière la maturité. Il est donc négativement corrélé à l'incertitude. Par ailleurs, une firme âgée dispose d'une historique de données assez longue qui facilite la prédiction future des dividendes.

- Volatilité de la rentabilité.
  Elle est mesurée par l'écart type sur une période de 25 jours ouvrables. Bien entendu, une volatilité élevée rend difficile la détermination de la valeur fondamentale limitant les actions des arbitragistes. La volatilité est donc positivement corrélée avec l'incertitude des dividendes futurs.

- Le volume d'échange.
  Il est mesuré par la moyenne des volumes d'échange journaliers durant les six (6) derniers mois en termes de pourcentage (actifs échangés/actifs détenus). C'est une mesure de dispersion des croyances d'investisseurs. Il est donc positivement corrélé avec l'incertitude liée à l'information. Dans la théorie de l'efficience, le volume d'échange est nul ou faible dû à l'homogénéité des croyances.

- La Duration.
  Elle est calculée en utilisant la méthodologie de Dechow et al. [31]. L'idée est la même que la duration utilisée dans les obligations sauf qu'elle utilise plus de paramètres. Plus la duration est grande plus l'incertitude liée à l'information est grande.

La deuxième étape consiste à regrouper en portefeuilles les firmes suivant les quatre critères cités ci-dessus. L'idée est de voir s'il existe une différence de performance entre les portefeuilles extrêmes rangés en décile. Globalement, leurs résultats confirment le fait que les effets momentum sont plus importants pour les actifs ayant une incertitude élevée. De plus, sur le long terme, ces mêmes actifs connaissent une rentabilité faible par rapport aux firmes plus matures. Cet effet de retour à la moyenne apparaît plus fréquemment sur les actifs qui sont difficilement évaluables (effet important des bruiteurs, effet limité des arbitragistes).

Une autre étude, regardant le lien entre incertitude liée aux informations et la rentabilité des actifs, est celle de Zhang [116]. La méthodologie est la même avec la présence de certains facteurs

supplémentaires. Par exemple, le facteur capitalisation est introduit dans son étude. Plus une firme est grande plus l'information la concernant est abondante. Ce facteur est donc négativement corrélé au facteur ambiguïté. Un autre point est le nombre d'analystes financiers combiné au facteur dispersion. Plus il y a de personnes qui font de la prédiction sur une firme avec moins de dispersion, plus cette cette firme comporte moins d'incertitude et de risque. Les résultats sont du même type que ceux de Jiang et al. [61]. Cependant, ces résultats ne permettent pas de créer des opportunités d'arbitrages en utilisant la stratégie de classement des portefeuilles à partir d'un critère donné comme dans le cas de Jiang et al. [61].

La deuxième application est relative à la relation Parité Call-Put. Elle a été utilisée à la fois pour tester la théorie de l'efficience au sens faible et le critère de limites d'arbitrage. Considérons un actif ne donnant pas de dividendes dont le cours à l'instant T est noté S_T. Pour tout réel K, il est facile de vérifier l'égalité mathématique suivante

$$max(S_T - K, 0) - max(K - S_T, 0) = S_T - K \tag{2.8}$$

Notons $C(S_t, K)$, $P(S_t, K)$ les prix des contrats à l'instant t ayant respectivement comme payoffs $max(S_T - K, 0)$ et $max(K - S_T, 0)$ à l'instant T. Ces contrats sont appelés Call et Put de type européen. Si la loi du prix unique est vérifiée alors l'égalité (2.8) devient à l'instant t

$$C(S_t, K) \; - \; P(S_t, K) \; = \; S_t \; - \; Ke^{-r(T-t)} \tag{2.9}$$

où r représente le taux sans risque, S_t le prix du sous-jacent à l'instant t. L'équation (2.9) est connue sous l'appellation de la relation Parité Call-Put. Le modèle se généralise facilement dans le cas où le sous-jacent donne des dividendes. Si l'égalité (2.9) que l'on peut récrire
$C(S_t, K) \; + \; Ke^{-r(T-t)} \; = \; S_t \; + \; P(S_t, K)$ est violée, on achète le membre le moins cher et on vend le plus cher. Par exemple, si $C(S_t, K) \; + \; Ke^{-r(T-t)} \; > \; S_t \; + \; P(S_t, K)$, l'arbitragiste vend un call et le combine avec un emprunt d'argent de sorte à pouvoir acheter un put et le sous-jacent. Cette opération est plus facile à réaliser que l'opération inverse faisant intervenir une vente à découvert du sous-jacent. Les études empiriques tentent de faire ressortir ce fait en montrant que la violation de la parité Call-Put n'est pas symétrique. On s'attend que l'inégalité faisant intervenir la vente à découvert $C(S_t, K) \; + \; Ke^{-r(T-t)} \; < \; S_t \; + \; P(S_t, K)$ soit beaucoup plus importante que l'autre inégalité en termes d'imperfections. Bien que la méthodologie utilisée soit la même, les études diffèrent soit du type du sous-jacent (action, indice, contrat future, etc.) soit du type d'option (européen, américain) ou soit de la fréquence des données (intra-journalière, journalière). Par exemple, Kamara et Miller [65] travaillent sur les options sur indice qui offrent plus de liquidité que les marchés d'options sur action. Ofek et al. [93] utilisent des options sur action pour voir l'impact des limites du critère d'arbitrage. Ils utilisent des options de type américain faisant intervenir dans le coté droit de l'équation (2.9) un terme supplémentaire représentant une prime. Celle-ci est due au fait que l'on peut exercer à tout instant l'option de type américain.
Pour la méthodologie, des filtres sont souvent utilisés pour enlever les contrats les moins liquides. Par exemple, des filtres de la forme $|\log(S/K)| \leq \alpha\%$ permettent de ne tenir en compte que les options qui sont proches de la monnaie. Dans le cas des données intra-journalières, il faudra

mettre un filtre qui permette de choisir des couples Call et Put cotés en des instants très proches, voir Deville et Riva [73]. Après la période filtre, la deuxième étape consiste à pouvoir expliquer les violations constatées dans la relation de parité Call-Put. Pour ce faire, on introduit le terme d'erreur E mesuré par

$$E_t \;=\; S_t \;+\; P(S_t, K) - C(S_t, K) \;-\; K\exp(-r(T-t) \;=\; S_t - S_t^*$$

où on a posé $S_t^* = \;\; C(S_t, K) \;+\; K\exp(-r(T-t) - P(S_t, K)$. D'autres versions peuvent être utilisées sans perte de généralité ($E_t = \frac{S_t}{S_t^*},\; E_t = \log S_t - \log S_t^*$). Pour les contraintes de vente à découvert on utilise des facteurs comme les taux de remise que l'on dispose en empruntant un titre (*rebate rate*). Pour des études entrant dans ce cadre, on peut voir aussi Lamont et Thaler [72] ou Ofek et al. [93].
Un autre facteur qui peut limiter les actions des arbitragistes est le facteur liquidité. Pour la parité Call-Put, on besoin de combiner plusieurs actions simultanées pour exploiter l'arbitrage. Si le marché est illiquide, la réalisation de cette tâche devient complexe et les agents rationnels préfèrent ne pas prendre de risque. De même, le volume des options, le volume du sous-jacent, le spread des options, l'échéance des options peuvent être choisis parmi tant d'autres comme des proxy de liquidité. Ayant spécifié les différents facteurs explicatifs, la dernière étape consiste à utiliser un modèle de régression pour voir si ces facteurs ont une tendance à accroître ou à réduire les violations de la parité Call-Put.

Deville et Riva [73] proposent une méthodologie assez différente au sens où ils ne travaillent pas sur la magnitude des erreurs E_t mais plutôt sur le temps nécessaire pour une opportunité d'arbitrage de disparaître. Ils regardent le premier instant où l'erreur prend forme ($E_t \neq 0$). On note cette date t_0 et on suppose sans perte de généralité que l'erreur est positive ($E_{t_0} > 0$). La durée nécessaire à la correction est mesurée par une variable notée TTNA qui est définie comme suit

$$TTNA \;=\; min\;\left\{s,\;\; s \geq t_0, E_s \leq 0\right\}.$$

Leur méthodologie utilise l'analyse de survie qui permet d'intégrer à la fois des données censurées et des variables explicatives comme dans les approches basées sur la taille des erreurs E_t.

Dans les deux premiers chapitres, on a rappelé les fondements de la Finance Néo-classique et de la Finance Comportementale. Précisément, dans le premier cas, on a une approche normative où tout investisseur maximise sa fonction d'utilité. Avec le critère d'arbitrage comme argument supplémentaire, des résultats importants ont été obtenus que ce soit en termes d'évaluation d'actifs financiers ou en termes d'allocation de portefeuille. Des études empiriques ont cependant contrasté les résultats précédents motivant le développement de la finance comportementale. Depuis lors, des controverses sont apparues surtout dans l'interprétation des anomalies. Dans le prochain chapitre qui suit, on revient sur cette problématique.

Chapitre 3

Quel paradigme choisir ?

La finance néo-classique a développé un cadre théorique basé sur les hypothèses de rationalité des agents et de marchés parfaits. Elle a fourni des résultats qui sont encore valides si on suppose la présence d'agents irrationnels ayant des actions non corrélées. Dans ce cas, seule l'action des agents rationnels va demeurer au niveau agrégé. Par contre, si les décisions des bruiteurs sont corrélées, on peut dévier des prédictions faites par la théorie. La finance néo-classique introduit alors les arbitragistes qui sont à l'affût des inefficiences et qui sont récompensés par leur travail de sorte que l'on retrouve le cadre précédent. Le débat principal repose sur ce dernier argument où il est supposé que les arbitragistes peuvent corriger le marché. Pour apporter des éclaircissements, des études empiriques ont été faites au niveau macro-économique pour voir si réellement les prix d'équilibre provenaient d'un marché où tous les agents sont supposés être comme rationnels. Un consensus ne s'est pas formé sur les résultats obtenus. On a eu à donner précédemment le cas des fonds fermés où la présence de décote a été expliquée à la fois par des comportements rationnels et irrationnels. Il en est de même pour les rentabilités variant selon les fins de cycle. Les partisans de la finance néo-classique précisent que plusieurs de ces phénomènes ont eu à disparaître entre temps expliquant ainsi l'importance du phénomène d'apprentissage.
Actuellement, il a été reconnu que le débat est lié principalement au fait que tester l'hypothèse de marché efficient c'est tester une hypothèse jointe au sens où un modèle d'équilibre doit être spécifié comme le CAPM ou tout autre modèle valide sous l'hypothèse d'efficience. Ce qui fait que si l'hypothèse jointe est rejetée, cela peut être dû à une inefficience ou simplement à une formulation incorrecte du modèle d'équilibre.
Ce que nous montrons ici est que le débat va au-delà de cette hypothèse jointe. Le réel problème est que les prix d'équilibre proviennent à la fois des arguments de la finance néo-classique et comportementale. Ce qui fait qu'il est difficile de discerner dans quel camp on se trouve. Nous illustrons d'abord nos propos à travers des études empiriques sur différents marchés pour montrer comment il est facile de basculer d'un paradigme à l'autre. En d'autres termes, la frontière qui sépare les deux camps est très fine au niveau macro-économique.[12] Ensuite, on propose un modèle théorique qui montre que même en évitant les critiques méthodologiques et la question de l'hypothèse jointe, il est difficile de trancher sur la question de l'efficience. Ces résultats nous motiveront à aborder dans la deuxième partie de ce document des modèles à agents virtuels qui puisse intégrer les deux types d'arguments.

12. Au niveau micro (description des agents), la différence est nette entre finance néo-classique et comportementale. Par contre, si on regarde les résultats fournis par ces deux paradigmes après agrégation, la différence se restreint considérablement.

3.1 Tests basés sur la marche aléatoire.

Comme rappelé à l'introduction, il existe trois types d'efficience de marché. Ici, seul l'efficience au sens faible est investie. Dans la première section, on s'intéresse aux tests liés au caractère aléatoire des rentabilités. On parlera entre autres des tests de corrélation, des tests basés sur la variance et des tests de signe (*run tests*).

Dans la deuxième section, on restera toujours dans le cadre de la prévisibilité des cours mais en proposant cette fois une stratégie pour exploiter la présence de corrélation statistique. On appliquera précisément une approche similaire à Jegadeesh et Titman [59].

L'objectif est de montrer dans les deux sections qui suivent comment on peut passer facilement d'un camp à l'autre sur de simples transformations.

3.1.1 Tests d'auto-corrélation.

Tester la marche aléatoire revient à voir si les rentabilités des titres sont indépendantes ou du moins non corrélées qui est une hypothèse moins forte. On a fait les études empiriques sur quatre marchés différents à savoir le CAC 40, le Dow Jones Industrial Average (DJIA), le Nasdaq Composite et le Nikkei 225. Les données sont prises sur le site finance de yahoo (http ://fr.finance.yahoo.com/). Elles vont du 1 Mars 1990 au 16 Octobre 2009.

On essaie de voir si à partir des rentabilités passées, on peut prédire les rentabilités futures. Pour un entier k donné, on mesure la corrélation qui existe entre la série initiale (r_t) avec la même série mais décalée cette fois d'un facteur k. Précisément, on a :

$$\rho_k \;=\; \frac{\sum_{t=1}^{N-k}(r_t-\bar{r})(r_{t+k}-\bar{r})}{\sum_{t=1}^{N}(r_t-\bar{r})^2} \;\;;\;\; \bar{r} \;=\; \frac{1}{N}\sum_{t=1}^{N} r_t. \tag{3.1}$$

L'objectif est de tester le modèle suivant :

$$r_t \;=\alpha \;+\; \rho_1\, r_{t-1} \;\cdots\; +\; \rho_s\, r_{t-s} \;+\; \epsilon_t. \tag{3.2}$$

L'hypothèse nulle est $H_0 : \rho_1 = \cdots = \rho_s = 0$.

Box et Pierce ont proposé la statistique suivante

$$Q \;=\; N\sum_{k=1}^{s}\rho_k^2, \tag{3.3}$$

où s représente le nombre total de retards.

Ljung et Box ont donné une version modifiée du test précédent pour tenir compte des échantillons de taille petite dont un mouvement brutal sur un ρ_k peut avoir un effet significatif sur la statistique. Ils proposent

$$Q^* = N(N+2)\sum_{k=1}^{s}\frac{\rho_k^2}{N-k}. \tag{3.4}$$

Sous l'hypothèse nulle, les statistiques Q et Q^* suivent une loi de Khi-deux $\mathcal{X}^2(s)$ à s degré de liberté. Il n'y a aucune méthode précise pour déterminer le paramètre s. Il doit doit être assez grand pour tenir compte d'éventuelles corrélations d'ordre supérieur mais pas trop grand au risque de diminuer la puissance statistique du test. Notons que les deux statistiques ci-dessus intègrent tous les coefficients de corrélations de 1 à s. C'est mieux qu'une statistique ne s'intéressant qu'à la signification statistique d'un seul coefficient.

Nous avons appliqué le test de Ljung et Box sur des données journalières, hebdomadaires, mensuelles couvrant la période de Mars 1990 à Octobre 2009 pour quatre différents retards $s = [20,\ 25,\ 30,\ 35]$. Les résultats sont fournis dans la table 3.1.1.

Données journalières		Nombre de retards s utilisé			
Mars 1990-Oct 2009		$s = 20$	$s = 25$	$s = 30$	$s = 35$
CAC 40 (4958 obs.)					
	Q^*	71,2336	79,6054	102,1391	123,4095
	Prob	≈ 0	≈ 0	≈ 0	≈ 0
DJIA (4950 Obs.)					
	Q^*	90,7744	101,4757	107,8306	136,7382
	Prob	≈ 0	≈ 0	≈ 0	≈ 0
Nasdaq (4950 obs.)					
	Q^*	90,5818	104,1485	119,3724	137,8932
	Prob	≈ 0	≈ 0	≈ 0	≈ 0
Nikkei (4831 obs.)					
	Q^*	43,1525	63,0218	66,6994	69,4189
	Prob	0,0020	≈ 0	≈ 0	≈ 0

TABLE 3.1 – Tests de corrélation sur des données journalières basés sur la statistique Ljung-Box.

On remarque que l'hypothèse nulle (absence de corrélation) est rejetée pour les quatre indices étudiés. Toutes les p-values sont quasi-nulles sauf pour l'indice Nikkei où on a une p-value de l'ordre de 10^{-3} pour le cas $s = 20$. Ce résultat montre que si la fréquence des données est élevée, des corrélations peuvent exister au niveau des rentabilités. Par contre, plus les fréquences sont basses, plus on tend vers la marche aléatoire. Les tables suivantes illustrent ces propos qui sont dus au théorème central limite.

Données hebdomadaires		Nombre de retards s utilisé			
Mars 1990-Oct 2009		$s = 20$	$s = 25$	$s = 30$	$s = 35$
CAC 40 (1025 obs.)					
	Q^*	45,8170	48,9305	56,3967	61,7900
	Prob	$0,0009$	$0,0029$	$0,0025$	$0,0034$
DJIA (1024 Obs.)					
	Q^*	44,9170	54,2739	68,5674	75,8889
	Prob	$0,0011$	$0,0006$	$0,0001$	$0,0001$
Nasdaq (1024 obs.)					
	Q^*	31,5256	32,2045	45.5761	55.5634
	Prob	$0,0486$	$0,1522$	$0,0341$	0,0150
Nikkei (1019 obs.)					
	Q^*	22,3226	31,8879	49,4733	55,2353
	Prob	0,3233	$0,1613$	$0,0141$	$0,01610$

TABLE 3.2 – Tests de corrélation sur des données hebdomadaires basés sur la statistique Ljung-Box.

Dans le cas des séries hebdomadaires, on voit que l'hypothèse nulle n'a pas été rejetée pour tous les indices. Par exemple pour le cas du Nasdaq avec $s = 25$, du Nikkei pour $s = 20$ et $s = 25$ on n'a pu la rejetée au seuil de signification $\alpha = 5\%$. Ce fait peut être expliqué par l'effet agrégation qui permet de blanchir le phénomène de corrélation. Dans la page suivante, on a représenté les résultats obtenus pour les données mensuelles.

Données mensuelles		Nombre de retards s utilisé			
Mars 1990-Oct 2009		$s = 20$	$s = 25$	$s = 30$	$s = 35$
CAC 40 (236 obs.)					
	Q^*	16,8044	18,2535	20,9950	26,6538
	Prob	$0,6656$	$0,8314$	$0,8880$	$0,8434$
DJIA (236 Obs.)					
	Q^*	12,1596	15,8438	17,9521	18,9648
	Prob	$0,9105$	$0,9194$	$0,9593$	$0,9875$
Nasdaq (236 obs.)					
	Q^*	21,3077	23,4122	27,1182	34,6900
	Prob	$0,3792$	$0,5535$	$0,6171$	0,4830
Nikkei (236 obs.)					
	Q^*	14,9710	17,6399	20,3097	21,8947
	Prob	0,7781	$0,8573$	$0,9082$	$0,9589$

TABLE 3.3 – Tests de corrélation sur des données mensuelles basés sur la statistique Ljung-Box.

Dans la table ci-dessus, Prob représente les p-values associées à la statistique Q^*. On voit que dans tous les quatre marchés, on n'a pas une information suffisante pour rejeter l'hypothèse nulle vu que les p-values sont très grandes. [13] Cette étude est faite pour illustrer un premier point qui rend difficile l'appréhension de la question de l'efficience. L'hypothèse d'absence de corrélation a plus de chance d'être acceptée pour des études dont la fréquence de données est faible (hebdomadaires, mensuelles). Pour les données hautes fréquences (journalières, intra-journalières), l'hypothèse nulle est souvent rejetée.
On vient de voir que les résultats sur l'efficience de marché peuvent changer si on modifie la fréquence des données. Dans ce qui suit, on montre même en fixant une fréquence par exemple journalière, on peut aussi aboutir à des résultats différents si on divise la série de données en deux. Nous illustrons nos propos cette fois en se basant sur les tests de ratio de Lo et MacKinlay

13. La taille des données qui s'est réduite considérablement a permis d'agrandir un peu aussi les p-values observées. Cependant, même si on augmente l'intervalle d'étude, l'hypothèse nulle n'est généralement pas rejetée due à l'effet agrégation.

[80] et sur les tests de changement de signe.

3.1.2 Tests sur le ratio de variance.

Le test est basé sur l'idée que dans le cas d'une absence de corrélation, la variance des rentabilités sur q unités de temps $(r_t - r_{t-q})$ est proportionnelle à celle d'une unité et le facteur de proportionnalité est donné par q c'est-à-dire $Var(r_t - r_{t-q}) = q\ Var(r_t - r_{t-1})$. En d'autres termes, la variance de la rentabilité est une fonction linéaire de l'intervalle de temps utilisée pour la calculer. On définit la variable VR par le ratio suivant

$$VR(q) = \frac{\frac{1}{q}Var(r_t - r_{t-q})}{Var(r_t - r_{t-1})} = \frac{\sigma(q)^2}{\sigma(1)^2}. \tag{3.5}$$

Sous l'hypothèse nulle d'absence de corrélation, cette variable est égale à 1. Dans le cas contraire la variable VR donnera par exemple une valeur inférieure (supérieure) strictement à 1 s'il existe des corrélations négatives (positives). Les paramètres $\sigma(q)$, $\sigma(1)$, étant inconnus, ils sont estimés en utilisant une historique de taille disons nq, $n \in \mathbb{N}$. Leurs valeurs sont données par :

$$\hat{\mu} = \frac{1}{nq}\sum_{k=1}^{nq}(r_k - r_{k-1}) = \frac{1}{nq}\ (r_{nq} - r_0)$$

$$\hat{\sigma}(1)^2 = \frac{1}{nq-1}\sum_{k=1}^{nq}(r_k - r_{k-1} - \hat{\mu})^2$$

$$\hat{\sigma}(q)^2 = \frac{1}{h}\sum_{k=q}^{nq}(r_k - r_{k-q} - q\hat{\mu})^2, \quad h = q(nq+1-q)(1-\frac{q}{nq}).$$

On note $\hat{VR}(q)$ la valeur estimée de $VR(q)$ basée sur les estimations non biaisées de $\sigma(1)^2$ et de $\sigma(q)^2$. Dans l'hypothèse de variance constante sur les incréments, Lo et MacKinlay [80] ont proposé la statistique Z suivante qui suit approximativement une loi normale centrée réduite

$$Z = \frac{\hat{VR}(q) - 1}{\sqrt{\phi(q)}}, \quad avec \ \ \phi(q) = \frac{2(2q-1)(q-1)}{3nq^2}.$$

Pour tenir en compte les effets d'hétéroscédasticité au niveau des variances, les auteurs ont proposé la statistique modifiée suivante

$$Z^* = \frac{\hat{VR}(q) - 1}{\sqrt{\phi(q)^*}}, \quad avec \quad \phi(q)^* = \sum_{j=1}^{q-1} \left[\frac{2(q-j)}{q}\right]^2 \hat{\delta}(j)$$

où

$$\hat{\delta}(j) = \frac{\sum_{k=j+1}^{nq}(r_k - r_{k-1} - \hat{\mu})^2(r_{k-j} - r_{k-j-1} - \hat{\mu})^2}{\sum_{k=1}^{nq}(r_k - r_{k-1} - \hat{\mu})^2}.$$

On a appliqué ce test en utilisant les données journalières du CAC 40, du Dow Jones, du Nasdaq et du Nikkei que l'on a subdivisé en deux séries de taille identique ou presque identique (pour les données impaires). L'idée est de voir pour un indice donné si on obtient les mêmes résultats pour les deux sous séries associées. Les résultats des deux sous-séries sont donnés dans les deux pages suivantes.

Données journalières	Nombre q de jours utilisés pour faire l'agrégation				
Série 1(Début Mar 90- Mi Dec 99)		$q = 3$	$q = 5$	$q = 10$	$q = 20$
CAC 40 (2479 obs.)					
	Z	1,9179	1,0016	-0,2263	0,1376
	Prob	0,0551	0,3165	0,8210	0,8906
	Z^*	1,4635	0,7673	-0,1763	0,1100
	Prob	0,1433	0,4429	0,8601	0,9124
DJIA (2475 Obs.)					
	Z	1,0879	-0,3170	-1,4636	-1,0639
	Prob	0,2766	0,7512	0,1433	0,2874
	Z^*	0,6926	-0,2108	-1,0224	-0,7806
	Prob	0,4886	0,8331	0,3066	0,4350
Nasdaq (2475 obs.)					
	Z	4,8179	4,1266	2,5842	2,5141
	Prob	≈ 0	≈ 0	0,0098	0,0119
	Z^*	2,9710	2,5827	1,6745	1,7159
	Prob	0,0030	0,0098	0,0940	0,0862
Nikkei (2416 obs.)					
	Z	-2,4014	-2,3509	-2,0047	-0,8710
	Prob	0,0163	0,0187	0,0450	0,3837
	Z^*	-1,7962	-1,7645	-1,5164	-0,6789
	Prob	0,0725	0,0776	0,1294	0,4972

TABLE 3.4 – Tests de corrélation sur des données journalières basés sur les statistiques Z et Z^*.

Comme précédemment, Prob représente les p-values associées aux statistiques Z et Z^* tenant compte et pas de l'hétéroscédasticité.

Données journalières	Nombre q de jours utilisés pour faire l'agrégation.				
Série 2(Mi Dec99- Mi Oct 09)		$q = 3$	$q = 5$	$q = 10$	$q = 20$
CAC 40 (2479 obs.)					
	Z	-3,0883	-3,5772	-3,7383	-2,7712
	Prob	0,0020	0,0003	0,0002	0,0056
	Z^*	-1,9627	-2,2233	-2,2872	-1,7010
	Prob	0,0497	0,0262	0,0222	0,0889
DJIA (2475 Obs.)					
	Z	-5,1973	-3,7461	-3,2529	-2,3030
	Prob	≈ 0	0,0002	0,0011	0,0213
	Z^*	-3,0220	-2,0863	-1,7588	-1,2345
	Prob	0,0025	0,0369	0,0786	0,2170
Nasdaq (2475 obs.)					
	Z	-3,1885	-2,8307	-2,5571	-1,3614
	Prob	0,0014	0,0046	0,0106	0,1734
	Z^*	-2,0873	-1,8308	-1,6456	-0,8805
	Prob	0,0369	0,0671	0,0998	0,3786
Nikkei (2415 obs.)					
	Z	-1,8570	-2,2141	-2,2930	-1,4412
	Prob	0,0633	0,0268	0,0218	0,1495
	Z^*	-0,9939	-1,1358	-1,1832	-0,7660
	Prob	0,3203	0,2560	0,2367	0,4437

TABLE 3.5 – Tests de corrélation sur des données journalières basés sur les statistiques Z et Z^*.

Si on considère un seuil de signification $\alpha = 5\%$, on note deux choses pour les deux tables ci-dessus. Concernant les indices Nasdaq et Nikkei, l'ordre des résultats est conservé à l'exception d'un ou de deux cas au maximum. Le résultat est conforme à ce que l'on s'attendait. Par contre, pour les deux autres indices restants (CAC 40 et Dow Jones) les résultats obtenus pour les deux sous- séries sont presque inversés à l'exception aussi d'un ou de deux cas au maximum. Par exemple pour le CAC 40, l'exception apparaît pour la statistique Z^* avec $q = 20$. Concernant le Dow Jones, il existe deux exceptions que l'on retrouve dans la statistique Z^* pour $q = 15$ et $q = 20$.
En d'autres termes, les résultats d'efficience peuvent changer sur des intervalles d'études différents même si on utilise la statistique Z^* tenant compte de l'hétérogénéité. On continue cette investigation en regardant cette fois un test non paramétrique qui peut sembler à priori plus robuste.

3.1.3 Tests de changements de signe (*Run tests*).

Si on suppose que les rentabilités des actifs suivent une marche aléatoire alors il y aurait autant de rentabilités positives que négatives. C'est l'idée des tests de changements de signe. Son intérêt est qu'il est un test non paramétrique. Il n'y a aucune hypothèse qui est faite sur les résidus (ϵ_t). Par exemple, le fait que les rentabilités des séries financières aient tendance à avoir des variances non constantes ou à suivre une distribution non normale importent peu ici.
Une rentabilité (ou variation des prix) à l'instant t est soit positive soit négative. On dénote par B_t, la variable binomiale prenant '+', '-' si la rentabilité à l'instant t est positive, négative respectivement. Si on suppose que les rentabilités des séries financières sont indépendantes et que le nombre total de données est T, alors on a simplement une loi binomiale $Z \sim \mathcal{B}(T, p)$ comme une somme de variables indépendantes B_t de type Bernoulli. Sous l'hypothèse nulle d'une marche aléatoire, la fréquence d'apparition des '+' et '-' doivent être identiques entrainant que $p = 1/2$.
En résumé, la statistique pour le test des signes est basée sur la statistique $Z \sim \mathcal{B}(T; 0,5)$ où T est le nombre de données disponibles sur les rentabilités. Si ce nombre T est grand comme c'est le cas dans notre cas, une statistique de type $z - test$ ou de type Khi-deux peut être utilisée de la forme

$$X = \frac{Z - p \cdot T}{\sqrt{p \cdot (1-p) \cdot T}}, \quad K = X^2 = \frac{(Z - p \cdot T)^2}{p \cdot (1-p) \cdot T}$$

où sous l'hypothèse nulle $p = 1/2$.

On a appliqué les tests de signe aux deux sous séries pour chaque indice considéré en utilisant la statistique noramle X définie ci-dessus. Les résultats sont les suivants :

Données journalières		Indices			
Série 1(Mars 90- Mi Dec 99)		$CAC40$	*Dow Jones*	*Nasdaq*	*Nikkei*
	X	-1,1186	0,1767	-6,9448	2,2674
	Prob	0,2633	0,8597	≈ 0	0,0233
Série 2(Mi Dec99- Mi Oct 09)					
	X	2,6762	3,2111	0,3902	1,5271
	Prob	0,0074	0,0013	0,6964	0,1267

TABLE 3.6 – Tests de corrélation sur des données journalières basés sur les signes.

Nous remarquons avec le test supplémentaire des signes que les tous les résultats obtenus dans la première sous-série se sont inversés dans la deuxième. Dans le cas des indices CAC 40 et Dow Jones, l'hypothèse de marche aléatoire est acceptable dans la première sous-série tandis qu'elle ne l'est pas dans l'autre moitié des données restantes. Ces résultats sont en conformité aussi avec ceux obtenus avec le ratio de variance donnés dans les tables 3.4 et 3.5. Pour les indices Nasdaq et Nikkei, un phénomène inverse apparaît. L'hypothèse nulle est d'abord rejetée au seuil de signification 5% pour la série de données allant de Mars 90 à Mi Decembre puis non rejetée pour l'autre moitié de la série restante.

Nous avons voulu montrer dans cette première section comment il est facile d'inverser les conclusions d'une étude en modifiant soit la fréquence des données (journalières, mensuelles) soit l'intervalle d'études (segmentation en deux parties égales). Dans la deuxième section, on revient sur les tests d'efficience qui proposent une stratégie pour exploiter la corrélation statistique qui peut exister dans les séries financières. L'objectif sera aussi de montrer comment on peut changer facilement de conclusions en faisant de simples transformations comme précédemment.

3.2 Tests basés sur une sélection de portefeuilles.

De nos jours, plusieurs tests ont été faits, basés sur une sélection de portefeuilles. La méthodologie est la suivante. On se donne un critère qui puisse permettre de classer des actifs en portefeuilles généralement au nombre de dix. On regarde ensuite le comportement des portefeuilles extrêmes en termes de rentabilités. Souvent, on détient le premier portefeuille et on vend le dernier ou on fait l'opération inverse. Cette méthodologie a été utilisée pour illustrer qu'un modèle à un seul facteur (portefeuille de marché) n'était pas adéquat pour expliquer la rentabilité des titres. Par

exemple, il a été noté que des portefeuilles avec un même béta avaient des rentabilités différentes selon leur capitalisation, selon leur ratio valeur comptable sur valeur de marché (B/M) ou selon leur ratio prix sur bénéfice par action (PER) entre autres. Dans ces exemples, le critère fourni est lié à une caractéristique de l'entreprise. On peut choisir aussi un critère lié à la performance récente des titres. Par exemple, le premier portefeuille ne va contenir que les actifs les plus performants tandis que le dernier les moins performants. Nous appliquons ici cette technique seulement au cas du marché français [14] en utilisant la méthodologie de Jegadeesh et Titman [59] pour voir s'il y a un effet momentum.

Pour faire cette étude, on considère souvent des centaines voire des milliers de titres que l'on regroupe en une dizaine de portefeuilles sur un intervalle d'études se mesurant en décennies. Ici, on a considéré que 30 actifs issus du CAC 40. Les données vont du 01 Janvier 2003 au 16 0ctobre 2009. Si on considérait une période plus longue, on aura un nombre encore plus petit d'actifs qui se soient maintenus dans l'indice. [15] La méthodologie utilisée pour tester l'effet momentum est la suivante. On considère que des données journalières au lieu d'une fréquence mensuelle vu le nombre assez faible de données que l'on dispose. Pour le test, on utilise trois types de données. Le premier est le taux sans risque représenté par le taux euribor 1 mois français tiré sur **Ecowin**. Les deux autres à savoir les rentabilités des différentes actions et celle de l'indice CAC 40 sont tirées dans le site finance de yahoo. On note $R_{i,t}$ la rentabilité de l'actif i à la date t. Ici, i prend les valeurs $1, \cdots, 30$ et T est égal à 1764. Par contre, pour l'indice CAC 40 et le taux euribor les données sont au nombre de 1739. On a dû compléter les données manquantes qui se trouvent généralement au niveau des fêtes de Noël et de la fête des travailleurs (01 Mai) en faisant les moyennes des données adjacentes. Dès lors, on note $R_{f,t}$, $R_{m,t}$ respectivement le taux euribor et la rentabilité de l'indice CAC 40 à la date $t = 1, \cdots, 1764$. Le taux euribor a été converti du taux mensuel au taux journalier en utilisant la transformation $R_f \leftarrow (1 + tx1m)^{\frac{1}{20}} - 1$ où $tx1m$ est le taux euribor 1 mois.
La deuxième étape consiste à définir une période de comparaison d'une longueur de m jours et une période d'évaluation de n jours. Ici, on a considéré des périodes de longueurs égales ($m = n$) qui ne s'entrelacent pas. Précisément, le nombre de données T est divisé en un nombre pair k d'intervalles contenant chacun n jours. Les intervalles ayant un nombre **impair** sont utilisés pour créer 6 portefeuilles [16] selon le critère de la performance. Le premier portefeuille $P1$ va contenir toujours les 5 actions les plus performantes et le dernier portefeuille $P6$ les 5 actions les moins performantes. Ensuite, on évalue la performance de ces portefeuilles extrêmes à savoir $P1$ et $P6$ seulement durant le prochain intervalle de numéro **pair**. Ci-dessous, on donne un schéma qui résume la méthodologie utilisée avec k supposé pair.

14. L'étude prenant plus de temps et plus d'espace que dans la section précédente, on a préféré se limiter au seul marché français.

15. Les actions du CAC utilisées par ordre alphabétique de leur Mnémo sont Accor (AC), Air Liquide (AI), Alcatel-Lucent(ALU), Danone(BN), BNP Paribas(BNP), Carrefour (CA), Capgemini(CAP), Axa (CS), Vinci (DG), Essilor INTL (EI), Total(FP), France Télécom (FTE), Société Générale (GLE), GDF Suez(GSZ), Lafarge (LG), LVMH(MC), Michelin(ML), Lagardère(MMB), L'Oréal(OR), PPR (PP), Pernod Ricard (RI), Renault(RNO), Sanofi-Aventis(SAN), Saint-Gobain(SGO), Schneider Electric(SU), Technip(TEC), Peugeot (UG), Unibail-Rodamco(UL), Vivendi (VIV), Vallourec(VK).

16. Chaque portefeuille contient donc 5 actifs. On a fait le même traitement en considérant 10 portefeuilles contenant chacun 3 actifs, les résultats restent les mêmes.

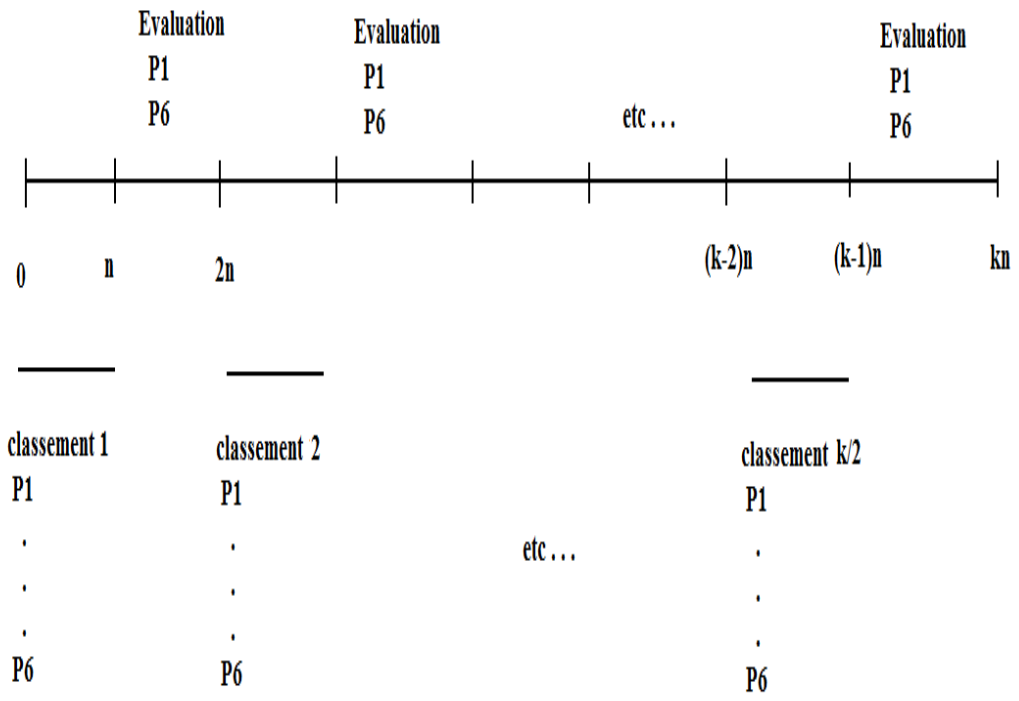

FIGURE 3.1 – Illustration de la méthodologie utilisée pour tester l'effet Momentum.

Le critère de classement est donné dans chaque intervalle j impair par $r_{i,j}$ où :

$$r_{i,j} = \frac{1}{n} \sum_{t=(j-1)n\ +1}^{jn} U_{i,t}, \quad i\ = 1, \cdots, 30, \quad avec\ \ U_{i,t}\ =\ R_{i,t}\ -\ R_{f,t}.$$

Plutôt que de travailler directement sur les rentabilités des titres $R_{i,t}$, on a préféré prendre la prime de risque donnée par $U_{i,t}$. On note $r_{gan,t}$, $r_{per,t}$ la moyenne des primes de risque respectivement du portefeuille des gagnants (P1) et des perdants (P6) à la date $t = nj+1, nj+2, \cdots, n(j+1)$ avec $j = 1, \cdots, floor(\frac{T}{n})$.[17] L'effet momentum peut alors être testé par une simple régression de la forme

$$r_{gan,t} - r_{per,t} = \alpha_1 + \epsilon_t, \tag{3.6}$$

où (ϵ_t) représente les termes d'erreurs.

On s'attend à ce que le coefficient α_1 soit significativement positif pour un effet momentum, négatif pour un effet reversal et nul dans le cas de l'efficience de marché. Dans le tableau ci-dessous,

17. floor (x) est la plus grande valeur entière inférieure ou égale à x.

on a pris $n = 5$ jours correspondant à une longueur d'une semaine pour classer et aussi évaluer la performance des portefeuilles.

$Ordinary\ Least - squares\ Estimates$			
$Dependent\ Variable$		$r_{gan,j} - r_{per,j}$	
$R - squared$		$0,0000$	
$Rbar - squared$		$0,0000$	
σ^2		$0,0002$	
$Durbin - Watson$		$2,0561$	
$Nobs, Nvars$		$880,1$	
$Variable$	$Coefficient$	$t - statistic$	$t - probability$
$\hat{\alpha_1}$	$-0,000434$	$-0,891362$	$0,372979$

On remarque que le coefficient α_1 est significativement nul au seuil de signification 5% ou 10%. Il n'existe donc pas d'excès de rentabilités entre les portefeuilles gagnants et perdants pour un critère de classement sur une semaine et commençant le 01 Janvier 2003.

Dans une seconde étape, on considère les mêmes données en faisant simplement une translation d'un jour. En d'autres termes, au lieu de commencer au 01 Janvier, on a pris le 02 Janvier 2003 puis un décalage d'une semaine (5 jours ouvrables) est toujours pris pour classer et évaluer les performances des actifs. Les résultats obtenus sont représentés dans le tableau ci-dessous.

$Ordinary\ Least - squares\ Estimates$			
$Dependent\ Variable$		$r_{gan,j} - r_{per,j}$	
$R - squared$		$0,0000$	
$Rbar - squared$		$0,0000$	
σ^2		$0,0003$	
$Durbin - Watson$		$2,0819$	
$Nobs,\ Nvars$		$880,1$	
$Variable$	$Coefficient$	$t - statistic$	$t - probability$
$\hat{\alpha_1}$	$0,003318$	$5,552410$	$0,000000$

On remarque que le coefficient α_1 est significativement non nul et positif au seuil de signification 10% et 5%. En d'autres termes, la stratégie qui consiste à acheter les actifs les plus performants et à vendre les moins performants durant la semaine précédente donne un excès de rentabilité significatif. On a donc un phénomène de momentum à court terme similaire à celui illustré par Jegadeesh et Titman [59].

Pour voir la robustesse de la transformation faite ci-dessus (décalage de jours), on a introduit la prime de risque de l'indice CAC 40 dont l'absence dans la régression (3.6) pouvait être à l'origine de cette différence nette. [18]

$$r_{gan,t} - r_{per,t} = \alpha_2 \ + \ \beta(R_{m,t} - R_{f,t}) \ + \ \epsilon_t. \tag{3.7}$$

Les résultats de la régression (3.7) sont donnés ci-dessous pour n= 5 jours et commençant au 01 Janvier 2003.

$Ordinary\ Least - squares\ Estimates$			
$Dependent\ Variable$		$r_{gan,j} - r_{per,j}$	
$R - squared$		$0,0000$	
$Rbar - squared$		$-0,0011$	
σ^2		$0,0002$	
$Durbin - Watson$		$2,0563$	
$Nobs,\ Nvars$		$880,2$	
$Variable$	$Coefficient$	$t - statistic$	$t - probability$
$\hat{\alpha_2}$	$-0,000430$	$-0,879360$	$0,379447$
$\hat{\beta}$	$0,003651$	$0,112369$	$0,910557$

On obtient un résultat similaire au cas où on avait qu'un seul facteur. Le paramètre β et α_2 ne sont pas significatifs. En faisant le décalage d'un jour, on arrive aussi à la même conclusion que dans le cas précédent à savoir que le coefficient α_2 devient significativement positif voir le tableau suivant.

18. Ce point a motivé le choix de $U_{i,t}$ plutôt que $R_{i,t}$.

$Ordinary\ Least-squares\ Estimates$			
$Dependent\ Variable$		$r_{gan,j}-r_{per,j}$	
$R-squared$		$0,0002$	
$Rbar-squared$		$-0,0010$	
σ^2		$0,0003$	
$Durbin-Watson$		$2,0820$	
$Nobs,\ Nvars$		$880,2$	
$Variable$	$Coefficient$	$t-statistic$	$t-probability$
$\hat{\alpha}_2$	$0,003301$	$5,503863$	$0,000000$
$\hat{\beta}$	$-0,014718$	$-0,369398$	$0,711920.$

On a refait la même technique appliquée à $n = 10$ (deux semaines) et $n = 20$ (1 mois). Dans les deux cas, on arrive à obtenir un effet momentum en faisant une translation des données de la série initiale. Par contre, si on commence au 01 Janvier ou sur certaines dates, cet effet momemtum disparaît comme dans le cas $n = 5$. Les conclusions restent identiques pour les deux types de régression (3.6) et (3.7). Les résultats ne sont donc pas reportées car similaires à ceux illustrés pour le cas hebdomadaire.

Dans cette partie empirique, on a voulu montrer la difficulté liée aux tests d'efficience. De simples transformations permettent d'un passer d'un camp à l'autre. On a illustré entre autres que la fréquence des données, l'intervalle d'études et les translations de données pour une fréquence fixée peuvent changer une conclusion d'étude.
Un autre aspect qui n'est pas investi ici (tests au sens semi-fort) est les modèles utilisant les techniques fondamentales pour prédire les rentabilités futures des titres (spread de taux, indices, dividendes, ratio B/M, PER, etc.). La prévisibilité constatée dans ces modèles peut être fallacieuse et être due simplement à de la chance. Par exemple, on peut voir facilement un phénomène de tendance dans une trajectoire d'un mouvement brownien si on en fait plusieurs simulations. Si cette trajectoire était celle de notre économie, on aurait à rejeter l'hypothèse d'efficience tout en ayant tort. Ce phénomène est connu sous le label de *data snooping* et peut expliquer les résultats obtenus ci-dessus en subdivisant l'intervalle d'études en deux. Cooper et Gulen [25] illustrent ce phénomène en montrant que même des variables instrumentales prises aléatoirement peuvent générer de la prévisibilité. Les auteurs montrent aussi l'instabilité de ces modèles basés sur l'approche fondamentale où de petits changements (ajout ou suppression d'une variable) peuvent faire changer les conclusions d'une étude. Le phénomène de data snooping s'applique aussi pour les modèles utilisant les techniques chartistes, voir par exemple Sullivan, Timmermann, White [107] pour d'amples informations.

3.3 Discussion.

Dans la section précédente, on a rappelé quelques tests liés à la marche aléatoire des rentabilités des actifs. Ceux-ci ne tenaient pas en compte le lien qui pouvait exister entre les rentabilités

des actifs et les fondamentaux de l'économie. Les rentabilités d'aujourd'hui étaient expliquées simplement par celles passés plus un bruit. En d'autres termes, le noyau d'évaluation ξ_{t+1} qui a été utilisé dans l'équation (1.10) n'apparaît pas dans la formation des prix d'équilibre. Ce fait n'est pas réaliste vu que les préférences des agents ainsi que les facteurs macro-économiques sont réduit à un bruit dans la formation des prix d'équilibre. C'est pourquoi, les chercheurs ont investi des tests d'efficience de marché faisant intégrer le noyau et dans ce cas la question de l'hypothèse jointe fait surface. On a vu dans le premier chapitre qu'il existait une relation entre les primes de risque de chaque actif risqué avec celle du noyau d'évaluation de la forme :

$$E_t(r_{j(t+1)} - r_t) = \frac{cov(r_{j(t+1)}, R_{L_{t+1}}|\mathcal{F}_t)}{V(R_{L_{t+1}}|\mathcal{F}_t)}\Big(E_t(R_{L_{t+1}}) - r_t)\Big) = \beta_t^{j,L} E_t(R_{L_{t+1}} - r_t) \tag{3.8}$$

où r_t est le taux déterministe sans risque, $\beta_t^{j,L}$ le facteur béta mesurant la corrélation entre le titre j et le portefeuille de référence L défini par $L_{t+1} = \xi_{t+1}(1 + r_t)$ avec ξ_{t+1} représentant le noyau d'évaluation. Pour des raisons empiriques, on doit spécifier une expression pour ce noyau en se basant sur la théorie de l'efficience de marché. On considère donc l'approche de la finance néo-classique fondée sur la maximisation de l'utilité espérée voir équation (1.1). Ici, l'utilité des agents provient essentiellement de leur consommation. A chaque instant, leur richesse est divisée en deux parties. La première est utilisée pour la consommation et la seconde pour l'investissement. On notera e_t^k, la dotation de l'agent k à l'instant t provenant de son salaire de travail par exemple. Pour faire simple, on considère un modèle à deux périodes avec un seul actif disponible en termes d'investissement dont le prix est noté S_t à l'instant t. On suppose que cet actif s'échange sans aucune contrainte. L'investisseur k résout le problème suivant :

$$\begin{aligned} &\underset{(c_t^k, c_{t+1}^k)}{Max} && E_{i,t}\Big[\ u(c_t^k)\ +\ \beta u(c_{t+1}^k)\ \Big]\ . \\ s.c: \quad & n_t^k S_t = e_t^k - c_t^k, && c_{t+1}^k = e_{t+1}^k + n_t^k S_{t+1}\ . \end{aligned} \tag{3.9}$$

La fonction d'utilité est séparable et de la forme $U(c_t^k, c_{t+1}^k) = u(c_t^k) + \beta u(c_{t+1}^k)$. Elle se généralise facilement dans le cas d'un horizon de temps infini. La première contrainte matérialise que la dotation initiale est répartie totalement entre consommation et investissement. Vu que le modèle est à deux périodes, à l'instant suivant, l'investisseur consomme toute sa richesse provenant de son investissement et de son salaire. La condition de premier ordre donne que

$$S_t u'(c_t^k) = E_t[\beta u'(c_{t+1}^k) S_{t+1}]\ . \tag{3.10}$$

L'interprétation de cette équation est qu'à l'équilibre, la perte d'utilité à l'instant t occasionnée par la détention d'un actif supplémentaire devrait être égale au gain d'utilité espéré à l'instant $t + 1$ actualisé par β pour avoir des termes comparables.

Cette équation est vraie dans un cadre général et ne dépend pas de la formulation du modèle décrite ci-dessus. En effet la perte d'utilité à l'instant t due à une détention supplémentaire h d'actifs avec h très petit est donnée par

$$u(c_t^k - hS_t) - u(c_t^k) = -hu'(c_t^k)S_t + o(h)$$

D'autre part, la détention de cette quantité supplémentaire d'actif augmentera notre utilité espérée à l'instant $t+1$ de

$$E_t\Big(u(c_{t+1}^k + hS_{t+1}) - u(c_{t+1}^k)\Big) = hE_t\Big(u'(c_{t+1}^k)S_{t+1})\Big) + o(h)$$

Dans un marché où les agents sont rationnels et où h peut être aussi petit que l'on veut, la somme de l'utilité perdue avec celle que l'on espère gagner après actualisation doit donner zéro. Ce qui redonne l'équation déjà obtenue (3.10) en notant que $o(h) \longrightarrow 0,\ si\ h \longrightarrow 0$. Le résultat est donc indépendant de l'horizon de temps fini supposé ici ou des dotations des agents introduites. Pour des études empiriques, on utilise généralement des variables agrégées. L'hypothèse la plus simple est de supposer que tous les agents sont identiques de sorte qu'on peut supprimer l'indice k. Dans ce cas, la valeur du prix S_t est donnée par

$$S_t = E_t\left[\frac{\beta u'(c_{t+1})}{u'(c_t)}S_{t+1}\right] \tag{3.11}$$

où c_t représente la consommation de l'agent représentatif. L'hypothèse d'homogénéité des agents peut à premier abord sembler restrictif. Il n'en est pas réellement vu que dans certaines situations un modèle d'équilibre avec hétérogénéité peut se ramener à un cadre d'homogénéité des agents au sens où les prix d'équilibre et la consommation agrégée restent les mêmes dans les deux types de modèle. Cette problématique est étudiée entre autres par Jouini et Napp [62].
En comparant l'équation (3.11) avec celle de (1.9), on peut en déduire une forme du noyau d'évaluation qui est donnée par

$$\xi_{t+1} = \frac{\beta u'(c_{t+1})}{u'(c_t)} \tag{3.12}$$

Il reste à spécifier une forme d'utilité pour l'agent représentatif. Souvent la fonction est choisie dans la classe HARA (Hyperbolic Absolute Risk Aversion) qui a la propriété d'être stable en termes d'agrégation. Un cas particulier très utilisé est la fonction d'utilité puissance définie par

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

où γ représente le coefficient d'aversion de l'agent. Dans ce cas, le noyau est donné par

$$\xi_{t+1} = \beta(\frac{c_{t+1}}{c_t})^{-\gamma}.$$

Connaissant ξ_{t+1}, on peut tester l'égalité statistique (3.8). Supposons qu'elle ne soit pas satisfaite ou si elle l'est, c'est pour des valeurs du coefficient d'aversion au risque γ non raisonnables. Dans ce cas, on peut supposer la présence d'irrationalité chez certains agents faisant que l'équation (3.12) n'est plus satisfaite. Il se peut aussi que celle-ci soit réellement satisfaite mais que c'est la fonction d'utilité puissance qui n'est pas commode pour modéliser les préférences des agents. Par exemple, on peut essayer les autres fonctions de type HARA pour voir si le puzzle va persister. S'il persiste, on revoit l'approche même d'utilité additivement séparable qui peut ne pas être convenable. La consommation en l'instant t peut avoir une influence sur l'utilité marginale en $t+1$. Dans ce cas, mieux vaut prendre des formes plus générales de fonction d'utilité non séparables.

Si le problème persiste toujours, on peut supposer que le problème se situe au niveau des données utilisées et pas au niveau des préférences des agents. Dans ce cas, il faudra ajouter d'autres facteurs macro-économiques ou indices qui peuvent influer l'utilité marginale des agents. On passe alors d'un facteur c_t à plusieurs facteurs $(c_t, f_t^1, \cdots, f_t^n)$. Pour plus de détails sur le noyau et son application en finance, on peut voir le livre de Cochrane [23] dont l'exposé ci-dessus est basé.

On voit donc la difficulté à tirer une conclusion si on arrive à trouver des anomalies dans une étude empirique. L'explication peut venir d'une irrationalité des agents comme simplement d'une mauvaise spécification de modèle d'équilibre. C'est le problème de l'hypothèses jointe.
Au-delà de ce fait, on a vu aussi que le traitement des données était sujet à des controverses (périodes d'étude, fréquence des données, etc.). Nous proposons un modèle qui annihile l'effet de ces deux points afin d'éclaircir le débat qui existe entre les deux paradigmes. Pour ce faire, on considère un modèle théorique pour éviter les problèmes méthodologiques et qui simule une violation de la loi du prix unique pour éviter l'hypothèse jointe. Remarquons que dans ce cas, on ne peut dire que l'anomalie est due à une mauvaise spécification de modèle. Même si tel est le cas, il s'appliquerait aux deux actifs donnant toujours des prix identiques ne reflétant pas les fondamentaux. La seule explication pour la théorie de l'efficience est de dire que les anomalies arrivent par chance. Les sous et surévaluations se répartissent de façon équitable de sorte qu'en moyenne les prix reflètent leurs valeurs fondamentales. Pour les partisans de la finance comportementale, la violation sera due à une présence d'agents irrationnels et le fait que les arbitrages sont risqués.
Notre objectif est de construire un modèle de sorte que les deux arguments soient valides. D'autre part, le modèle trouvé peut être utilisé en termes de modélisation de prix d'actifs vu qu'il permet de tenir compte à la fois des arguments de la finance néo-classique et comportementale.

Pour construire un tel modèle de marché, on suppose qu'il existe deux actifs risqués A et B de mêmes caractéristiques (substituts) et un actif sans risque C donnant un taux d'intérêt constant

r. Le marché est segmenté au sens où il existe des investisseurs de type A et des investisseurs de type B. Les investisseurs de type A (B) n'investissent que sur l'actif risqué A (B) et sur l'actif sans risque C. Les actifs A et B peuvent être vus comme une même action cotée dans deux places financières. Il existe des arbitragistes qui, eux, peuvent investir à la fois dans les deux secteurs où A et B sont échangés. On suppose que les investisseurs de type A sont tous rationnels. On en déduit que le prix de l'actif A reflète toujours sa valeur fondamentale qui est supposé suivre un modèle à volatilité stochastique pour garder un cadre de généralité.
Les investisseurs de type B sont composés quant à eux d'agents rationnels et irrationnels de sorte que le prix de l'actif B peut dévier de sa valeur fondamentale. Dans cette étude, le cadre est supposé continu pour faciliter l'étude de la stratégie optimale d'un arbitragiste vu la disponibilité du calcul stochastique. Ici, la dynamique de la valeur fondamentale de A ne joue qu'un rôle secondaire vu que l'actif B aura également une dynamique similaire. Mais on ajoutera une autre partie dans la dynamique de B pour tenir compte de l'effet des agents irrationnels qui sont présents dans le marché.
On note respectivement $X = (X_t)_{t\in[0,T]}$ et $Y = (Y_t)_{t\in[0,T]}$ les processus de prix pour les actifs A et B où $X_0 = Y_0 = C > 0$ et T représente la date d'échéance pour les investisseurs. La valeur fondamentale de l'actif A est à tout instant donnée par :

$$dX_t = (r + \eta V_t)X_t dt + \sqrt{V_t} X_t dB_t^1 \ . \tag{3.13}$$

$$dV_t = (\alpha - \beta V_t)dt + \sigma\sqrt{V_t} dB_t^2 \ . \tag{3.14}$$

Cette dynamique pour la volatilité a été introduite par Cox, Ingersoll, Ross [27]. Elle permet d'assurer que la volatilité est toujours positive.[19] Les paramètres r, η, α, β, σ sont des constantes positives. β est le coefficient indiquant la vitesse à la quelle la volatilité retourne à sa moyenne à long terme matérialisée par $\frac{\alpha}{\beta}$. B^1, B^2 représentent deux mouvements browniens standards corrélés (coefficient de corrélation ρ) pour faire ressortir l'aspect effet de levier que l'on trouve fréquemment dans les séries financières. On a pris la prime de risque proportionnelle à la volatilité pour spécifier le fait que plus le marché est volatile plus les investisseurs demandent une prime élevée. Pour spécifier la dynamique du processus Y, on se base sur celle de X. En effet, il existe un lien entre les deux actifs A et B vu qu'ils sont similaires. Précisément, ils sont égaux jusqu'à un certain instant aléatoire T_1 avant qu'ils ne diffèrent. De plus, les deux processus de prix doivent avoir une même rentabilité espérée (drift) pour refléter qu'ils sont de mêmes caractéristiques. Formellement, on choisit la dynamique suivante pour Y

$$dY_t = (r + \eta V_t)Y_t dt + \sqrt{V_t} Y_t dB_t^1 + Z_t Y_{t-} dN_t \tag{3.15}$$

où N est un processus de poisson d'intensité λ avec $N_0 = 0$. La dynamique de V_t est donnée par l'équation (3.14). La variable aléatoire Z_t mesure le signe et la taille des sauts. Elle a un support sur $[-b_1, b_1]$ où b_1 est tel que $0 < b_1 < 1$. Cette dernière condition permet d'assurer la positivité du processus Y.
Dans notre modèle, les actifs A et B ayant mêmes caractéristiques, ils réagissent donc de la

19. La variance sera strictement positive si la condition de Feller est satisfaite i.e $2\alpha > \sigma^2$.

même manière face aux nouvelles informations qui arrivent. Ce qui veut dire que l'instant des sauts n'est pas dû à l'apparition d'une information fondamentale comme c'est le cas dans les modèles à sauts mais plutôt à l'apparition de biais dans le marché. C'est pourquoi on doit faire des hypothèses particulières sur la variable Z_t mesurant l'amplitude des sauts.

$$\forall t, \quad Z_t \textit{ est telle que } E(Z_t) = 0. \tag{3.16}$$

$$Support(Z) = [-b_1, b_1] \textit{ avec } \; 0 < b_1 < \epsilon \tag{3.17}$$

où ϵ est un réel petit strictement inférieur à 1 et $Support(Z) = [inf\ Z, sup\ Z]$.

La condition (3.16) contraint la variable Z_t d'être d'espérance nulle pour permettre les actifs A et B d'avoir la même tendance. En effet si $E(Z_t) = \mu \neq 0$, alors on devrait retrancher l'effet tendance qu'auraient occasionné les sauts. Dans ce cas les dynamiques X et Y n'auront pas le même drift et ce qui est contradictoire avec l'hypothèse qu'ils sont similaires.
La condition (3.17) permet d'assurer que l'amplitude des sauts ne peut pas être très grande. En effet comme on l'a souligné tantôt les sauts sont utilisés pour modéliser l'apparition de biais tels que le pessimisme ou l'optimisme des agents vis à vis de l'actif B. Ces biais apparaîtront aux temps d'arrêts (T_k) suivants :

$$\begin{aligned} T_1 &= inf\Big\{t > 0,\ \Delta N_t \neq 0\Big\} \wedge T \\ T_2 &= inf\Big\{t > T_1,\ \Delta N_t \neq 0\Big\} \wedge T. \end{aligned}$$

De façon itérative, on définit pour $k \geq 1$

$$\begin{aligned} T_{2k+1} &= inf\Big\{t > T_{2k},\ \Delta N_t \neq 0\Big\} \ \wedge T \\ T_{2k+2} &= inf\Big\{t > T_{2k+1},\ \Delta N_t \neq 0\Big\} \ \wedge T \end{aligned}$$

L'intervalle $[T_k, T_{k+1}[$ mesure la durée entre l'apparition de deux chocs. La persistance du biais a lieu si pour deux temps d'arrêts successifs (ou plus) la réalisation de la variable Z_t garde un même signe constant. Le fait que cet événement soit non négligeable rend les arbitrages risqués. On a simulé quelques scénarios des processus X et Y pour voir de façon plus claire le problème auquel sont confrontés les arbitragistes.

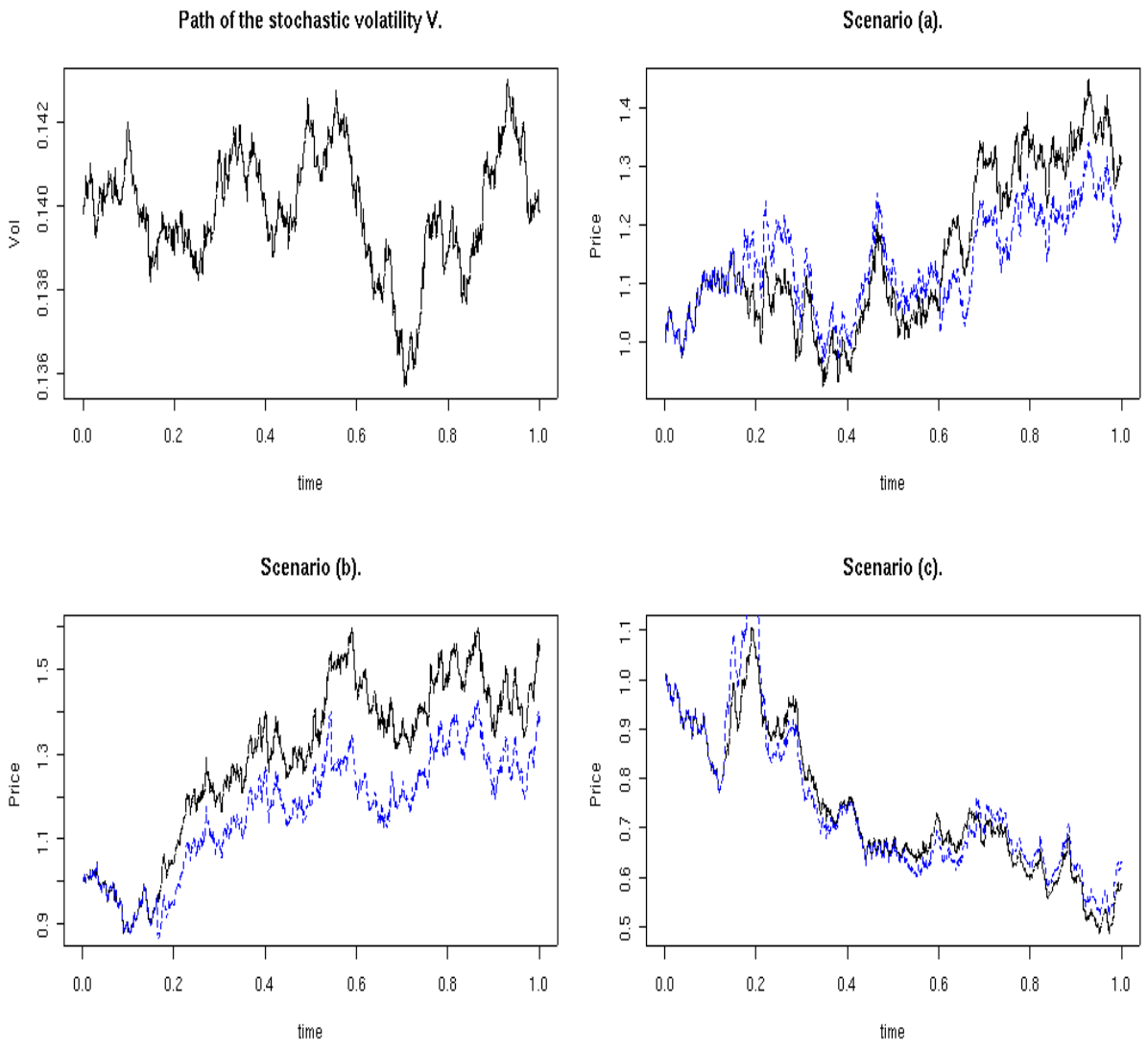

FIGURE 3.2 – Différents scénarios de X et Y pour $Z_t \sim \mathcal{U}[-0.1, 0.1]$.

Dans la figure 3.2, on a choisi pour la variable Z_t une loi uniforme $\mathcal{U}[-0.1, 0.1]$ et pour les paramètres du modèle, on a pris :

$$X_0 = Y_0 = 1,\ V_0 = 0.14,\ r = 0.05,\ \eta = 0.86,\ \alpha = 0.15,\ \beta = 1.1,\ \sigma = 0.02,\ \rho = 0.$$

Le processus X est représenté par un trait plein et le processus Y par des traits espacés. La figure 3.3 est obtenue pour les mêmes données sauf pour la loi uniforme où on a réduit l'amplitude des

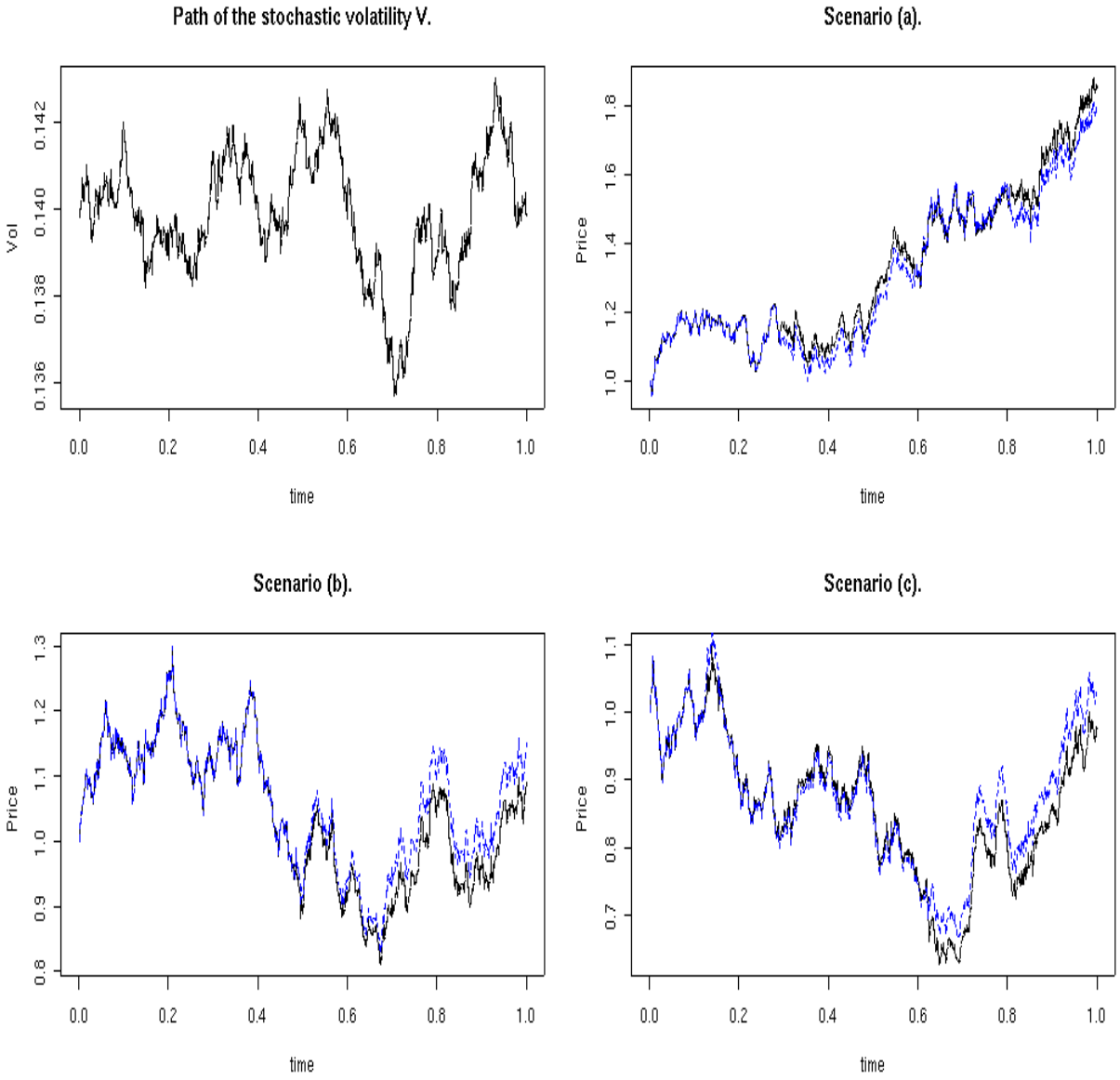

FIGURE 3.3 – Différents scénarios de X et Y pour $Z_t \sim \mathcal{U}[-0.05, 0.05]$.

sauts en choisissant $\mathcal{U}[-0.05, 0.05]$.
Dans les deux figures, l'intensité des sauts a été prise égale à $\lambda = 5$, la maturité à $T = 1$ que l'on a subdivisé en pas de $h = 10^{-3}$ pour la simulation (Méthode d'Euler).
On remarque que sur certains scénarios, la stratégie consistant à jouer contre les arbitragistes entraîne une perte, voir figure 3.2, scénario (b). Sur d'autres, on peut à la fois y gagner et y perdre durant l'intervalle $[0, 1]$. C'est le cas des scénarios (a), (c) toujours de la figure 3.2. On retrouve à peu prés les mêmes types de trajectoires pour la figure 3.3, où on n'a fait que diminuer l'amplitude des sauts. Par exemple, dans le scénario (c) de la figure 3.3, on a un phénomène similaire à un marché efficient. En effet, les deux processus se séparent à l'instant de saut $T_1 = 0,12$ pour se recouper à l'instant $T_2 = 0.15$. Dès fois, il peut s'écouler beaucoup de temps avant que le recollement n'ait lieu. C'est le cas du scénario (a) toujours de la figure 3.3 où les processus se séparent à l'instant $T_1 = 0,29$ pour se retrouver à l'instant $T_2 = 0,59$.
Dans un tel marché décrit ci-dessus, le marché est-il efficient ou inefficient ?
Nous montrons que la réponse à cette question est problématique indépendamment de l'hypothèse jointe ou des problèmes méthodologiques.

Le Marché est efficient.
Pour répondre à cette question, on peut distinguer deux cas.

- **Cas 1 : λ grand, b_1 petit.**

  Si l'intensité des sauts λ est grande et que la taille des sauts b_1 est petite alors le processus Y se comporte comme un bruit du processus X. En effet, un processus avec une intensité élevée de sauts symétriques se comporte comme un mouvement brownien. Si de plus la taille des sauts est petite, on obtient un processus similaire à un bruit. On retrouve les arguments de l'efficience de marché qui stipulent que les anomalies sont dues simplement à de la chance. D'autre part, on n'a aucune information pour anticiper sur le signe du choc (variations imprévisibles). Les agents rationnels contrôlent le marché et les bruiteurs, même s'ils existent, n'ont aucune influence sur le marché vu que leur action est considéré comme un bruit. On peut voir, par exemple, le scénario (c) de la figure 3.2 ou le scénario (a) de la figure 3.3 pour illustration.

- **Cas 2 : λ petit ou b_1 grand.**

  Si l'intensité des sauts est faible ou si la taille des sauts est très grande alors l'argument chance ne tient plus, voir par exemple le scénario (b) de la figure 3.2. Un argument que l'on peut utiliser consiste à dire que les anomalies arrivent de façon symétrique de sorte qu'en moyenne les prix des actifs reflètent leurs valeurs fondamentales, voir Fama [37]. Sous cet angle, le marché demeure aussi efficient.

Pour les partisans de la finance comportementale, le marché est inefficient. La psychologie apparaît déjà à travers la partie saut ajoutée au processus Y et qui permet aux deux actifs A et B d'avoir des prix différents. Remarquons que la dynamique du processus Y est similaire à celle de DeLong et al. [81] rappelée dans le chapitre dédié à la finance comportementale au sens où il y a une composante fondamentale pour les prix (dynamique de X) plus une partie bruit (agents

irrationnels).

La seconde étape consiste à voir si la description faite du marché avec la présence d'imperfections peut réellement exister dans un cadre où les arbitragistes sont présents. La réponse est oui car les agents irrationnels influencent les prix d'une telle manière qu'il est risqué pour les arbitragistes de jouer contre eux (*Limits to arbitrage*). Pour le montrer, on a étudié la ou les stratégie(s) optimale(s) pour un arbitragiste d'exploiter les inefficiences. On suppose que l'arbitragiste dispose d'une somme initiale $x > 0$ qui lui est fournie par des investisseurs. Avec cette somme, il doit mettre une stratégie qui lui permette d'exploiter les mauvaises spécifications de prix qui débutent à partir de la date T_1. Pour ce faire, il va investir sur un portefeuille constitué des actifs risqués A et B dont les prix sont représentés par les processus X et Y. Si on note W_t la valeur de son portefeuille à l'instant t, ζ_t et ϕ_t la proportion en montant du portefeuille investie respectivement dans l'actif A et dans l'actif B alors la dynamique du portefeuille W est donnée par

$$dW_t = \zeta_t W_t \frac{dX_t}{X_t} + \phi_t W_{t-} \frac{dY_t}{Y_{t-}} \ .$$

On considère que les stratégies de l'arbitragiste sont autofinancées en d'autres termes, il n'y a ni apport ni retrait de fonds durant la période d'étude $[0, T]$. En utilisant (3.13), (3.15) et le fait que $\zeta_t = 1 - \phi_t$, la dynamique du portefeuille peut être exprimée en fonction de ϕ_t seulement. Précisément, on a :

$$dW_t = W_t \Big(r + \eta V_t \Big) dt + W_t \sqrt{V_t} dB_t^1 + W_{t-} \phi_t Z_t dN_t \ . \tag{3.18}$$

Les arbitragistes, faisant partie de la catégorie des rationnels, il est supposé que leurs préférences vérifient les axiomes de VNM ou de Savage et peuvent donc être représentées par des fonctions d'utilité. On note $\mathcal{U}$ l'ensemble des fonctions U strictement concaves, croissantes et de classe $\mathcal{C}^2$ notamment que $U' > 0,\ U'' < 0$. On suppose que tout arbitragiste a sa fonction d'utilité dans cet ensemble $\mathcal{U}$. Son objectif sera de maximiser l'utilité terminale de son portefeuille. Si U est sa fonction d'utilité, il ou elle résout le problème suivant :

$$\sup_{\phi} \quad E_P[U(W_T)] \tag{3.19}$$

La proposition suivante montre que le processus d'investissement sur l'actif B dépend de l'amplitude des sauts.

Proposition Supposons que pour tout $t \in [0, T]$, nous ayons

$$0 < E_t \Big[\exp \Big(- \int_t^T \lambda V_s ds \Big) \Big] < 1 \tag{3.20}$$

alors le processus (ϕ_t) est borné et on a $\forall t, \quad -\frac{1}{b_1} \leq \phi_t \leq -\frac{1}{b_1}$.

Preuve Voir Liu et al. (2000).

L'idée est de voir que la condition (3.20) traduit le fait que la probabilité qu'il n'y ait pas de sauts durant l'intervalle $[t, T]$ pour $t \in [0, T[$ est comprise strictement entre 0 et 1. Autrement dit, les événements E_1 = "avoir une trajectoire continue sur [t,T]" et E_2 = "avoir une trajectoire discontinue sur $[t, T]$" pour $t < T$ sont tous non négligeables. Conditionnellement à l'événement E_1, c'est-à-dire la partie saut de l'équation (3.18)est nulle, on a la relation suivante

$$W_T = W_t \exp\left(\int_t^T \left\{\left(r + \eta V_s - \frac{V_s}{2}\right)ds + \sqrt{V_s}dB_s^1\right\}\right)$$

Pour que l'investisseur puisse s'assurer $W_T > 0$, on doit avoir W_t toujours strictement positive pour tout t.

D'autre part, sur l'événement E_2, on a à l'instant du saut $W_t = W_{t-}(1 + \phi_t Z_t)$. Pour que la condition $W_t > 0$ soit satisfaite pour tout t, on doit exiger que

$$1 + \phi_t Z_{inf} > 0, \quad 1 + \phi_t Z_{sup} > 0$$

où Z_{inf} et Z_{sup} représentent la borne inférieure, supérieure respectivement de la variable Z_t. Le résultat de la proposition en découle si on remarque que $Z_{inf} = -b_1$ et $Z_{sup} = b_1$.

Remarquons de cette proposition que l'arbitragiste a un vaste choix de stratégies. En effet, la proposition nous dit qu'il peut détenir à tout instant t une quantité ϕ_t d'actif B parcourant au moins l'intervalle $[-\frac{1}{\epsilon}, \frac{1}{\epsilon}]$ qui est grand vu que ϵ est petit.

La prochaine étape consiste à déterminer la solution optimale associée au problème de l'arbitragiste (3.19). Nous utilisons l'approche Hamilton Jacobi Bellman (HJB). On introduit la fonction valeur $J(W, V, t)$ définie par

$$J(W, V, t) = \sup_{\phi_s, t \leq s \leq T} E_t^P\left[U(W_T) \middle| W_t = W, V_t = V\right]. \tag{3.21}$$

Notons que la fonction valeur J ainsi définie hérite des propriétés de la fonction d'utilité U. On l'appelle aussi la fonction d'utilité indirecte associée à U. L'équation HJB correspondant à (3.21) est donnée par

$$\max_{\phi} \Big\{ J_t + W(r+\eta V)J_W + (\alpha - \beta V)J_V + \tfrac{W^2}{2} V J_{WW} + \tfrac{\sigma^2}{2} V J_{VV}$$

$$+\rho\sigma V W J_{VW} + \lambda V E\big[J(W(1+\phi Z_t), V, t) - J(W, V, t)\big]\Big\} = 0$$

où l'espérance est prise par rapport à la variable aléatoire Z_t.
On remarque que le contrôle ϕ n'apparaît que dans la partie liée au saut. Cela s'explique par le fait qu'on ait considéré que des actifs risqués au lieu du cadre général où on inclut de plus un actif sans risque dans le portefeuille.
La solution peut être déterminée de façon explicite en donnant une forme particulière à la fonction valeur J. On la choisit de la forme $J(W, V, t) = U(W)F(V, t)$ où F est une fonction régulière, strictement positive telle que $F(V, T) = 1$. Donc $J\Big(W(1+\phi Z_t), V, t\Big) = U\Big(W(1+\phi Z_t)\Big)F(V, t)$ et une condition nécessaire pour que ϕ^* soit optimal est [20] :

$$E\Big[W Z_t\, U'\Big(W(1+\phi^* Z_t\Big)F(V, t)\Big] \;= 0\ . \tag{3.22}$$

Il reste à déterminer les racines de la fonction g

$$g:\ \phi \longrightarrow\ E\Big[Z_t\, U'\Big(W(1+\phi^* Z_t\Big)\Big] \;= 0$$

vu que $F(V, t)$ et W sont constants (indépendant de Z_t) avec $F > 0$ et $W > 0$.

Cette condition est aussi suffisante vu que $U \in \mathcal{U}$. Précisément,

$$g'(\phi) = W E\Big[Z_t^2\, U''\Big(W(1+\phi^* Z_t\Big)\Big] < 0.$$

L'inégalité vient du fait que $U'' < 0$ et $W > 0$.
Donc Il existe une et une seule stratégie ϕ^* optimale pour ce problème. Elle est donnée par $\phi^* = 0$. En effet, on a :

$$g(0) = E\Big[Z_t\, U'(W)\Big] \;=\; U'(W)\ E[Z_t] = 0. \tag{3.23}$$

20. Nous supposons de plus que U et U' sont tels qu'on peut inverser dérivée et intégrale.

par l'hypothèse (3.16).[21] Ce qui veut dire que toutes les personnes rationnelles ne vont pas intervenir dans le marché. En effet, la stratégie qui consiste à acheter l'actif le moins cher et à vendre le plus cher leur rapporte une utilité moindre à celle de ne détenir que l'actif A.
L'explication vient du fait que les arbitrages sont risqués dans ce modèle limitant l'action de tous les arbitragistes. L'effet *noise trader risk* voir Shleifer et Summers [104] est présent dans notre modèle et correspond à une succession de sauts positifs ou négatifs (accroissement de l'écart).
D'autre part, le fait que les arbitragistes soient inactifs valide d'une certaine façon la description des prix de notre modèle. En effet si les arbitragistes étaient actifs, il aurait fallu un modèle d'équilibre pour intégrer leurs actions sur les dynamiques des processus X et Y vu que celles-ci seront toutes affectées. Ce phénomène apparaîtrait par exemple si la variable Z_t était telle que $E(Z_t) \neq 0$. Dans ce cas, tous les agents rationnels interviennent dans le marché, voir l'équation (3.23) où la stratégie nulle n'est plus optimale car $g(0) \neq 0$. D'où l'intérêt de l'hypothèse $E(Z_t) = 0$ pour que le modèle descriptif soit valide ainsi que les arguments des deux paradigmes.

Notre approche, consistant à prendre pour le processus Y, la dynamique de X plus une partie saut d'espérance nulle, est descriptive. La seconde étape est de voir si cette description est réaliste au sens où elle permet d'intégrer les principales anomalies constatées dans la littérature. D'abord, l'effet momentum est pris en compte. Il apparaît dans notre modèle s'il y a une succession de sauts positifs ou négatifs correspondant à l'effet *noise trader risk*. Ensuite, l'effet renversement est aussi pris en compte. En effet, quand l'intensité des sauts λ est grande et qu'on débute avec une succession de sauts positifs (négatifs), on terminera par des sauts négatifs (positifs) occasionnant un effet retour à la moyenne grâce à l'hypothèse que la variable Z_t est centrée. Enfin, le modèle intègre le fait que les anomalies peuvent être importantes comme dans le cas des actions jumelles Royal Dutch et Shell. En effet la stratégie optimale $\phi^* = 0$ est indépendante du support $[-b_1,\ b_1]$. Prendre donc un b_1 un peu plus élevé donnera des imperfections significatives sans pour autant modifier les arguments des deux paradigmes. On voit que le modèle arrive à reproduire les principales observations empiriques en jouant sur ses paramètres λ et b_1.
Une conséquence qui découle de cette étude est que les modèles de diffusion ayant une composante saut d'espérance nulle permettent de bien tenir en compte les arguments des deux paradigmes. Précisément, la partie saut va intégrer les comportements des agents irrationnels en apportant un excès de volatilité puisque le processus (N_t) est orthogonal à celui du mouvement brownien (B_t). La partie continue de la diffusion jouera quant à elle le rôle de dynamique des prix pour la valeur fondamentale de l'actif.

Bien que le modèle continu soit assez intéressant en termes de possibilités de calcul offertes, il est difficile d'y introduire un modèle d'équilibre spécifiant le comportement des agents. C'est pourquoi dans la partie qui va suivre, on s'intéresse aux modèles discrets faisant intervenir la description d'agents (virtuels). L'objectif sera toujours de décrire un modèle qui puisse intégrer les arguments des deux théories tout en mettant l'accent sur la dynamique des prix qui en découle au niveau macro-économique. Celle-ci pourra alors être utilisée en guise de validation pour voir, d'une part, si elle peut générer les principales propriétés des séries financières et, d'autre part, si elle peut être utilisée comme une stratégie d'investissement pour battre la performance du marché.

21. Une autre façon de voir le résultat est que si nous avons deux variables aléatoires X et $Y = X + \epsilon$ telles que $E(\epsilon|X) = 0$, alors tout investisseur averse au risque choisira X sur Y. Le modèle translate ce fait en temps continu en remplaçant variables aléatoires par processus.

Deuxième Partie :
Modèles
A base
D'agents virtuels.

Chapitre 4

Structure des modèles à agents.

Dans le premier chapitre, on a rappelé la théorie de la finance néo-classique. Ses implications ont été beaucoup testées. Si les premiers résultats lui étaient favorables jusqu'à la fin de la décennie septante, il en demeure moins pour la décennie suivante. Les études empiriques ont montré l'existence de primes de risque élevées, un excès de volatilité et des prévisibilités sur la rentabilité des actions. La finance comportementale a apporté des explications à ces faits en se fondant sur l'approche semi-normative consistant à relâcher ou alléger quelques axiomes de Von Neuman et Morgenstern ou en se basant sur le domaine de la psychologie. Un exemple illustratif dans le dernier cas est celui de Barberis, Huang et Santos [10] basé sur la théorie de Kahnemann et Tversky. Il est possible de trouver aussi des explications rationnelles pour ces faits basées sur l'hétérogénéité voir par exemple Kurz et Motelese [71].
Les modèles à agents se basent aussi sur ce facteur d'hétérogénéité mais en supposant que les agents ne sont pas totalement rationnels. Les investisseurs sont supposés faire de l'apprentissage pour trouver dans quel état se trouve la dynamique de l'économie et parfois communiquent pour tenir en compte l'effet d'interaction sociale. Même si c'est une discipline assez récente, beaucoup de catégories de modèles ont été développées. La méthodologie est de partir du bas vers le haut (bottom-up). On commence d'abord par décrire le comportement des agents au niveau micro pour ensuite remonter aux prix d'équilibre au niveau macro-économique. Ensuite, les prix obtenus sont comparés aux données empiriques réelles des marchés financiers en guise de validation. Les thèses de Grothmann [50] et de Martinez-Jaramillo [85] offrent une revue de littérature assez importante. On se base sur ces références en plus de celle de Lebaron[74] pour décrire les différents compartiments de ces modèles.

4.1 Description des agents.

Il existe deux types de description de façon générale. La première consiste à supposer que les décisions des agents sont indépendantes au sens où il n'y a aucune communication qui prend place entre les acteurs. Les investisseurs restent chacun dans leur coin en émettant des ordres. Aucune influence n'est présente. Ils n'ont en commun que le tableau de bord où sont affichés les prix d'équilibre. Dans cette catégorie entrent presque tous les modèles d'apprentissage. On peut citer le modèle artificiel de Sante Fe développé par Arthur et al. [6] utilisant une approche génétique pour faire évoluer les stratégies de base des agents. Brock et Hommes [17] utilisent quant à eux une approche adaptative sur des stratégies fixes dans le temps. Des modèles d'apprentissage utilisant la programmation génétique commencent à être à la mode, voir la thèse de

Martinez-Jaramillo [85] et les références qui y sont citées.
Toujours dans le cas de l'indépendance des investisseurs, on trouve des modèles où il n'y a pas du tout de phénomène d'apprentissage. Par exemple, les agents émettent des ordres à tout instant de façon aléatoire. On parle alors d'intelligence zéro, voir Gode et Sunder [46].

Dans la seconde catégorie, on retrouve les modèles à agents où une structure de communication est présente. Généralement, les modèles utilisent la théorie du graphe pour modéliser comment les agents sont inter-connectés. Ayant la structure de communication, il restera à préciser comment celle-ci se matérialise en termes de décision. On peut supposer que tous les agents inter-connectés vont avoir la même décision d'achat ou de vente d'actifs. Une autre possibilité est de dire que tout agent va faire une pondération équitable de la croyance des individus avec lesquels il communique. On peut également faire une pondération non équitable où les poids vont varier selon les personnes. On donnera plus de poids au sujet ayant plus de connections matérialisant sa fiabilité. On peut voir, par exemple, Iori [58] qui utilise une structure de lattice pour la communication combinée avec des comportements à seuils. Hoffmann Jager et von Eije [57] proposent plusieurs formes de communication et les prix d'équilibre sont déterminés à l'instar des marchés réels (carnet d'ordres). La motivation de cette seconde approche est surtout de prendre en compte les faits de mimétisme ou de comportement moutonniers que l'on retrouve souvent dans les marchés financiers.

4.2 Mécanisme de marché.

Dans les modèles à agents, il est supposé généralement l'existence d'un actif sans risque et d'un actif risqué. L'actif sans risque peut être, par exemple, une obligation émise par un état ou simplement du cash. L'actif risqué est considéré comme une action ou un taux change. La limitation à deux actifs est faite pour des raisons de commodité. Il faut noter que les modèles à agents basés surtout sur l'apprentissage évolutionniste sont très coûteux en termes de temps d'exécution. On préfère donc s'intéresser à des descriptions de marché simples limitant par la même les pouvoirs explicatifs des modèles, voir LeBaron [74].
Si la description des actifs composant le marché est assez uniforme, le mécanisme de formation des prix est par contre assez hétérogène. Une première approche consiste à faire recours aux fonctions de demande. Celles-ci apparaissent fréquemment dans les modèles à agents où il n'y pas de phénomène d'apprentissage c'est-à-dire ni de fonctions objectives ni critères de performance pour choisir entre différentes alternatives. Le choix des individus est alors complètement aléatoire ou repose sur une information publique qui est interprétée pour prendre des actions telles qu'acheter, vendre ou ne rien faire. Les décisions sont souvent normalisées dans ce cas pour prendre des valeurs dans l'ensemble $\{-1,\ 0,\ 1\}$. Précisément, si on note I_t le signal public qui arrive au marché à l'instant t et ϵ_k le signal perçu par l'agent k intégrant d'autres facteurs tel que la communication (herd) ou des facteurs psychologiques, on matérialise sa décision par

$$\hat{n}_{k,t} = \begin{cases} 1 & si\ \epsilon_k > V_t \\ -1 & si\ \epsilon_k < -V_t \\ 0 & si\ |\epsilon_k| < V_t \end{cases}$$

Ici, V_t est souvent un facteur positif proportionnel à la volatilité courante dont une approximation est donnée par $|r_t| = |\log S_t - \log S_{t-1}|$, voir par exemple Iori [58] ou Ghoulmie et al. [43]. Ensuite, l'excès de demande E_t est déterminé par la somme de l'offre $O(t)$ et de la demande $D(t)$ c'est-à-dire :

$$E_t \;=\; O(t) \;+\; D(t)\ ,$$

où

$$O(t) = \sum_{i:\ \hat{n}_{i,t}<0} \hat{n}_{i,t}, \qquad D(t) = \sum_{i:\ \hat{n}_{i,t}>0} \hat{n}_{i,t}\ .$$

Le prix à la date t, dénoté S_t, est alors ajusté à la date $t+1$ en utilisant l'équation

$$S_{t+1} = f(S_t, E_t)\ .$$

Ici f est une fonction croissante par rapport à sa deuxième variable et vérifie de plus $f(x,0) = x$. Ce qui veut dire que le prix est revu à la hausse (baisse) si l'excès de demande est positif (négatif). Des exemples de fonctions f sont donnés par

$$f(x,y) = x + \lambda y \;\; i.e \;\; S_{t+1} = S_t + \lambda(E_t), \qquad f(x,y) = xe^{\lambda y} \;\; i.e \;\; \log S_{t+1} = \log S_t + \lambda(E_t) \quad (4.1)$$

Des études empiriques ont été faites pour déterminer la forme la plus adéquate pour f. Il a été préconisé des fonctions de type concave comme les fonctions racine carrée ou tangente hyperbolique.

Pour les modèles basés sur un apprentissage, l'approche précédente peut être utilisée aussi, voir Chen et Yeh [21] bien que rare. Souvent, on fait recours à la procédure suivante. A l'instant $t = 0$, il est supposé l'existence d'un nombre fixe de titres offerts disons N et une richesse initiale pour chaque agent k dénotée par $W_{k,0}$. Avec l'argent disponible, chaque agent k fait une demande de titres $\hat{n}_{k,t}$ de sorte à maximiser sa fonction d'utilité. Pour obtenir des formules analytiques, on peut supposer que les investisseurs ont des fonctions d'utilité de type exponentiel et que le prix des actions et des dividendes suivent une loi normale. Dans ce cas, la demande de l'agent k est donnée par

$$n_{k,t} = \frac{E_{k,t}(S_{t+1} + d_{t+1}) - S_t(1+r)}{a_i V_{k,t}(S_{t+1} + d_{t+1})}\ . \tag{4.2}$$

où $E_{k,t}$ et $V_{k,t}$ représentent l'espérance et la variance conditionnelle de l'agent k. Notons que cette stratégie est indépendante de la richesse de l'investisseur. Un modèle général consiste à

résoudre le système suivant

$$\underset{n_{i,t}}{Max} \quad E_{k,t}\Big[U_k(W_{k,t+1})\Big] \ . \tag{4.3}$$

où U_k est la fonction d'utilité de l'agent k qui peut être de type CRRA pour faire dépendre la stratégie optimale de la richesse $W_{k,t}$.

Ensuite le prix d'équilibre S_{t+1} est déterminé en égalisant l'offre et la demande c'est-à-dire

$$\sum_{k=1}^{M} \hat{n}_{k,t} \quad = \quad N. \tag{4.4}$$

L'équation ci-dessus peut être récrite sous la forme $f(S_{t+1}, \theta) = 0$ où θ est constitué de paramètres connus (dividendes, coefficient d'aversion, etc.). On pourra donc déterminer la variable S_{t+1} de façon explicite ou par des méthodes numériques dépendant de la structure de la fonction f. Un exemple d'illustration est le modèle de Levy et al. [78]. On peut voir aussi leur livre [77] pour des informations complémentaires. Tout agent maximise une fonction d'utilité puissance en utilisant un portefeuille composé d'un actif risqué et d'un actif sans risque. Ils partent de l'hypothèse d'homogénéité à celle d'hétérogénéité en regardant à chaque fois la caractéristique des prix obtenus. L'hétérogénéité consiste à définir deux groupes d'agents : les partisans des techniques fondamentales et les partisans de l'hypothèse d'efficience. Les premiers font leurs décisions en se fondant sur les dividendes futurs attendus. Si l'actif est surévalué (sous évalué), ils vendent (achètent) une quantité d'actifs risqués variant selon leur coefficient d'aversion (présence d'incertitude dans les prédictions). Pour le second groupe, l'actif reflète toujours sa valeur fondamentale et leur problème se résume à la diversification optimale entre les deux actifs présents dans le marché.

Enfin, il existe un dernier mécanisme qui simule le fonctionnement des marchés réels. Il est supposé que chaque investisseur propose deux prix (bid et ask). Un prix d'équilibre est ensuite déterminé numériquement de sorte qu'il y ait autant d'échanges possibles. Si dans les deux premiers mécanismes, les transactions sont synchrones au sens où on récolte à chaque instant t la décision de tous les agents en même temps, ici des décalages apparaissent. Un carnet d'ordres est utilisé permettant aux investisseurs de voir le comportement des autres.

4.3 Apprentissage.

Les modèles d'apprentissage sont nombreux dans la littérature. Leur complexité varie selon les hypothèses (stratégie fixe ou évolutionniste).

4.3.1 Stratégies fixes.

Brock et Hommes [17] considèrent le cas où chaque investisseur à un nombre fixe de stratégies disons H. Les stratégies sont publiques et partagées par tous les autres investisseurs. Il n'y pas de communication et chaque agent prend ses décisions en ignorant le comportement des autres. Les stratégies disponibles sont souvent basées sur des techniques fondamentalistes et chartistes. L'apprentissage consistera pour chaque agent à regarder la stratégie qui performe dans un passé récent. La performance est relative à un critère donné (fonction fitness). Vu que les stratégies sont fixes et partagées par tous, les auteurs regardent le facteur $n_{h,t}$ représentant la fraction d'agents utilisant la stratégie h. On s'attend à ce que plus une stratégie se comporte bien plus il y ait d'individus qui l'utilisent. Précisément, la fraction d'agents utilisant la stratégie $h = 1, \cdots, H$ est donnée par

$$n_{h,t} = \frac{\exp[\beta U_{t,h}]}{Z_t}, \qquad Z_t = \sum_{h=1}^{H} \exp[\beta U_{t,h}] \tag{4.5}$$

où Z_t est juste un facteur de normalisation. Dans la pratique, on prend généralement $H = 2$ qui consiste à dire que chaque agent est soit fondamentaliste ou chartiste. Le terme $U_{t,h}$ représente le critère de performance de la stratégie h et β est une mesure d'intensité qui est toujours positive ($\beta \geq 0$). Notons que plus la performance d'une stratégie h est grande ($U_{t,h}$ grand) plus la fraction $n_{h,t}$ l'est aussi. Dans le cas extrême $\beta = 0$, tous les agents sont équi-probablement répartis sur les stratégies c'est-à-dire $n_{h,t} = 1/H$ et dans le cas $\beta = \infty$, tous les investisseurs utilisent une seule stratégie (la plus performante). Le critère de performance $U_{t,h}$ est défini de façon récursive en utilisant son terme précédent plus une mesure de profit correspondant à la date la plus récente. A chaque instant, une mise à jour des fractions des agents données par l'équation (4.5) est faite en tenant compte de la déviation des prix d'équilibre de leur valeur fondamentale.

Un autre exemple est celui de Routledge [99] qui applique la technique d'apprentissage au modèle répété de Grossman et Stiglitz [49]. Le principe est le même. Ici, l'apprentissage consistera à voir si on est prêt à débourser de l'argent pour acquérir une information et si oui à quelle proportion. Bien entendu, plus on investit, plus on améliore la prédiction. On peut aussi ne rien dépenser si on a confiance en sa capacité individuelle d'inférence. Dans une telle situation, il faudra chercher la bonne fonction parmi celles que l'on dispose pour améliorer les résultats de prédiction. Les stratégies ne sont pas identiques entre les agents comme dans le modèle précédent de Brock et Hommes. L'hypothèse d'indépendance des agents est également relâchée en permettant le mimétisme dans l'apprentissage des agents. Un individu qui n'est pas satisfait de ses prédictions peut aller copier les stratégies des autres.
Pour l'approche de l'apprentissage adaptatif, seul le critère sélection est utilisé (fitness). Dans la sous-section suivante, d'autres étapes telles que le croisement ou la mutation sont introduites.

4.3.2 Stratégies évolutionnistes.

Comme le nom l'indique, les stratégies subissent ici certaines transformations dans le temps. Les modifications apportées proviennent essentiellement des algorithmes génétiques. L'idée est de tenir compte du fait que le raisonnement des investisseurs peut changer au fil de leur expérience.

La description de ces modèles est la suivante.
On part d'un élément de base appelé individu qui joue le rôle de chromosome dans la génétique. Dans le domaine de la finance, cet individu peut être une stratégie que l'on utilise en termes d'allocation de portefeuilles. Elle peut être aussi une fonction (choix de paramètres) pour des études économétriques. L'objectif est de chercher l'individu qui va maximiser un critère donné. Si la fonction à maximiser à de bonnes propriétés, on peut faire recours aux techniques usuelles telles que les méthodes de gradients. Dans le cas contraire, on utilise un algorithme de recherche stochastique. Une de ses particularités est qu'on travaille sur un nuage d'individus. Si l'algorithme arrive donc à converger, on aura plusieurs solutions au lieu d'en avoir une seule. Souvent la population initiale est choisie de façon aléatoire dans un domaine admissible (espace de recherche) que l'on notera $\mathcal{D}$. Pour pouvoir appliquer les algorithmes génétiques, on passe à un système de codage des individus qui est généralement binaire. Un individu A est donc de la forme $A = \{a_1, a_2, \cdots, a_{l-1}, a_l\}$ où pour tout $i \in \{1, \cdots, l\}$ $a_i \in \{0, 1\}$. Le paramètre l représente la taille du chromosome et permet de gérer l'aspect précision dans les calculs.
Après avoir mis les stratégies sous forme de bits, la seconde étape consiste à définir une fonction de performance notée f qui va permettre d'accorder des poids différents aux individus. Dans l'exemple de Brock et Hommes, la fonction exponentielle a été prise. De façon générale, toute fonction de $\mathbb{R}$ à valeurs dans $\mathbb{R}_+$ peut être prise. Il faudra simplement trouver au préalable une fonction d pour transformer tout individu A (sous forme binaire) en une forme réelle puis choisir la fonction f définie de $\mathcal{D} = d(\{0, 1\}^l) \subset \mathbb{R}$ à valeurs dans $\mathbb{R}_+$. Le choix de la fonction d dépendra du problème économique.
Ayant défini le critère d'évaluation f, on peut s'intéresser aux individus qu'il faudrait sélectionner pour la prochaine génération. La logique voudrait que plus un individu performe plus on doit augmenter ses chances d'être reconduit (probabilité de survie p_i). Un critère similaire à celui de Brock et Hommes peut être pris appelé la roue de la loterie biaisée. Il consiste à faire une sélection proportionnelle i.e

$$p_i = \frac{f(d(c_i))}{\sum_{i=1}^{N} f(d(c_i))}.$$

Cette technique revenant à choisir des individus à partir de leur poids peut donner des solutions sous-optimales s'il existe un individu dominant très distant de la solution optimale. Pour ne pas ignorer l'effet que pourrait apporter les autres individus, il a été proposé, dans la littérature, la méthode du tournoi. Elle consiste à tirer aléatoirement deux individus puis en prendre que le meilleur pour la prochaine génération. Elle capte l'idée de la sélection proportionnelle en donnant plus de poids aux stratégies moins performantes.

Dans cette première étape de fitness, on travaille toujours sur les mêmes stratégies et on peut supposer dans un sens large que l'on se trouve dans un cadre adaptatif. L'approche évolutionniste est intégrée dans les deux points suivants à savoir une partie enjambement encore appelée croisement ou recombinaison qui va permettre d'explorer de nouveaux types de stratégies et une partie mutation qui va transformer des stratégies existantes par complémentation binaire par exemple. Dans le cas de l'étape croisement, deux individus (chromosomes) tirés dans la population sont couplés pour donner des enfants. Précisément, ces enfants sont obtenus par un système d'échange d'information entre les parents à un ou différents segments de leurs parties. Si aucun échange n'apparaît alors les enfants sont identiques aux parents. Les échanges (croi-

sement) se font à travers une probabilité p_c. Plus cette probabilité est grande, plus on élargit notre ensemble de recherche. Cette étape n'est cependant pas toujours suffisante pour parcourir tout l'ensemble admissible surtout si la population initiale n'est pas très diversifiée. Il est utile de permettre à l'algorithme d'aller chercher dans d'autres horizons. C'est ce que permet l'étape mutation. Contrairement à l'étape croisement, ici un seul individu est tiré aléatoirement avec une certaine probabilité p_m. Dans le cas du codage binaire, une complémentation est faite i.e les bits 1 sont remplacés par 0 et les bits 0 par 1. Cette étape permet de donner la propriété d'ergodicité aux algorithmes génétiques en permettant une exploration de tous les points du domaine. Souvent la valeur p_m est petite en pratique pour ne pas trop perturber la dynamique de convergence.

L'algorithme génétique, décrit ci-dessus, peut être combiné avec un système classificateur. Cela permet d'accroître la vitesse de convergence. Un tel algorithme a été utilisé par Arthur et al. [6]. Dans leur modèle, l'information est modélisée par une suite de bits (12 précisément). Chaque bit représente un indicateur qui peut être fondamentaliste ou chartiste. A tout instant t, selon l'état du marché, il existera des stratégies qui seront activées et le choix ne s'opérera que sur cet ensemble à travers le critère de performance. Leur objectif était de voir si à partir d'agents hétérogènes ayant plusieurs modèles linéaires de prévision, il était possible d'obtenir une convergence vers une croyance commune qui sera bien entendu linéaire aussi et de la forme $p_t = ad_t + b$. Les paramètres (a, b) de cette croyance commune peuvent être déterminés de façon explicite et vont constituer une sorte de référence pour la comparaison. Les résultats obtenus sont que si les agents font une mise à jour espacée de leurs stratégies alors le phénomène d'apprentissage conduit vers une converge de croyance (a', b') égale à celle de (a, b) et donc confirmant l'hypothèse de l'efficience de marché. Les stratégies chartistes ne sont pas rentables dans ce cadre. Par contre, quand on augmente la fréquence d'adaptation des stratégies, on se retrouve dans un marché assez complexe où des phénomènes de crash ou de bulles apparaissent temporairement. Le marché devient inefficient et des stratégies chartistes arrivent à performer.
A côté de ce modèle combinant génétique et système classificateur, d'autres combinent programmes informatiques et génétique. Dans ce cas, l'individu est remplacé par une liste d'instructions. Un exemple de programme informatique sera par exemple de la forme :

Programme : Règle de trading Numéro X

SI COND1 ET COND 2

Alors DEC1

SINON SI COND 3 OU COND 4

Alors DEC2

SINON

DEC3.

Dans le domaine de la finance, les conditions peuvent être des indicateurs de type chartiste ou fondamentaliste et les décisions des ordres tels que achat, vente ou ne rien faire. Le programme

est ensuite codé sous forme d'arbre puis des algorithmes évolutionnistes lui sont appliqués comme précédemment. Par exemple, pour une approche de test de la théorie de l'efficience, Chen et Yeh [20] utilisent la programmation génétique pour voir si l'instruction Achat et Conservation est optimale. Pour des données ayant une historique longue, la stratégie nulle équivalente à ne rien faire prédomine confirmant la théorie de l'efficience. Cependant, pour un petit échantillon, cette stratégie n'est plus optimale et des prévisibilités apparaissent. Ces résultats, obtenus avec des algorithmes poussés dans le cas Chen et Yeh [20] aussi bien dans le cas de Arthur et al. [6], collaborent nos propos à savoir que la frontière entre efficience et inefficience de marchés est très légère, voir aussi les études empiriques faites dans le chapitre 3. Etant donné qu'il n'est pas possible de donner une réponse binaire, une alternative est de considérer le concept d'efficience en moyenne où on retrouve à la fois les arguments des partisans de l'efficience de marché et de ceux de la finance comportementale. C'est cette idée que l'on formalise dans le chapitre qui suit. Avant cela, on termine la description des modèles à agents en précisant leur aspect validation.

4.4 Critères de validation.

Après avoir décrit au niveau micro le comportement des agents et avoir spécifié la procédure de formation des prix, la prochaine étape est l'aspect validation. C'est une étape importante car un modèle est évalué par sa capacité à rendre compte les faits empiriques constatés comme l'avait mentionné Friedman [40]. Il est donc utile de vérifier si les prix que le modèle a produit possèdent les propriétés notées sur les séries financières encore appelés faits stylisés. Il en existe, aujourd'hui, plusieurs recensés. On se limite ici aux plus répandus. Avant de les énumérer, j'ai représenté l'évolution des prix des quatre indices étudiés précédemment à savoir l'indice CAC 40 (France), l'indice NIKKEI 225 (Japon), l'indice Dow Jones Industrial Average (DJIA) (Etat-Unis) et l'indice Nasdaq composite (Etats- Unis) dans la figure 4.1. Les données sont journalières et couvrent la période du 1 Mars 1990 au 16 Octobre 2009.

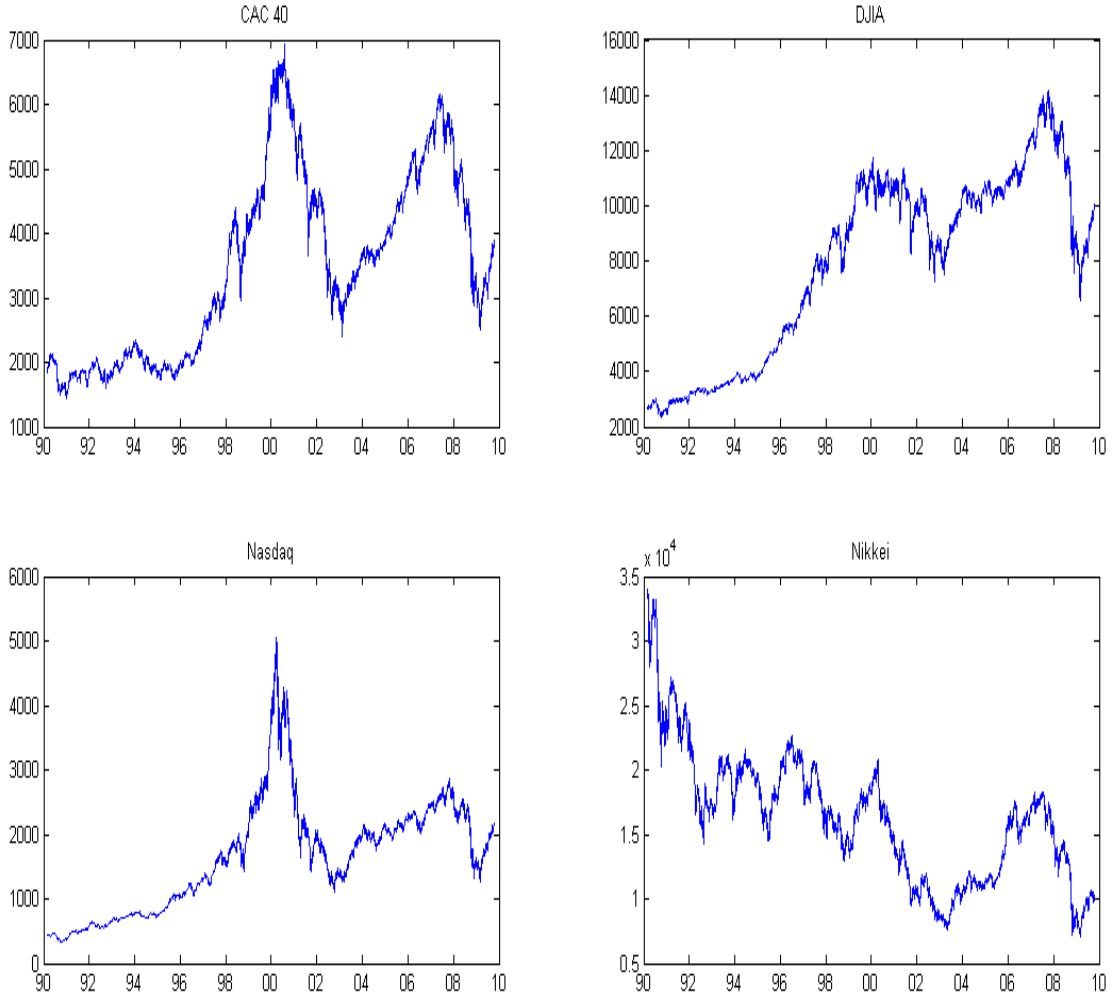

FIGURE 4.1 – Evolution des prix d'indice CAC 40, Dow Jones, Nasdaq et NIKKEI.

Avec la série des prix, il est difficile de noter la similarité des courbes. Pour voir les faits stylisés, il est plus commode de travailler sur la série des rentabilités $r_t = \log S_t - \log S_{t-1}$ qui permet de rendre les séries au moins stationnaire par rapport au moment d'ordre 1. Nous avons représenté dans la prochaine figure la série des rentabilités des différents indices (première ligne de la figure) ainsi que les estimations des valeurs d'auto-corrélation de $signe(r_t)$ et r_t (deuxième et troisième ligne de la figure).

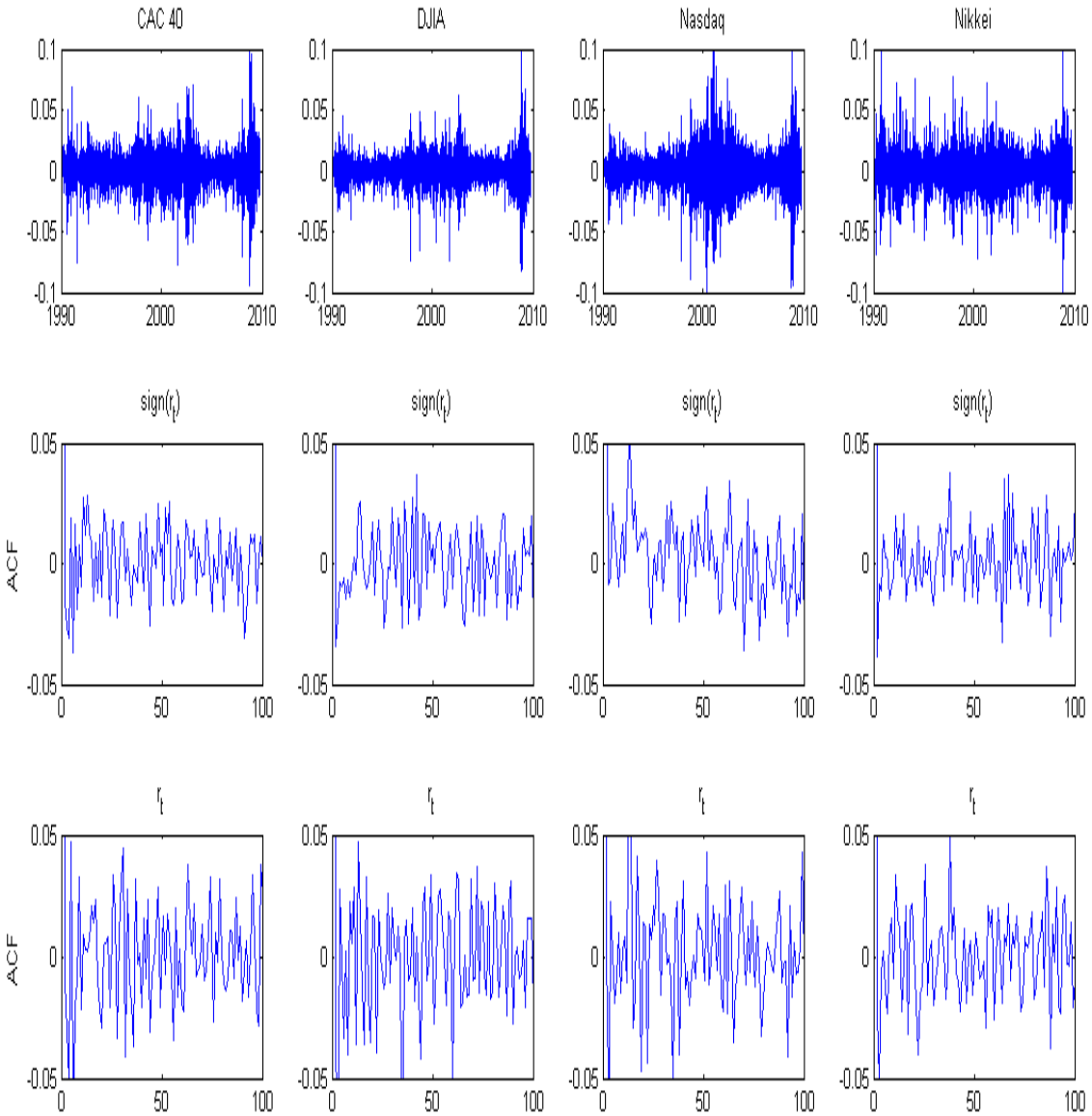

FIGURE 4.2 – Rentabilités et fonction d'auto-corrélations(ACF) pour les 4 indices.

Nous remarquons les faits suivants :

Fait 1. Quasi marche aléatoire.
Pour toutes les séries, la positivité ou la négativité des rentabilités r_t arrive de façon quasi aléatoire comme le montre les valeurs estimées de la fonction d'auto-corrélation de $signe(r_t)$. Nous remarquons que les valeurs sont toutes petites en valeur absolue et inférieures à 0.05. Il en est de même pour les valeurs estimées de la fonction d'auto-corrélation de r_t (équation 3.1). Il est donc difficile de prédire les rentabilités des titres en utilisant des modèles linéaires.

Fait 2. Effet de regroupement de volatilité.
Le graphique 4.2 montre que les périodes de grandes et petites volatilité ont tendance à se regrouper ensemble. En d'autres termes, si à un instant donné la rentabilité d'une action est assez forte, il y a de forte chance que la suivante l'est aussi même si elle peut changer de signe. La même remarque s'applique pour les petites volatilités. Une rentabilité faible a tendance à être suivie par une rentabilité faible. Une manière de mettre en relief ce regroupement de volatilité est de voir la fonction d'auto-corrélation des séries de la forme $|r|^q$ où q est un réel positif. Nous illustrons ce fait dans la prochaine figure pour $q = 1$ et $q = 2$.

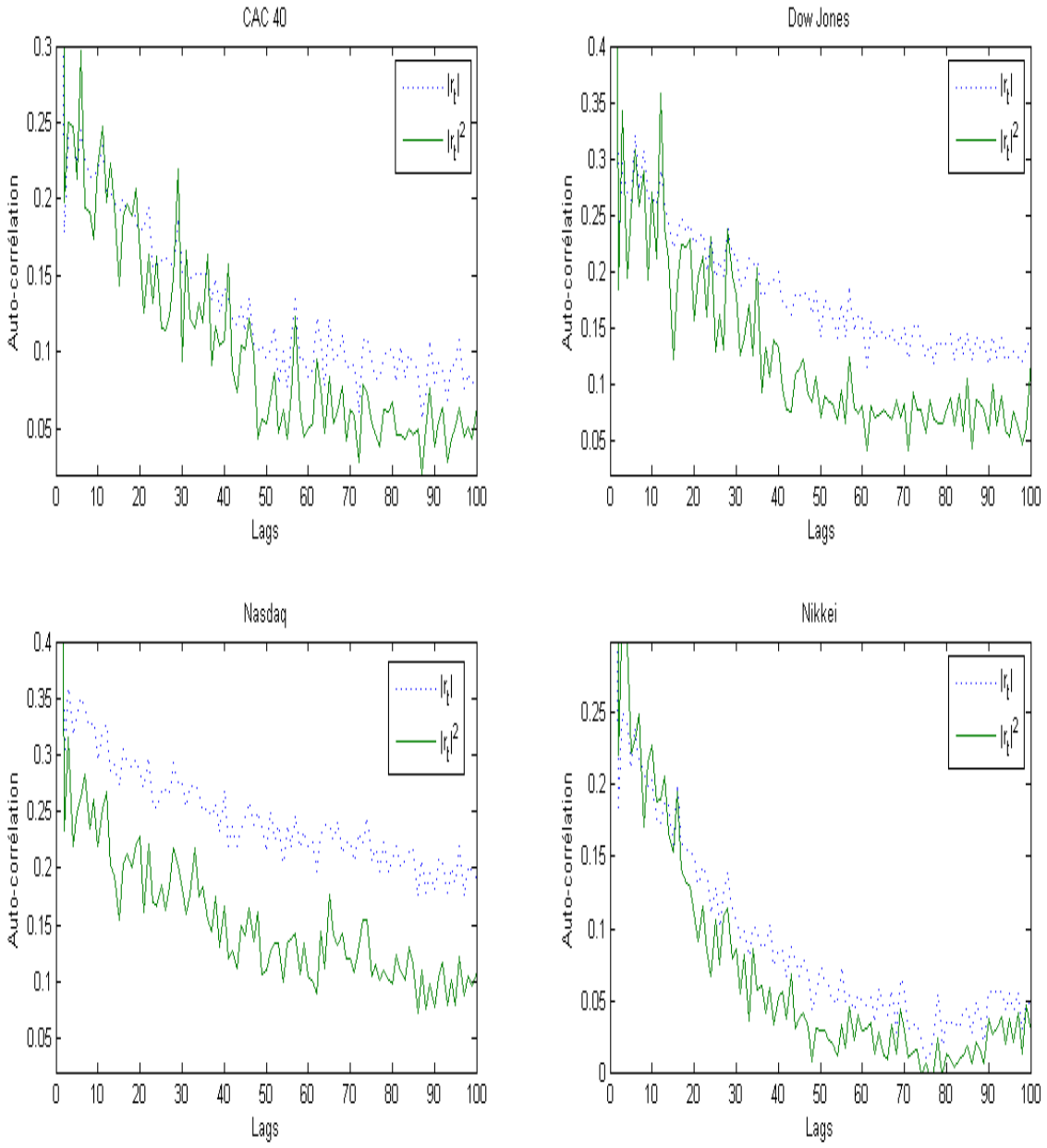

FIGURE 4.3 – Valeurs des estimations de la fonction d'auto-corrélation de $|r_t|^q$, $q = 1, 2$.

Pour toutes les courbes, un phénomène de décroissance vers zéro apparaît avec une certaine persistance plus prononcée pour $q = 1$ et ceci indépendamment de l'indice utilisé.
Ce fait stylisé a motivé le développement des modèles (G)ARCH ou des modèles à volatilité stochastiques comme celui de la dynamique X donnée dans l'équation (3.13).

Fait 3. Distribution non gaussienne de r_t.
La distribution de la rentabilité des actifs s'écarte de la loi normale. Les événements extrêmes apparaissent plus fréquemment donnant des queues de distributions plus épaisses. De plus, on n'a pas une symétrie de distribution comme dans une loi normale standard. La skewness est généralement négative. Dans le prochain tableau, nous représentons la skewness, le kurtosis, la statistique de Jacque Bera et sa p-value correspondante.

r_t	$CAC40$	$DJIA$	$Nasdaq$	$NIKKEI225$
$Skewness$	$-0,0387$	$-0,1165$	$-0,0491$	$-0,0154$
$Kurtosis$	$7,7585$	$11,5586$	$8,8684$	$8,2586$
JB	$4,6780$	15116	7103	$5,565$
$p-value$	$< 1,00e-03$	$1,00e-03$	$1,00e-03$	$1,00e-03$

Les tests de Jacque Bera rejettent l'hypothèse nulle de normalité pour toutes les quatre séries au seuil de signification 1%.

Fait 4. Distribution de $|r_t|$.

La moyenne et l'écart type de la série des valeurs absolues $|r_t|$ sont presque égales. Nous avons calculé les quatre premiers moments de la série $(|r_t|)$.

$\|r_t\|$	$CAC40$	$DJIA$	$Nasdaq$	$NIKKEI$
moyenne	$0,0101$	$0,0077$	$0,0108$	$0,0114$
$\text{écart} - type$	$0,0099$	$0,0083$	$0,0116$	$0,0111$
skewness	$2,5708$	$3,2273$	$2,6932$	$2,6162$
kurtosis	$14,5297$	$22,2214$	$14,9574$	$16,6938$

On voit que pour les quatre indices, la moyenne et l'écart type sont assez proches. Ce fait est plus apparent au niveau des indices CAC 40 et Nikkei.

La validation des modèles à agents virtuels se fait en regardant si les prix qu'ils fournissent reflètent les faits stylisés mentionnés ci-dessus parmi tant d'autres. Cela a été le cas de pour plusieurs modèles. On peut citer, en guise d'illustration, les méthodes basées sur l'apprentissage adaptatif comme celui de Brock et Hommes [17], les méthodes génétiques de Lebaron et al. [75], le modèle Levy et al. [78]. Cependant, des critiques ont été apportées sur certains modèles dont le nombre de paramètres entrant en jeu était élevé. Cela donne une certaine flexibilité qui permet de reproduire beaucoup de choses. C'est pourquoi, il est préconisé de tester la robustesse des modèles au niveau des paramètres. On pourra modifier par exemple le nombre d'agents ou le nombre de stratégies initiales pour les modèles basés sur l'algorithme génétique pour regarder si on obtient les mêmes résultats. D'autres modèles sont arrivés à reproduire les principales caractéristiques des séries financières avec moins de complexité. C'est le cas de Ghoulmie et al. [43] basé sur des comportements à seuil des agents. On peut combiner cette technique d'activation avec une structure de communication comme dans Iori [58]. Des modèles simples d'interaction comme celui de Alfarano et Lux [3] reproduisent aussi les faits stylisés tout en permettant une approche analytique. La liste est loin d'être exhaustive. On peut voir la revue de Samanidou et al. [100] ou celle de Cont [24] pour plus d'informations.

Chapitre 5

Tentative de réconciliation des deux paradigmes.

Dans le chapitre précédent, on a rappelé le fonctionnement des modèles à agents virtuels. De nos jours, il existe plusieurs types de modèles qui ont été développés arrivant à reproduire les principales propriétés des séries financières. Cet aspect est une sorte de validation de la méthodologie utilisée. Cependant pour faciliter l'interprétation des résultats, il est important de réduire la complexité des modèles pour savoir réellement quels sont les principaux facteurs responsables de ces faits stylisés. C'est pourquoi, beaucoup de modèles tentent de nos jours de simplifier la description des agents. Cela permet d'avoir des modèles semi-analytiques facilitant l'interprétation des résultats. Un consensus qui découle de ces modèles à agents virtuels est que les faits stylisés proviennent des interactions entre groupes d'agents. Précisément, l'interaction se passe entre fondamentalistes et bruiteurs. Les individus passent d'un groupe à l'autre dû à la présence d'incertitude au niveau de la dynamique des prix, voir Alfarano et Lux [3]. Alfarano [2] a investi dans sa thèse les modèles avec phénomène de switching puis les a relié aux modèles à volatilité stochastique. Ici, on établit un lien avec les modèles auto-régressifs à coefficient aléatoire.

5.1 Description et prix d'équilibre.

Nous proposons un modèle qui suit la description de Ghoulmie et al.[43]. La seule différence est que l'on introduit un phénomène d'interaction entre deux groupes : les fondamentalistes et les bruiteurs. Ce fait permet d'avoir en tout instant des acheteurs et des vendeurs. On se positionne dans un cadre où les données de marchés arrivent avec une fréquence haute. L'intervalle entre deux instants est mesuré en termes de secondes ou minutes par exemple. On peut dans un tel contexte négliger l'effet du taux sans risque et des dividendes. Les prix des actifs bougent sous l'influence du flux d'information qui apparaît à chaque instant dans le marché. On notera cet ensemble d'information qui est supposé publique $(I_t)_{t\in\mathbb{T}=\{0=t_0,\cdots,t_n,\cdots\}}$. La durée entre deux transactions est notée $h = t_k - t_{k-1}$ et est supposée constante. A l'instant t_{n-1}, l'information I_{t_n} est reçue par chaque agent indexé disons par $k \in \mathbb{J} = \{1, \cdots, M\}$ qui va l'interpréter à sa guise en tenant compte de l'état du marché. Cette interprétation lui donnera un nouveau signal $R^k_{t_n}$ qui va varier d'un investisseur à l'autre dû à l'hétérogénéité. Formellement, on a

$$R^k_{t_n} = I_{t_n} + \epsilon^k_{t_n}, \quad \epsilon^k_{t_n} = \sigma_k p^N_{t_n} \cdot Z^k_{t_n}. \tag{5.1}$$

où

$$p_{t_n}^N = \sum_{i=1}^{N} \omega_i |r_{t_{n-i}}|, \quad r_{t_{n-i}} = \log S_{t_{n-i}} - \log S_{t_{n-i-1}}, \quad \forall i \ \omega_i \geq 0, \quad \sum_{i=1}^{N} \omega_i = 1. \tag{5.2}$$

Le terme $p_{t_n}^N$ représente l'approximation de la volatilité courante du marché et le terme N sur p_t^N spécifie le nombre de termes utilisés pour approximer cette volatilité courante. Par exemple, dans Ghoulmie et al. [43] ou Iory [58], celle-ci est prise égale à $|r_{t_{n-1}}|$ correspondant dans notre cas à $N = 1$ et $\omega_1 = 1$. Ensuite, la décision des agents est faite proportionnellement à leur signal reçu. Pour l'agent k, on a par exemple :

$$n_{t_n}^k = c_k \, R_{t_n}^k, \qquad c_k \in \mathbb{R}_+^*, \tag{5.3}$$

où c_k est une constante strictement positive. En d'autres termes, l'agent k achète (vend) $|n_{t_n}^k|$ parts de l'actif si l'information reçue $R_{t_n}^k$ est positive (négative). On considère enfin les assomptions suivantes.

Assomption 1.
Pour tout agent $k \in \mathbb{J}$, $(Z_t^k)_{t\in\mathbb{T}}$ et $(I_t)_{t\in\mathbb{T}}$ sont des processus bruits blancs Gaussiens mutuellement indépendants avec $I_t \sim \mathcal{N}(0, \sigma_I^2)$, $Z_t^k \sim \mathcal{N}(0, 1)$.

Assomption 2.
Pour tout temps $t \in \mathbb{T}$, les variables aléatoires des agents $(Z_t^k)_{k\in\mathbb{J}}$ sont mutuellement indépendantes.

Assomption 3.
Le paramètre σ_k donné par l'équation (5.1) est plus petit que σ_I (écart type de I_t).
$0 < \sigma \leq \sigma_k < \sigma_I$.

Avant de donner les propriétés qui découlent du modèle à agents, nous discutons les suppositions faites ci-dessus.

Le terme $\epsilon_{t_n}^k$, provenant de (5.1), est un bruit idiosyncratique qui apparaît lorsque les investisseurs évaluent l'information fondamentale I_{t_n}. Donc (I_t) et (Z_t^k) doivent être mutuellement indépendants pour tout $k \in \mathbb{J}$ afin d'éviter toute causalité entre ces deux processus. C'est le sens de la première assomption.

L'assomption 2 prend en compte le manque de communication entre les agents. Ces derniers, dans notre modèle, font leurs décisions sans interactions. C'est pourquoi, les variables aléatoires

$(Z_t^k)_{k\in\mathbb{T}}$ sont supposées être mutuellement indépendantes pour tout $t \in \mathbb{T}$.

La dernière assomption introduit une rationalité bornée.[22] Elle implique que le biais $\epsilon_{t_n}^k$ est souvent plus petit que I_{t_n} en valeur absolue. Comme conséquence, la probabilité que $R_{t_n}^k = I_{t_n} + \epsilon_{t_n}^k$ soit de même signe que I_{t_n} est grande. Précisément, si le nombre d'agents M est très élevé alors les ensembles $\{k \in \mathbb{J},\ R_{t_n}^k \cdot I_{t_n} \geq 0)$ et $\{k \in \mathbb{J},\ R_{t_n}^k \cdot I_{t_n} < 0)$ sont tous non vides (presque surement). Cette assertion est formellement montrée dans la propriété 1.

Propriété 1 : Présence d'acheteurs et de vendeurs.

Nous devons montrer que lorsque le nombre de traders M est grand (tendant vers l'infini), la probabilité d'obtenir que des acheteurs ou que des vendeurs tend vers zéro à tout instant t_{n-1}, où $t_{n-1} \geq h$.

Puisque $n_{t_n}^k = c_k R_{t_n}^k$ avec $c_k > 0$, nous avons besoin de montrer seulement que $P(R_{t_n}^k \leq 0,\ \ k = 1, \cdots, M) \to 0$ si $M \to +\infty$. On a :

$$P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \leq 0\}\Big) = P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \leq 0\} \cap \{I_{t_n} > 0\}\Big) + P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \leq 0\} \cap \{I_{t_n} \leq 0\}\Big)$$

$$= P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} > 0\}\Big)P(\{I_{t_n} > 0\}) + P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} \leq 0\}\Big)P(\{I_{t_n} \leq 0\})$$

$$= \prod_{k=1}^{M} P\Big(\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} > 0\}\Big)P(\{I_{t_n} > 0\}) + \prod_{k=1}^{M} P\Big(\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} \leq 0\}\Big)P(\{I_{t_n} \leq 0\})$$

où dans la dernière égalité, nous avons utilisé le fait qu'à l'instant t_{n-1}, étant donné I_{t_n}, les variables aléatoires $R_{t_n}^1, \cdots, R_{t_n}^M$ sont indépendantes, voir (5.1) et l'assomption 2. Posons :

$$A_k = P\Big(\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} > 0\}\Big), \quad \mathbf{A} = max\{A_k, k \in \mathbb{J}\}. \tag{5.4}$$

$$B_k = P\Big(\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} \leq 0\}\Big) \quad \mathbf{B} = max\{B_k, k \in \mathbb{J}\}. \tag{5.5}$$

L'assomption 3 implique que $\mathbf{A}$, $\mathbf{B}$ sont presque surement dans $]0, 1[$ tandis que l'assomption 1 donne $P(I_t > 0) = P(I_t \leq 0) = \frac{1}{2}$. Donc

22. La rationalité complète (absence de biais) est obtenue en posant $\sigma_k = 0$ pour tout agent $k \in \mathbb{J}$.

$$P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \leq 0\}\Big) = \frac{1}{2}\Big[\prod_{k=1}^{M} P\Big(\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} > 0\}\Big) + \prod_{k=1}^{M} P\Big(\{R_{t_n}^k \leq 0\}\Big|\{I_{t_n} \leq 0\}\Big)\Big]$$

$$\leq \frac{1}{2}(\mathbf{A}^M + \mathbf{B}^M) \longrightarrow 0 \;\; quand \;\; M \longrightarrow +\infty.$$

De façon similaire, on montre que la probabilité de n'avoir que des acheteurs $P\Big(\bigcap_{k=1}^{M}\{R_{t_n}^k \geq 0\}\Big)$ tend vers zéro si M tends vers l'infini.

On en conclut qu'à tout instant nous avons des acheteurs et des vendeurs si le nombre d'investisseurs M est élevé. Cette propriété est satisfaite presque surement. L'exception apparaît s'il existe $t_n \in \mathbb{T}$ tel que $p_{t_n}^N = 0$ qui est un événement négligeable pour une variable aléatoire continue. On pourra appeler un agent, fondamentaliste à la date t_n, si son signal perçu $R_{t_n}^k$ est de même signe que l'information fondamentale I_{t_n} c'est-à-dire $R_{t_n}^k \cdot I_{t_n} \geq 0$. Dans le cas contraire, on l'appellera bruiteur ($R_{t_n}^k \cdot I_{t_n} < 0$).

Propriété 2 : Présence de phénomène de permutation (switching).

On doit montrer qu'un agent fondamentaliste ne peut le demeurer toujours. En d'autres termes, il passera dans un instant futur dans le camp des bruiteurs et vice-versa pour les bruiteurs.

Montrons d'abord que l'on ne peut être toujours dans le camp des bruiteurs. Nous avons grâce à (5.1)

$$\{R_{t_n}^k \cdot I_{t_n} < 0\} = \{\sigma_k p_{t_n}^N |Z_{t_n}^k| > |I_{t_n}|\} \cap \{I_{t_n} \cdot Z_{t_n}^k \leq 0\} \tag{5.6}$$

Donc $\{R_{t_n}^k \cdot I_{t_n} < 0\} \subset D_{t_n}$ où $D_{t_n} = \{I_{t_n} \cdot Z_{t_n}^k \leq 0\}$. Les événements D_{t_n}, $t_n \in \mathbb{T}$ sont indépendants d'après l'assomption 1. Cette dernière donne aussi que la probabilité d'obtenir D_{t_n} est $P(D_{t_n}) = P(I_{t_n} \geq 0, Z_{t_n}^k \leq 0) + P(I_{t_n} \leq 0, Z_{t_n}^k \geq 0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$. Donc

$$P(\bigcap_{n=1}^{T}\{R_{t_n}^k \cdot I_{t_n} < 0\}) \leq P(\bigcap_{n=1}^{T} D_{t_n}) = \prod_{n=1}^{T} P(D_{t_n}) = \frac{1}{2^T} \longrightarrow 0, \quad quand\ T \longrightarrow +\infty. \tag{5.7}$$

D'autre part, on ne peut pas aussi rester toujours fondamentaliste. En effet, d'après (5.6), on a :

$$\{R_{t_n}^k \cdot I_{t_n} \geq 0\} = \{R_{t_n}^k \cdot I_{t_n} < 0\}^c \;=\; \{\sigma_k p_{t_n}^N |Z_{t_n}^k| \leq |I_{t_n}|\} \cup \{I_{t_n} \cdot Z_{t_n}^k > 0\}$$

où A^c représente le complémentaire de l'ensemble A. On pose $c_{min} = min\{p_t^N, t \in \mathbb{T}\}$ qui est strictement positive presque surement puisque l'événement $\{c_{min} = 0)\}$ est identique à celui de $\cup_{t_n \in \mathbb{T}}\{p_{t_n}^N = 0\}$. Donc

$$\{R_{t_n}^k \cdot I_{t_n} \geq 0\} = \{\sigma_k p_{t_n}^N |Z_{t_n}^k| \leq |I_{t_n}|\} \cup \{I_{t_n} \cdot Z_{t_n}^k > 0\} \subseteq \{\sigma_k c_{min} |Z_{t_n}^k| \leq |I_{t_n}|\} \cup \{I_{t_n} \cdot Z_{t_n}^k > 0\} = F_{t_n}$$

où F_{t_n}, $t_n \in \mathbb{T}$, sont des événements indépendants (Assomption 1). Pour toute date t_n, la probabilité $P(F_{t_n})$ est dans l'intervalle $]\frac{1}{2}, 1[$ puisque $c_{min} > 0$. D'où,

$$P(\bigcap_{n=1}^{T}\{R_{t_n}^k \cdot I_{t_n} \geq 0\}) \leq P(\bigcap_{n=1}^{T} F_{t_n}) = \prod_{n=1}^{T} P(F_{t_n}) \longrightarrow 0, \ quand\ T \longrightarrow +\infty. \tag{5.8}$$

Les équations (5.7) et (5.8) nous disent qu'un agent k ne peut toujours demeurer un bruiteur ou un fondamentaliste. Par conséquent, le modèle à agents virtuels fournit un phénomène de permutation.

Enfin, nous montrons une dernière propriété de notre modèle à agents virtuels.

Propriété 3. Dynamique RCA(p) pour le log des prix.

Pour trouver la dynamique des prix, on a besoin de spécifier la fonction d'impact sur l'excès de demande. Précisément, cette fonction f est choisie ici de façon log-linéaire, voir l'équation (4.1) où $\log S_{t_n} - \log S_{t_{n-1}} = \frac{E_{t_n}}{\lambda}$. Le paramètre λ est une constante positive servant à normaliser l'excès de demande E_{t_n}.

En utilisant les équations (5.1) et (5.3), cet excès E_{t_n} est déterminé par :

$$E_{t_n} = \sum_{k=1}^{M} n_{t_n}^k = \sum_{k=1}^{M} c_k R_{t_n}^k = I_{t_n}(\sum_{k=1}^{M} c_k) + p_{t_n}^N \sum_{k=1}^{M} c_k \sigma_k Z_{t_n}^k. \tag{5.9}$$

Donc,

$$\log S_{t_n} = \log S_{t_{n-1}} + \frac{E_{t_n}}{\lambda} = \log S_{t_{n-1}} + \tilde{I}_{t_n} + p_{t_n}^N Z_{t_n} \tag{5.10}$$

où

$$\tilde{I}_{t_n} = cI_{t_n}, \quad c = \frac{\sum_{k=1}^{M} c_k}{\lambda} \quad Z_{t_n} = \frac{1}{\lambda}\sum_{k=1}^{M} c_k \sigma_k Z_{t_n}^k. \tag{5.11}$$

Dans la suite, nous illustrons nos propos pour $p_t^N = |r_{t_{n-1}}|$ (N=1, $\omega_1 = 1$) car l'idée étant la même que dans le cadre général. Puisque $p_{t_n}^N = |r_{t_{n-1}}| = 1_{\{r_{t_{n-1}} \geq 0\}}(\log S_{t_{n-1}} - \log S_{t_{n-2}}) +$

$1_{\{r_{t_{n-1}}<0\}}(\log S_{t_{n-2}} - \log S_{t_{n-1}})$, l'équation (5.10) devient

$$\log S_{t_n} \;=\; \log S_{t_{n-1}}\Big(1 + \Big(1_{\{r_{t_{n-1}}\geq 0\}} - 1_{\{r_{t_{n-1}}<0\}}\Big)Z_{t_n}\Big) \;+\; \tilde{I}_{t_n} \;+\; \log S_{t_{n-2}}\Big((1_{\{r_{t_{n-1}}<0\}} - 1_{\{r_{t_{n-1}}\geq 0\}})Z_{t_n}\Big).$$

où $1_A(x)$ est la fonction indicatrice de A (prend 1 et 0 respectivement si $x \in A$, $x \notin A$).
Il reste à poser que $a^1_{t_n} = \Big(1 + \Big(1_{\{r_{t_{n-1}}\geq 0\}} - 1_{\{r_{t_{n-1}}<0\}}\Big)Z_{t_n}\Big)$ et $a^2_{t_n} = \Big((1_{\{r_{t_{n-1}}<0\}} - 1_{\{r_{t_{n-1}}\geq 0\}})Z_{t_n}\Big)$ pour obtenir la structure auto-régressive à coefficient aléatoire pour le log des prix

$$\log S_{t_n} \;= a^1_{t_n} \log S_{t_{n-1}} \;+\; a^2_{t_n} \log S_{t_{n-2}} \;+\; \tilde{I}_{t_n} \tag{5.12}$$

avec $(\tilde{I}_t)$ un bruit blanc gaussien, voir (5.11) et l'assomption 1.

Les log des prix ne possèdent une racine unitaire qu'en moyenne. En effet, dans notre cas simplifié (N=1, $\omega_1 = 1$), on a

$$E(a^1_{t_n}) = E\Big[E(a^1_{t_n})|\mathcal{F}_{t_{n-1}}\Big] = E\Big[1 \;+\; \Big(1_{\{r_{t_{n-1}}\geq 0\}} \;-\; 1_{\{r_{t_{n-1}}<0\}}\Big)E(Z_{t_n}|\mathcal{F}_{t_{n-1}})\Big]$$

où $\mathcal{F}_{t_{n-1}} = \sigma\{S_0, \cdots, S_{t_{n-1}}\}$ est la filtration naturelle générée par les prix. Mais la variable aléatoire Z_{t_n} définie dans (5.11) est indépendante de $\mathcal{F}_{t_{n-1}}$. Donc $E(Z_{t_n}|\mathcal{F}_{t_{n-1}}) = E(Z_{t_n}) = 0$ puisque $E(Z^k_{t_n}) = 0$ pour tout k. Ce qui donne que

$$E(a^1_t) = E\Big[1 \;+\; \Big(1_{\{r_{t_{n-1}}\geq 0\}} \;-\; 1_{\{r_{t_{n-1}}<0\}}\Big)E(Z_{t_n})\Big) = 1$$

La même méthodologie montre aussi que $E(a^2_{t_n}) = 0$ et que notre structure RCA est précisément un modèle à racine unitaire stochastique (SUR). Dans le cas général, p^N_t avec $N > 1$, on a un modèle de la forme

$$\log S_{t_n} \;=\; \sum_{l=1}^{p} \log S_{t_{n-l}}\; X^l_{t_n} \;\; + I_{t_n} \tag{5.13}$$

où $p \geq 2$, (X^l_t) est tel que $E(X^1_t) = 1$; $\forall\, l \geq 2$ on a $E(X^l_t) = 0$ et $cov(X^l_t, I_t) = 0\ \forall t \in \mathbb{T}$.

Si on se résume, la description des événements donnant cette dynamique, est la suivante :
Au temps $t_{n-1} = t_n - h$, l'information fondamentale I_{t_n} est divulguée à tous les agents. La prochaine étape est l'interprétation de ce signal pour prédire les changements futurs des prix. Celle-ci aboutit pour l'agent k au signal $R^k_{t_n}$ après avoir intégré ses opinions personnelles $(\epsilon^k_{t_n})$. Il fait ensuite la décision suivante $n^k_{t_n} = c_k R^k_t$ puis l'excès de demande E_{t_n} est calculé. Enfin, le prix précédent est réajusté à l'instant t_n en utilisant l'équation (5.10). Nous donnons une représentation graphique du modèle à agents (voir figure 5.1) dans le cas simple où la volatilité courante est approximée par $|r_{t-h}|$ pour un certain $t \in \mathbb{T}$.

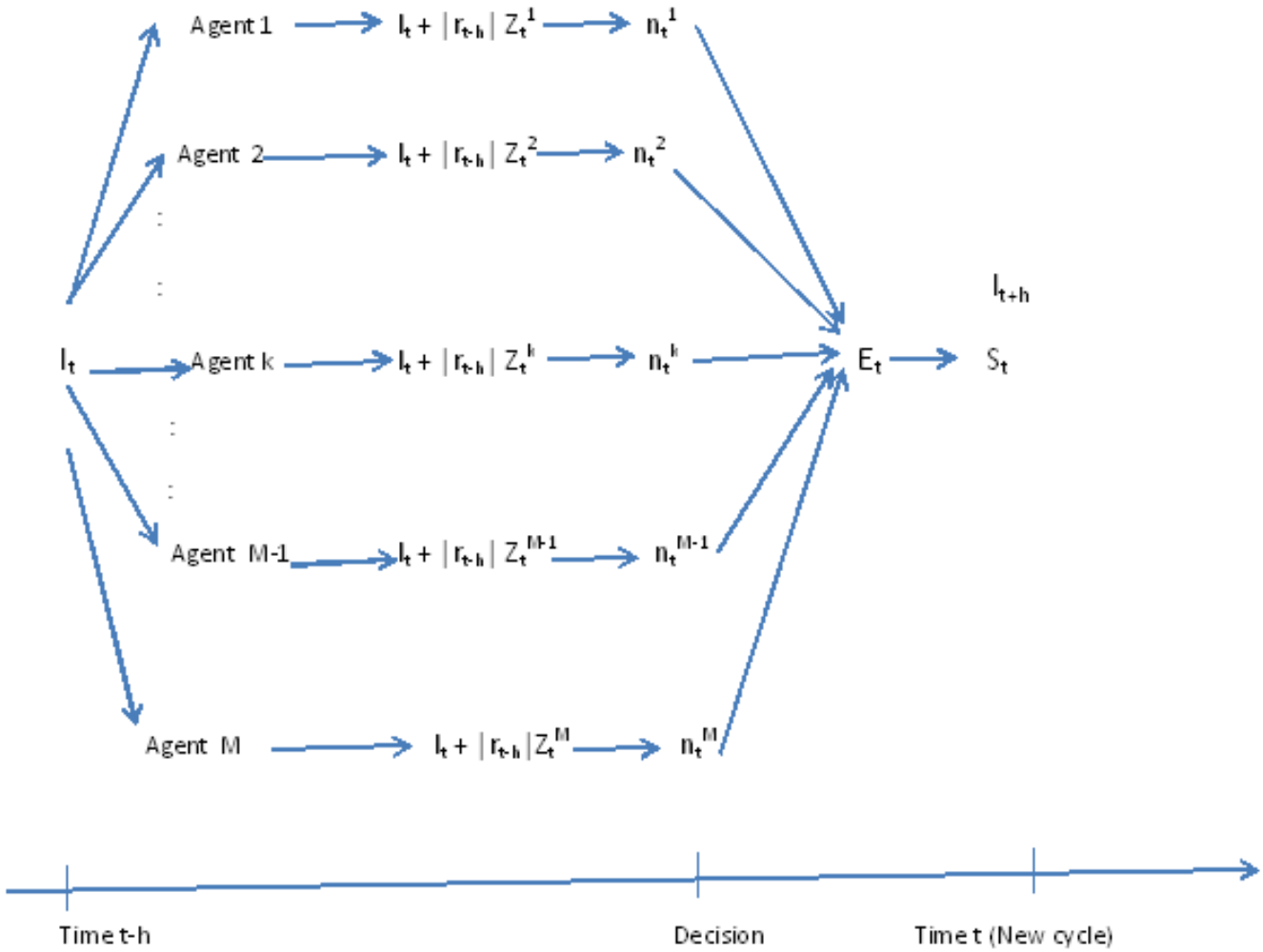

FIGURE 5.1 – Représentation graphique d'un cycle du modèle à agents entre $t-h$ et t, $t \in \mathbb{T}$.

time series	Mean	Variance	minimum	maximum	skewness	kurtosis	statistique JB	p
r_t	1,90 e-05	3,07e-02	-0,24	0,23	0,05	7,83	1,952 e+04	$<$ 0,001
$\|r_t\|$	2,24 e-02	2,12e-02	0	0,24	2,64	16,4	1,7469 e+05	$<$0,001

TABLE 5.1 – Quelques statistiques des séries (r_t) et $(|r_t|)$ basées sur l'équation (5.14).

Avant de faire des simulations, on précise trois faits qui sont dès fois reprochés aux modèles à agents. Le modèle utilise plusieurs paramètres en entrée. En effet tout individu $k = 1, \cdots, M$, est caractérisé par les paramètres σ_k intervenant dans le bruit idiosyncratique, c_k pour la décision, voir la figure 5.1. La modification de ces derniers ne change pas radicalement la structure des prix qui reste toujours de type $RCA(p)$. Seule la loi des variables aléatoires X_t^l, $l = 1, \cdots, p$ est changée. Une deuxième remarque est liée au nombre d'agents M qui peut dès fois rendre le système instable. Dans notre cas, ce paramètre ne perturbe pas aussi la dynamique des prix vu que les agents sont indépendants. On conserve la structure $RCA(p)$ avec une modification des lois de X_t^l ou du paramètre p. Un autre point est celui du choix du coefficient λ. On peut le modifier en prenant un réel positif quelconque. La structure $RCA(p)$ reste inchangée. Il n'y a que l'amplitude des rentabilités qui est modifiée.

Pour la simulation, on a utilisé directement l'équation (5.13) qui est issue de notre modèle à agents. Cela permet de gagner en termes de temps d'exécution. En effet, on n'a plus besoin de simuler le comportement de chaque agent au niveau micro pour faire ensuite une agrégation au niveau macro comme c'est le cas souvent dans la littérature. Ceci est un des avantages majeurs de notre modèle où les simulations ne durent que quelques secondes. Précisément, on considère le modèle le plus simple ($p_t^N = |r_{t_{n-1}}|$) correspondant au modèle $RCA(2)$ i.e :

$$\log S_{t_n} = \log S_{t_{n-1}} \Big(1 + u_{t_n}\Big) \quad + \quad \log S_{t_{n-2}} v_{t_n} \quad + \quad I_{t_n}. \tag{5.14}$$

Pour obtenir des valeurs réalistes de rentabilités, nous avons choisi de petites variances pour toutes les variables. On a supposé que $u_t \sim \mathcal{N}(0, (0.02)^2)$, $v_t \sim \mathcal{N}(0, (0.002)^2)$ et $I_t \sim \mathcal{N}(0, (0.02)^2)$ pour $t \in \{t_0, \cdots, t_T\}$. Les covariances entre les variables sont données par $cov(u_t, v_t) = 0.005^2$ et $cov(u_t, I_t) = cov(v_t, I_t) = 0$. Le nombre de pas de simulation est fixé à $T = 20000$ et les valeurs d'initialisation égales à $\log S_0 = \log S_h = 0.05$.

Dans la table 5.1 ci-dessus, nous avons donné les propriétés des séries $r_{t_n} = \log S_{t_n} - \log S_{t_{n-1}}$ et $|r_{t_n}| = |\log S_{t_n} - \log S_{t_{n-1}}|$.

Nous remarquons que le kurtosis de la série (r_t) est grand par rapport à la distribution normale. L'hypothèse nulle de normalité est rejetée avec une p-value plus petite que $0,001$ (Test de Jacque Bera). De même la moyenne et l'écart type de $|r_t|$ sont très proches. On retrouve donc les faits

stylisés mentionnés précédemment au niveau des séries (r_t) et $(|r_t|)$.

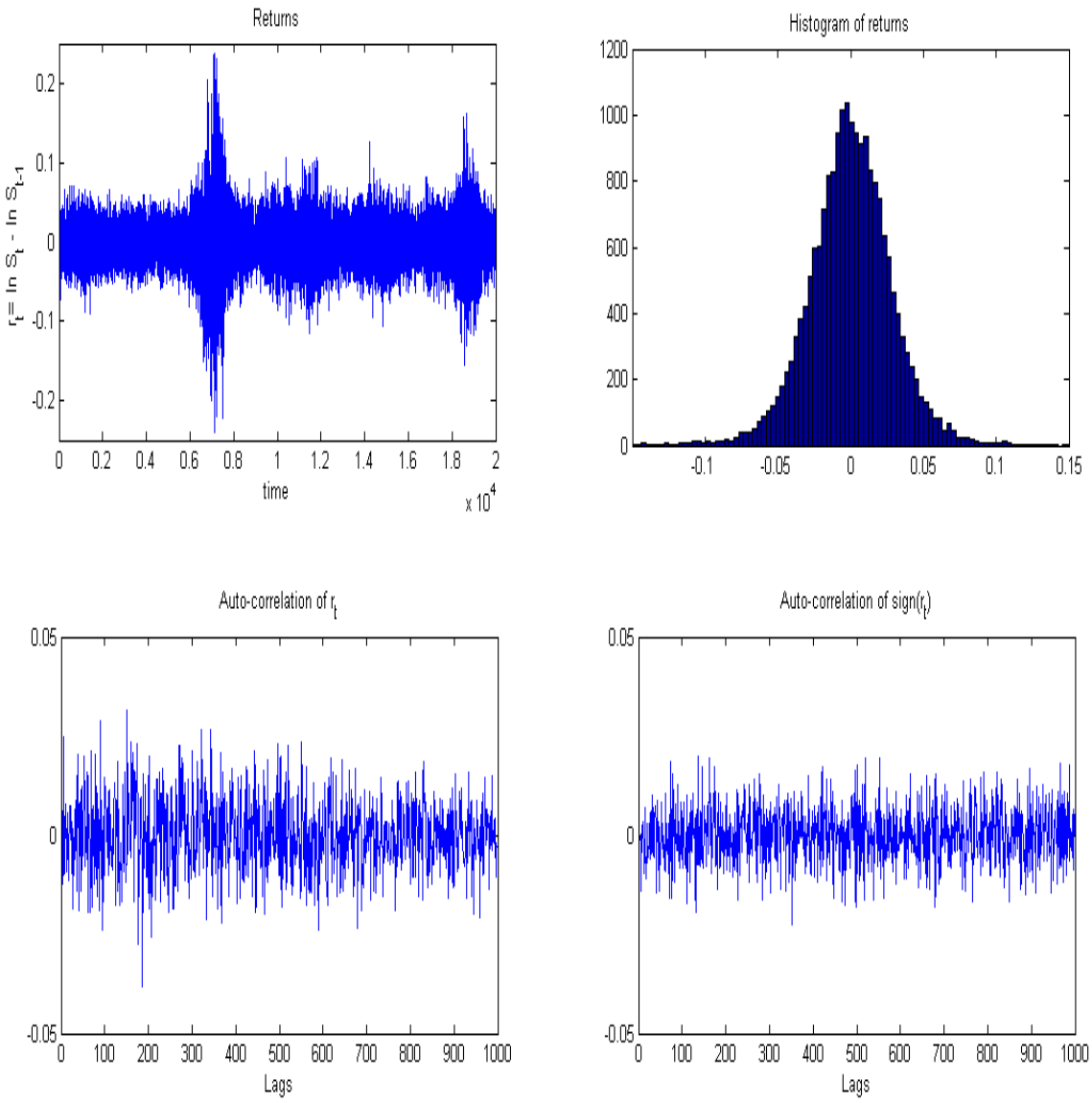

FIGURE 5.2 – Statistique de la série temporelle basée sur l'équation (5.14).

La figure 5.2 montre que les valeurs estimées des fonctions d'auto-corrélation (r_t) et sign(r_t) ne sont pas significatives confirmant la presque marche aléatoire des rentabilités. Un effet de regroupement de volatilité apparaît aussi dans la série des rentabilités.

La prochaine figure 5.3 montre les valeurs estimées des fonctions d'auto-corrélation de $|r_t|^q$ pour $q = 0,25;\ 0,5;\ 0,75;\ 1,\ 1,25;\ 1,5;\ 1,75$.

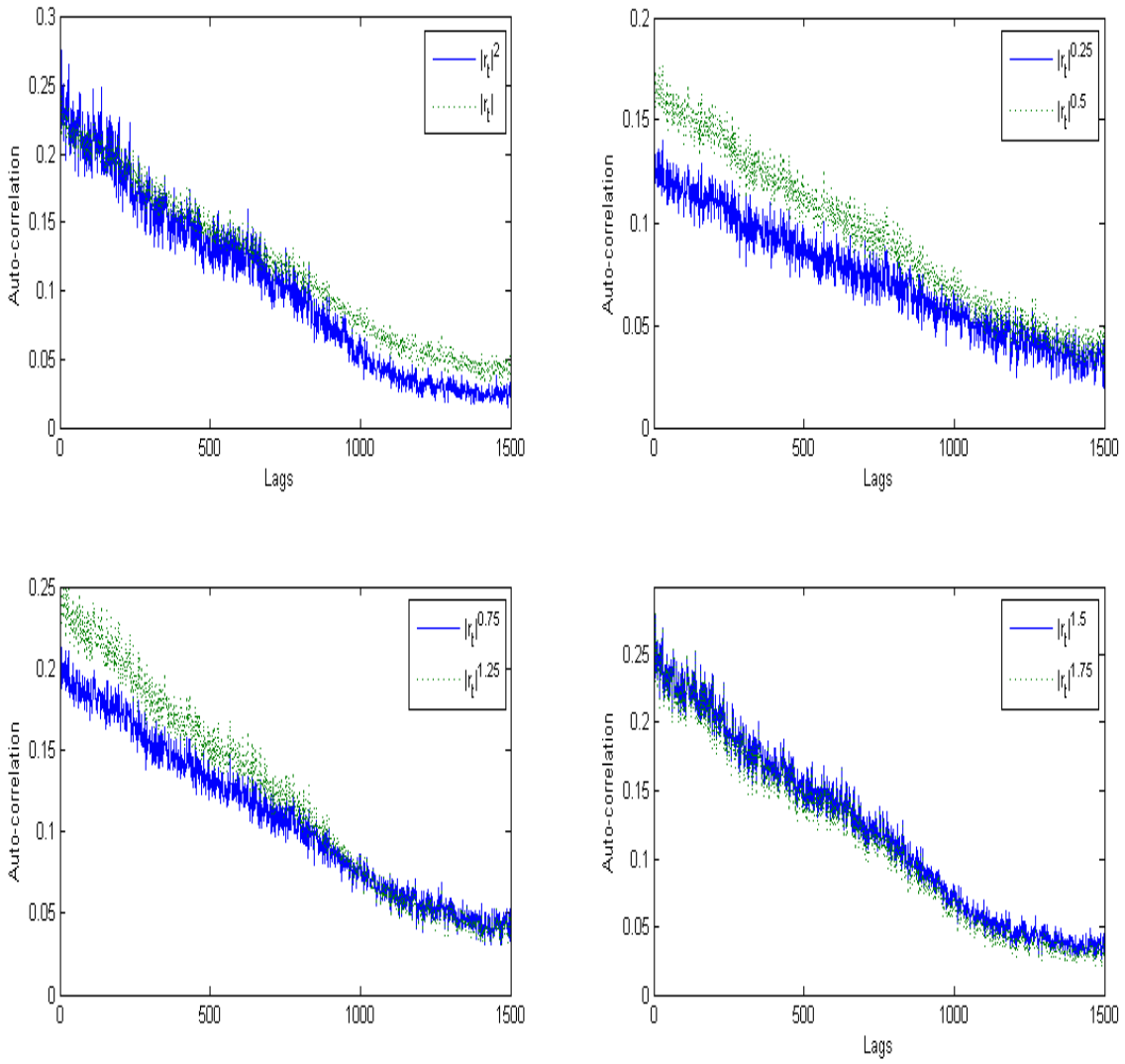

FIGURE 5.3 – Valeurs estimées des fonctions d'auto-corrélation de $|r_t|^q$, pour $q = 0,25;\ 0,5;\ 0,75;\ 1;\ 1,25;\ 1,5;\ 1,75$, basées sur (5.14).

On remarque que toutes ces courbes décroissent vers zéro avec une certaine persistance confirmant l'effet regroupement de volatilité. On note, comme dans les séries de données d'indices, que la persistance est plus prononcée pour $q = 1$ que pour $q = 2$. Si on considère toutes les courbes, la persistance est plus marquée pour $q = 0.25$. Un tel fait a été noté généralement dans les séries de taux de change.

Dans la prochaine étape, on regarde la robustesse du modèle. On a vu que modifier les paramètres d'entrée revenaient à changer les lois de X_t^l et peut être du paramètre p. On a testé plusieurs variantes du modèle $RCA(p)$ consistant à modifier p et la matrice de variance-covariance associée aux différentes variables. Les résultats sont presque les mêmes et reproduisent tous les faits stylisés. On a pris un seul exemple en guise d'illustration. Il s'agit d'un modèle $RCA(3)$ de la forme :

$$\log S_{t_n} \;=\; \sum_{l=1}^{3} \log S_{t_{n-l}}\, X_{t_n}^l \;+\; I_{t_n}, \quad p = 3 \tag{5.15}$$

Ici, $[I_{t_n}\ X_{t_n}^1\ X_{t_n}^2\ X_{t_n}^3]'$ suit une distribution multivariée $\mathcal{N}(0, A)$ où

$$A = \begin{pmatrix} 0,02^2 & 0 & 0 & 0 \\ 0 & 0,01^2 & 0,004^2 & 0,003^2 \\ 0 & 0,004^2 & 0,005^2 & 0,002^2 \\ 0 & 0,003^2 & 0,002^2 & 0,03^2 \end{pmatrix}$$

Les résultats des simulations sont représentés dans les figures 5.4 et 5.5 où on a pris cette fois $T = 10000$. On retrouve les mêmes propriétés obtenues dans les figures 5.2 et 5.3.

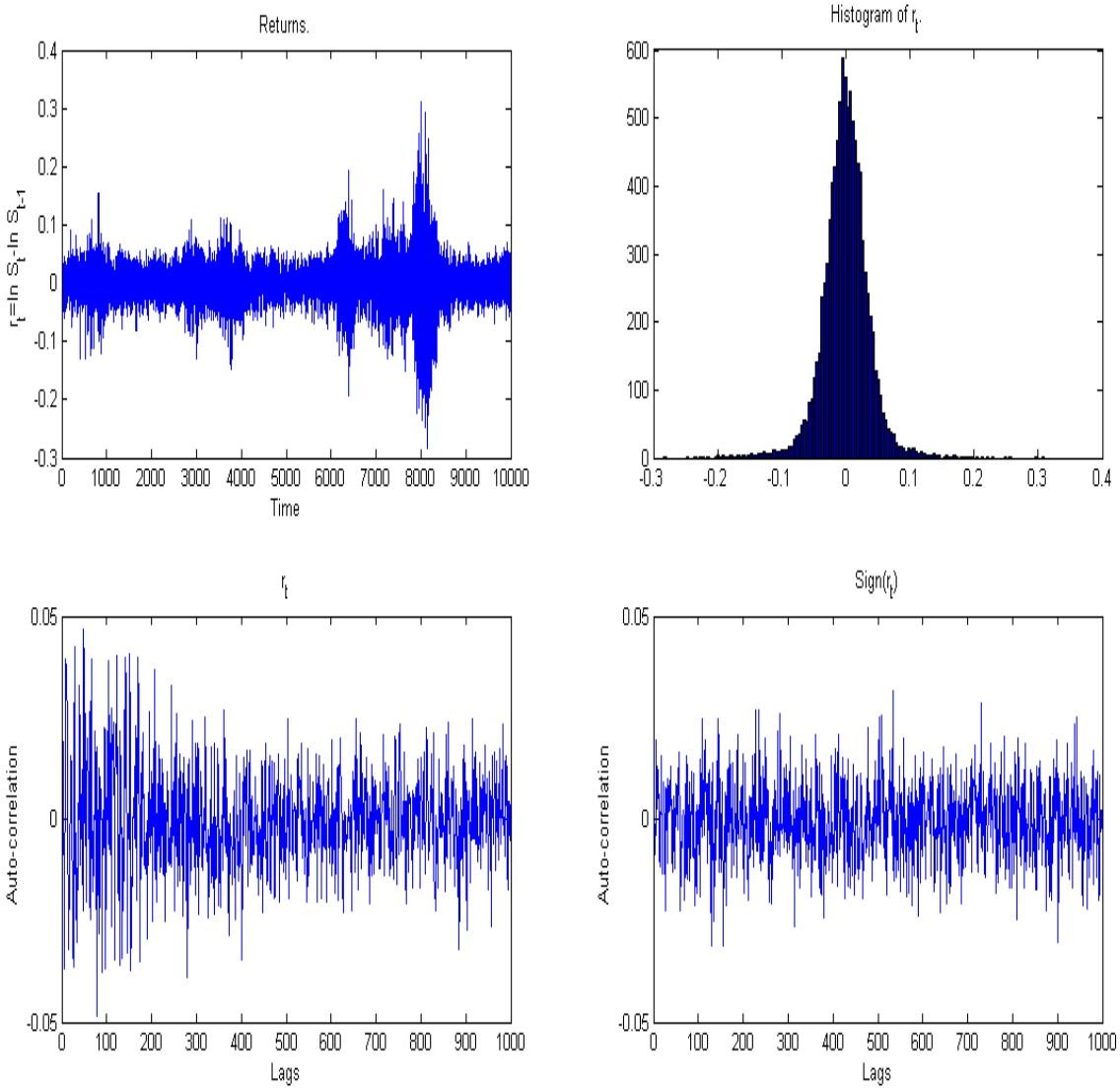

FIGURE 5.4 – Propriétés de la série temporelle (r_t) basées sur l'équation (5.15).

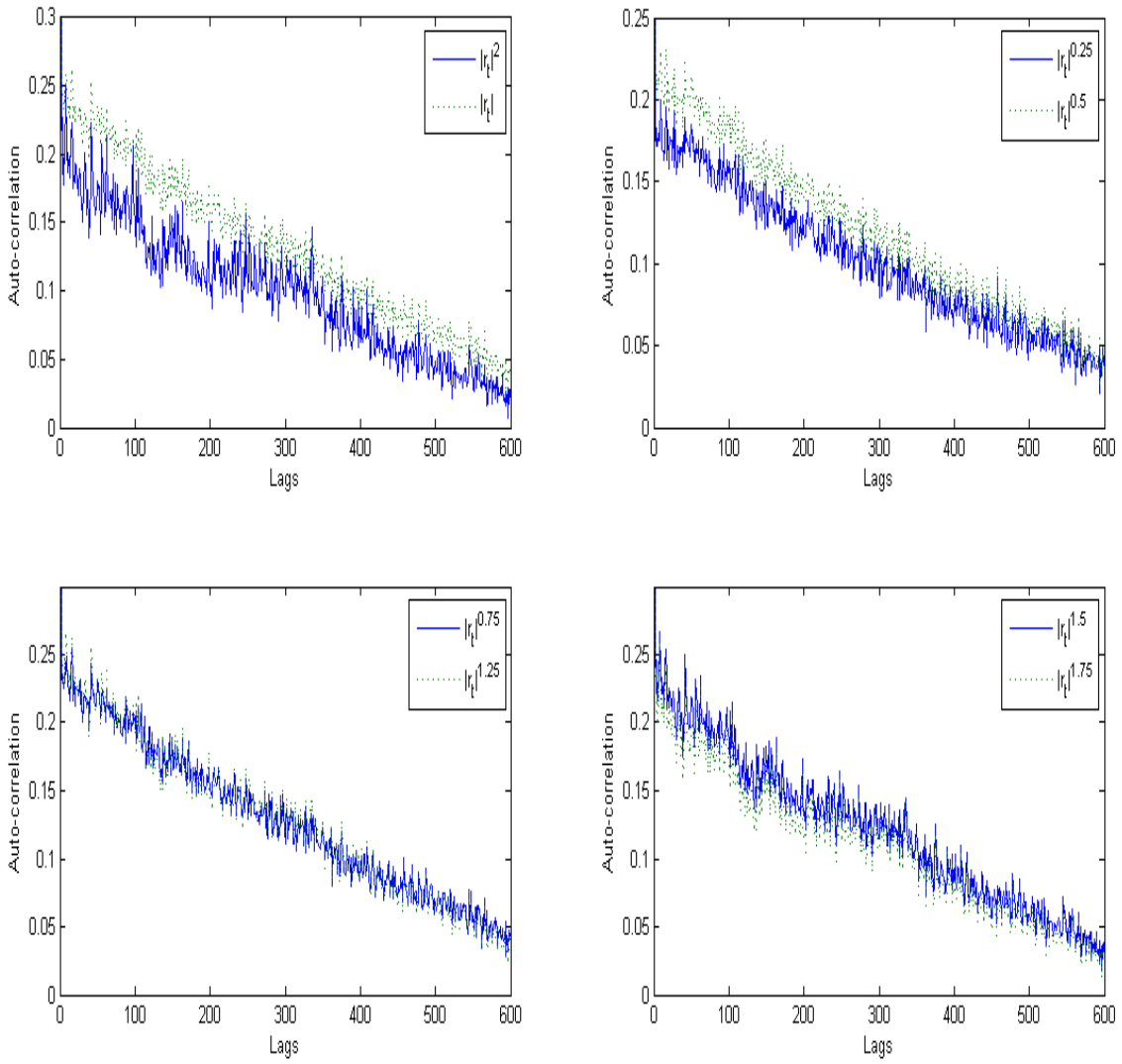

FIGURE 5.5 – Valeurs estimées des fonctions d'auto-corrélation de $|r_t|^q$, pour $q = 0,25;\ 0,5;\ 0,75;\ 1;\ 1,25;\ 1,5;\ 1,75$, basées sur (5.15).

Time series	Mean	Variance	minimum	maximum	skewness	kurtosis	JB statistic	p
r_t	2,03 e-04	3,75 e-02	-0,28	0,31	-7,20 e-02	8,84	14.265	<0,001
$\|r_t\|$	2,64 e-02	2,66 e-02	0	0,31	2,86	16,5	90.424	<0,001

TABLE 5.2 – Statistique des séries (r_t) et $(|r_t|)$ basée sur l'équation (5.15).

Dans la table 5.2 ci- dessus, on a représenté les statistiques des séries (r_t) et $(|r_t|)$. Nous remarquons que les résultats sont similaires aussi à la table 5.1.

5.2 Extension et Commentaires.

Notre modèle peut être relié à des processus ayant un comportement similaire au modèle ARCH proposé par Engle [35]. En effet, l'équation (5.10)devient à l'instant t_n

$$r_{t_n} = \ln S_{t_n} - \ln S_{t_{n-1}} = \frac{E_{t_n}}{\lambda}$$

où $\frac{E_{t_n}}{\lambda} \sim \mathcal{N}(0,\ c^2\sigma_I^2 + \frac{(p_{t_n}^N)^2}{\lambda^2}\sum_{i=1}^M c_k^2\sigma_k^2)$, voir Assomption 1 et l'équation (5.11).

Donc la dynamique de (r_t) est donnée par

$$\begin{cases} r_{t_n} &= \phi_{t_n} Y_{t_n} \\ \phi_{t_n}^2 &= c^2\sigma_I^2 + \frac{(p_{t_n}^N)^2}{\lambda^2}\sum_{i=1}^M c_k^2\sigma_k^2 \end{cases} \tag{5.16}$$

où les variables aléatoires $Y_t, t \in \mathbb{T}$ sont i.i.d avec $Y_t \sim \mathcal{N}(0,1)$.

Si $N = 1$, on a $(p_{t_n}^N)^2 = r_{t_{n-1}}^2$ et évidemment (r_t) est un processus de type ARCH(1). Pour le cas général $N > 1$, la dynamique r_t se comporte comme un processus de type ARCH(N) avec cependant des termes additionnels au niveau de la variance conditionnelle. Ces termes sont précisément des doubles produits de la forme $|r_{t_{n-i}}||r_{t_{n-j}}|$ $1 < i < j \leq N$, voir (5.2).

Sato et Takayasu [102] ont aussi dérivé un processus ARCH(1) à partir d'un modèle à agents virtuels avec un cadre différent. Dans leur modèle, les agents utilisent des règles déterministes tandis qu'ici elles sont aléatoires. Cependant les deux modèles partagent la même idée qui est

que la décision des agents dépend des dernières rentabilités.
Notre modèle est relié aussi à celui de Takayasu et Takayasu [108] ou de Ghoulie et al. [43] sur comment les équations des prix sont obtenues. Précisément, la fonction d'impact sur les prix est supposée log linéaire. Si on travaille sur des données intra-journalières où $h = t_i - t_{i-1}$ est petit, on peut espérer que l'excès de demande E_t soit proche de zéro et que toute fonction non linéaire aura une approximation acceptable par rapport à un développement limité d'ordre 1 de Taylor.

La seconde extension est liée à la loi normale. On a utilisé que sa propriété de stabilité par rapport à une somme de variables aléatoires indépendantes. Ce qui veut dire que la généralisation peut se faire sur l'ensemble des variables aléatoires symétriques stables de loi L. Dans ce cas, on restera toujours dans la classe RCA(p) avec un changement de la normale remplacée par celle de L. En d'autres termes, on aura une équation de type (5.16) où la variable aléatoire Y_{t_n} est maintenant de loi L.

Pour les commentaires, Yoon [113] est l'un des premiers à remarquer que les modèles ayant une racine unitaire en espérance (Stochastic Unit Root models) avait la capacité de reproduire les principales propriétés des séries financières. Cependant aucune explication n'est donnée de l'élément fondamental qui est responsable de ces faits stylisés. Vu que le modèle économétrique SUR a été relié à un modèle à agents virtuels, cet élément peut être investi. Si on suppose que les agents ont tous une rationalité illimitée, alors ils prédisent toujours le signal I_{t_n} quelque soit l'état du marché. Dans ce cas, leur décision serait de la forme $c_k I_{t_n}$ et donc les prix des actifs vont suivre une marche aléatoire. Par suite, il n'y aurait pas d'effet regroupement de volatilité. Cette hypothèse est assez contraignante. On peut la relâcher en supposant que les agents évaluent l'information fondamentale en faisant des erreurs aléatoires indépendamment de l'état du marché. Dans ce cas, leur décision peut être modélisée par le nouveau signal noté $\tilde{R}^k_{t_n} = I_{t_n} + \sigma_k Z^k_{t_n}$. L'équation (5.10) devient alors :

$$\log S_{t_n} = \log S_{t_{n-1}} + \frac{\tilde{E}_{t_n}}{\lambda} = \log S_{t_{n-1}} + \tilde{I}_{t_n} + Z_{t_n} = \log S_{t_{n-1}} + \gamma_{t_n} \tag{5.17}$$

où $\tilde{E}_{t_n} = \sum_{k=1}^{M} c_k \tilde{R}^k_{t_n}$ et $(\gamma_{t_n}) = (\tilde{I}_{t_n} + Z_{t_n})$, $t_n \in \mathbb{T}$ est un bruit blanc gaussien grâce à l'assomption 1. La notation $\tilde{I}_t$ est la même que celle définie dans l'équation (5.11)

On retombe donc sur une marche aléatoire dans l'équation (5.17). Le seul moyen pour générer ces faits stylisés est d'introduire un effet feedback par rapport aux prix passés. C'est ce que nous avons illustré ci-dessus en montrant qu'un modèle composé d'agents indépendants ayant des comportements rétrospectifs ($R^k_{t_n} = I_{t_n} + \sigma_k p^N_{t_n} Z^k_{t_n}$) est suffisant pour reproduire les principales propriétés des séries financières. Dans ce cas, la dynamique des prix d'équilibre appartient à la classe $RCA(p)$. Cette dernière ne peut être attribuée simplement aux arguments de la finance comportementale ou à ceux de la finance néo-classique. Elle est obtenue par une interaction des deux arguments. Les investisseurs suivent souvent le signe de l'information fondamentale I_{t_n} (théorie de l'efficience) tout en faisant en même temps des extrapolations (biais psychologiques). Le modèle peut donc jouer un rôle réconciliateur dans le cadre discret.

Troisième Partie : Applications.

On a établi deux cadres convenables pour modéliser la dynamique des prix. Le terme convenable est relatif au fait d'intégrer à la fois les arguments de la Finance néo-classique et ceux de la Finance comportementale. Dans le cadre continu, on avait trouvé les modèles de diffusion avec une partie saut d'espérance nulle. Dans le cadre discret, on s'est retrouvé avec les modèles RCA(p) avec une racine unitaire en moyenne.
Par ailleurs, on a vu au chapitre 2 que les corrélations linéaires des rentabilités financières étaient très faibles. C'est pourquoi, si on veut prédire la direction des prix à partir de prix passés, il serait commode d'utiliser des modèles non linéaires. C'est en ce sens que la première application sera dédiée à l'utilisation du modèle non linéaire RCA(1) pour des fins d'investissements.
Dans la deuxième application, au lieu de prédire la direction des prix ou des rentabilités, on s'intéresse plutôt à l'aspect volatilité qui est un enjeu majeur aussi pour tout investisseur. Précisément pour tenir en compte la complexité de la dynamique des prix d'actifs financiers, on a préféré généraliser le modèle standard GARCH en rendant ses coefficients stochastiques.

Chapitre 6

Stratégies d'investissement.

Dans le chapitre précédent, on a fait des simulations pour valider le modèle économétrique RCA(p). Même si ce point est une phase importante, ce qui intéresse le plus les opérateurs économiques, c'est de pouvoir proposer des stratégies gagnantes. C'est pourquoi, notre première application sera d'utiliser le modèle RCA(1) pour voir si on peut battre la performance du marché.

6.1 Description de la méthodologie

Nous considérons le modèle RCA(1) suivant :

$$y_t \;=\; (1+b_t)y_{t-1} \;+\; \epsilon_t \tag{6.1}$$

$$E(b_t) \;=\; E(\epsilon_t)=0, \quad E(b_t^2)=\omega^2, \quad E(\epsilon_t^2)=\sigma^2, \quad cov(b_t,\epsilon_t)=0$$

où (ϵ_t) est un bruit blanc gaussien et $y_t = \log S_t$.

Cette formulation est motivée par la dynamique trouvée dans le chapitre précédent, voir l'équation (5.13), où les variables b_t et ϵ_t n'étaient pas corrélées. Le modèle général suivant est souvent utilisé

$$y_t \;=\; (\phi+b_t)y_{t-1} \;+\; \epsilon_t$$
$$E(b_t) \;=\; E(\epsilon_t)=0, \quad E(b_t^2)=\omega^2, \quad E(\epsilon_t^2)=\sigma^2, \quad cov(b_t,\epsilon_t)=\psi\omega$$

où il est supposé que $\phi^2+\omega^2<1$. Dans ce cas, le processus (y_t) est stationnaire, voir Nicholls et Quinn [90]. L'objectif de ces études et de tester l'hypothèse nulle d'un modèle auto-régressif(AR)

contre un modèle AR avec des coefficients aléatoires.

$$H_0 : \quad \omega^2 \;=\; 0 \; .$$

Le problème a été aussi étudié dans le cadre non stationnaire qui nous intéresse ici vu que $\phi = 1$. On peut voir, par exemple, Nagakura [88] et les références qui y sont citées. On va mettre plus l'accent ici sur la prévision de la direction des prix qui est un enjeu majeur pour les opérateurs économiques. La difficulté dans le cas où $\phi = 1$ est que l'on obtient une série qui n'admet pas de moments finis. L'explication vient du fait que la distribution asymptotique des queues est de la forme puissance avec un exposant égal à 1 i.e

$$P(|y_t| > x) \;\approx\; cx^{-1}$$

où c et x sont deux constantes avec x assez grand.

Il serait donc commode d'utiliser un nombre assez petit de données pour faire de l'estimation sinon on risque de voir les deux premiers moments ne pas exister. C'est ce que nous faisons en utilisant toujours les données les plus récentes pour estimer le paramètre b_t. Bien entendu celui-ci est une variable aléatoire et son estimation devrait être plutôt non-paramétrique. Pour que le problème d'estimation soit faisable, on rend le problème paramétrique. Cette approche a été utilisée par Granger et Swasson [48] qui expriment le coefficient $1 + b_t$ sous la forme

$$1 + b_t \;=\; \exp(\alpha_t); \qquad \alpha_t \; i.i.d \;\sim\; N(m, \sigma_\alpha^2)$$

$$\alpha_t = \mu \;+\; \rho\alpha_{t-1} + \eta_t$$

où (η_t) est un processus gaussien ($\eta_t \sim N(0, \sigma_\eta)$) qui est indépendant de (ϵ_t) défini dans l'équation (6.1). La contrainte $E(b_t) = 0$ est matérialisée par $E(\exp(\alpha_t)) = 1$ équivalant à dire que $m + \frac{1}{2}\sigma_\alpha^2 = 0$. Ils utilisent des approximations de la fonction de vraisemblance pour estimer les paramètres $\mu,\ \rho, \sigma_\eta$.

Sollis et al. [106] utilisent, quant à eux, une version paramétrique directement sur b_t de la forme

$$b_t = \rho b_{t-1} + \eta_t$$

qui satisfait bien la condition $E(b_t) = 0$. Le paramètre ρ est estimé en utilisant les méthodes de filtre de Kalman.

Les deux approches ci-dessus utilisent un nombre consistant de données pour l'estimation des paramètres. Ici, nous proposons une méthode d'estimation locale de b_t n'utilisant que les données récentes. La motivation vient du fait qu'on a établi une relation entre les modèles RCA(p) et certains modèles à agents avec phénomène de switching. Vu sous cet angle de dynamiques complexes, on a besoin de méthodologies locales pour capter l'évolution du système. Précisément, on suppose que b_t suit une loi binomiale prenant les valeurs α et $-\alpha$ avec une même probabilité égale à $\frac{1}{2}$. On peut récrire b_t comme suit $b_t = \alpha X_{t-1}$ où pour tout t, $P(X_t = 1) = P(X_t = -1) = 0,5$. On remarque que le terme b_t a été pris comme variable prévisible et cela permet de pouvoir faire de la prévision. Ici, on a un décalage d'un pas comme dans les modèles ci-dessus.
La variance inconditionnelle de b_t est donnée par α^2. On suppose par contre que sa variance conditionnelle bouge dans le temps à l'instar des modèles ARCH, voir Engle [35]. A l'instant t, pour estimer le paramètre α_t, on utilise l'approche des moindres carrées. L'équation (6.1) devient :

$$r_t = \log S_t - \log S_{t-1} = y_t - y_{t-1} = c + \alpha X_{t-1} y_{t-1} + \epsilon_t, \tag{6.2}$$

où on a ajouté une constante c, comme il est d'usage, dans les régressions linéaires.

Pour le moment, on ne peut appliquer directement la méthode car la variable X_{t-1} n'est pas connue (variable binomiale). On peut contourner ce problème en faisant une régression conditionnellement à la trajectoire du processus (X_t). Par exemple, si dans la régression (6.2), on choisit une longueur de n données pour faire l'estimation, on aura 2^n trajectoires possibles car chaque X_t ne prend que deux valeurs -1 ou 1. Pour chaque trajectoire donnée, on peut associer un modèle linéaire de régression composée d'une variable explicative $(X_{t-1}y_{t-1})$ plus la variable constante. On aura donc au total 2^n modèles.[23] Pour une régression linéaire, on a besoin que le paramètre n soit grand pour une précision des estimations mais pas trop dans notre cas pour éviter de perdre l'aspect feedback.

Ci- dessous, on a représenté pour n fixé le nombre total de modèles linéaires disponibles générés par les différentes valeurs possibles de (X_t). On remarquera que, dans ce tableau, les modèles de numéro j et $2^n - j + 1$ fournissent les mêmes résultats pour $j = 1, \cdots, 2^{n-1}$.

23. Réellement on a $2^{n-1} = \frac{2^n}{2}$ modèles car chacun se reproduisant deux fois. En effet, on ne peut distinguer le modèle caractérisé par $\{\alpha_t, (X_{t-1}, \cdots, X_{t-n})\}$ du modèle suivant $\{-\alpha_t, (-X_{t-1}, \cdots, -X_{t-n})\}$.

Modèle 1	Modèle 2	$\cdots$	$\cdots$	Modèle $2^n - 1$	Modèle 2^n
-1	-1			1	1
-1	-1			1	1
-1	-1			1	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
-1	-1			1	1
-1	-1			1	1
-1	1	$\cdots$	$\cdots$	-1	1

TABLE 6.1 – Caractérisation de chaque modèle linéaire à travers les valeurs binaires de (X_t).

Faire 2^n régressions peut être coûteux en temps pour n assez grand et contraignant aussi pour stocker la matrice représentée ci-dessus.[24] De plus, utiliser tous les modèles diminue la performance des stratégies dû à la corrélation qui existe entre les modèles. C'est pourquoi, on choisit de prendre parmi les 2^n modèles disponibles que n modèles non corrélés. Ce qui revient à choisir une base orthogonale de $\mathbb{R}^n$ formé que de vecteurs colonnes avec des valeurs -1 et 1. Pour ce faire, l'entier n doit être tel que $n/2$, $n/12$ ou $n/20$ soit de la forme 2^k, $k \in \mathbb{N}$. Dans ce cas, la base orthogonale binaire peut être obtenue en utilisant la fonction matlab *hadamard*. Par exemple pour $n = 2$, la matrice de Hadamard est donnée par

$$H_2 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Le produit scalaire de ces deux vecteurs colonnes vaut bien zéro. Puis de façon récursive, on peut calculer la matrice H_4, H_8, $\cdots$, en utilisant la formule suivante :

$$H_{2n} = \begin{bmatrix} H_n & -H_n \\ H_n & H_n \end{bmatrix}$$

Cette approche de base orthogonale binaire permet d'obtenir une complexité linéaire par rapport au paramètre n tout en améliorant la performance.[25] Pour chacun de ces n modèles linéaires (donc la trajectoire (X_t) connue), la direction de hausse ou de baisse est déterminée en estimant l'équation (6.2) par la méthode des moindres carrés puis une prévision de la variation des prix est faite à l'instant $t + 1$ par

24. Pour n dépassant la vingtaine, on a obtenu des débordements de capacité de mémoire. La taille de la matrice est de l'ordre $n \times 2^n$.

25. On a remarqué dans l'étude empirique que prendre un nombre élevé de modèles linéaires ne tendaient pas à augmenter les performances des stratégies. Dans la pratique, on a besoin de suivre que quelques modèles représentatifs de l'ensemble $\{1, \cdots, 2^n\}$.

$$\hat{r}_{t+1} = \hat{y}_{t+1} - y_t = \hat{c} + \hat{\alpha}_t \hat{X}_t y_t. \tag{6.3}$$

On décide de tester la théorie de l'efficience en proposant des stratégies que l'on évaluera par rapport à celle consistant à ne détenir que l'indice (Achat et Conservation). Chaque stratégie est caractérisée par deux paramètres n et m où ces entiers matérialisent le nombre de données utilisées dans la régression (6.2). La stratégie (m, n) consistera d'abord à choisir m (n) modèles parmi les 2^{m-1} (2^{n-1}) modèles linéaires différents. Parmi ces m (n) modèles linéaires, on prendra le meilleur en termes de variance résiduelle minimale. On sera dans une position longue (courte) si le meilleur modèle basé sur un échantillon de m données ainsi que celui basé sur un échantillon de n données prédisent tous une baisse (hausse) des prix. En d'autres termes, la stratégie (m, n) essaie de capter les effets revirements plutôt que des effets tendances. Pourquoi cette approche contrarienne ? La première raison est qu'on travaillera sur des entiers petits pour les paramètres m et n et donc capter un effet tendance reste difficile. La deuxième raison est que la variable expliquée, ici r_t, est proche d'une martingale ou d'une marche aléatoire. Vu que l'on travaille sur le court terme, il est préférable de miser sur un effet de retour vers sa moyenne et donc d'utiliser une stratégie contrarienne. En résumé, pour la stratégie (m, n) on regarde la prédiction des deux meilleurs modèles obtenus pour la classe des régressions avec m données (dénoté mod_1) et la classe des régressions avec n données (dénoté mod_2). Si la tendance prédite par ces meilleurs modèles mod_1 et mod_2 est une hausse (baisse) alors la décision prise sera une vente (achat) de l'actif. Dans le cas contraire (prédiction différente entre les deux meilleurs modèles), on ne fait rien et on conserve notre position précédente.

Pour une stratégie (m, n), on introduit la variable $position_{m,n}(t)$ qui contient la position de l'agent à l'instant t. Cette variable est définie d'après les explications ci-dessus par

$$position_{m,n}(t) = \begin{cases} -sign(\hat{r}^{(1)}_{t+1}) & si\ sign(\hat{r}^{(1)}_{t+1}) = sign(\hat{r}^{(2)}_{t+1}) \\ position_{m,n}(t-1) & sinon. \end{cases}$$

La fonction sign est ici définie par $sign(x) = 1$ si $x > 0$, $sign(x) = -1$ si $x < 0$ et $sign(0) = 0$. La notation $\hat{r}^{(j)}_{t+1}$ signifie la prévision faite par le meilleur modèle mod_j où $j = 1, 2$.

On a testé la performance de cet algorithme sur l'indice CAC 40 et sur l'indice $S\&P$ 500. Ce dernier est un indice de référence pour des études liées à l'efficience de marché. Pour les données journalières du S&P 500, l'échantillon des données va du 2 Janvier 2002 au 31 Décembre 2010 correspondant à un nombre total de données égal à $N = 2267$. Pour le CAC 40, on a pris également un nombre de données égale à celui du $S\&P$. L'échantillon va précisément du 25 Février 2002 au 31 Décembre 2010.[26] Nous mesurons la performance à travers le ratio de Sharpe (RS) qui est égale au rapport de la rentabilité sur l'écart type associé. Si on note $diff_t = S_t - S_{t-1}$, la

26. Le décalage au niveau de la date de début s'explique par le fait que l'échantillon du CAC 40 comporte plus de données que celui de l'indice $S\&P$ 500 sur la période 2002-2010.

variation du prix de l'indice à l'instant t alors pour un nombre de prévisions N fait, la stratégie (m, n) fournit les valeurs suivantes :

$$RM(m,n) \quad = \quad \frac{1}{N}\sum_{t=1}^{N} z_t, \quad \text{où} \quad z_t = position_{m,n}(t) \cdot diff_{t+1}. \tag{6.4}$$

$$RS(m,n) \;=\; \frac{RMoy(n,k)}{\sigma(m,n)}, \quad \text{où} \quad \sigma(m,n) = \left[\frac{1}{N}\sum_{t=1}^{N}\Big(z_t - RM(m,n)\Big)^2\right]^{0,5}. \tag{6.5}$$

La performance de l'indice quant à elle est mesurée par la stratégie Achat et Conservation qui correspond à acheter l'indice et ne rien faire après. En d'autres termes, on est toujours dans une position longue i.e position(t)=1, $\forall\ t$.

Vu que le modèle économétrique RCA(p) a été dérivé à travers une hypothèse de feedback des agents sur les prix passées des titres, les paramètres m et n sont contraints de ne pas être des entiers grands. De plus, ces entiers doivent être de sorte qu'une base de Hadamard puisse exister. Ceci nous mène à proposer seize (16) stratégies basées sur les paramètres (m, n) avec $m \in \{4, 8, 12, 16\}$ et $n \in \{20, 24, 32, 40\}$. Il serait intéressant de comparer ces stratégies par rapport à celles basées sur des techniques graphiques. Vu que la littérature dans ce domaine est assez vaste, on va se limiter dans cette étude, aux techniques de moyennes mobiles. Ces dernières consistent à observer les signaux générés par le croisement d'une moyenne mobile court-terme et d'une moyenne mobile long-terme. Les moyennes mobiles considérées dans cette étude sont de deux types : moyenne simple (pondération équitable) et moyenne exponentielle (les données récentes ont plus de poids que celles lointaines). Dans les deux cas, la stratégie consiste à acheter (vendre) lorsque la moyenne mobile court-terme (long-terme) traverse à la hausse (à la baisse) la moyenne mobile long-terme (court-terme). Dans chacun des deux cas, 16 stratégies ont également été proposées qui sont paramétrées par deux entiers m et n avec $m \in \{5, 10, 15, 20\}$ et $n \in \{25, 50, 75, 100\}$. La moyenne sur les m (n) dernières données correspondent à la moyenne mobile court-terme (long terme). Pour plus de détails sur les techniques chartistes, on réfère le papier de Sullivan, Timmermann, White [107] et les références qui y sont citées.

6.2 Résultats

Dans le tableau ci-dessous, on montre les résultats des différents algorithmes sur les indices CAC 40 et S&P. L'échantillon (in sample) contient les trois quart de l'intervalle d'étude à savoir $1700 = \ 2267 * 3/4$ données. Le dernier quart restant (567 données) est réservé comme hors échantillon. Cette décomposition est un peu standard où beaucoup de données sont réservées à la partie apprentissage (in sample). Précisément, on a pris à peu près les six premières années pour trouver les meilleurs modèles pour chaque catégorie et deux années pour les évaluer hors échantillon. Ici, on en a trois catégories (classe RCA(1), moyenne mobile simple, moyenne mobile exponentielle) en plus de la stratégie qui ne consiste à ne rien faire (Achat et Conservation). Nous commençons, d'abord, par donner la performance de toutes les stratégies considérées dans la phase apprentissage pour l'indice français. Pour les données hors échantillon, on utilisera que les meilleurs modèles dans chaque catégorie.

TABLE 6.2 – Ratio de Sharpe des différentes stratégies sur le CAC 40.

In sample (Fev 2002, Oct 2008)				
RCA(1)				
Stratégie(m,n)	20	24	32	40
4	0,05415922	0,03397293	0,02075063	0,05308532
8	0,04039933	0,01940447	0,00013559	0,04575169
12	0,067998	0,03347935	0,01563643	0,05871787
16	0,03359021	0,0401175	0,01148712	0,03945495
Achat/Conservation			-0,0094	
Moyenne Mobile simple				
Stratégie(m,n)	50	100	150	200
5	0,01186702	0,01152782	0,02466787	0,01152782
10	0,01838265	0,01403258	0,03083357	0,01403258
15	0,0373459	0,02241777	0,0380776	0,02241777
20	0,02224289	0,00916671	0,02523294	0,00916671
Moyenne mobile exponentielle				
Stratégie(m,n)	50	100	150	200
5	0,02902561	0,02778984	0,04052782	0,02778984
10	0,03114447	0,04271336	0,045677	0,04271336
15	0,02910525	0,04723122	0,0389833	0,04723122
20	0,0339175	0,0546568	0,04576496	0,0546568

Pour les stratégies basées sur le modèle économétrique RCA(1), il est à noter qu'elles fournissent toutes une rentabilité positive matérialisée par un ratio de Sharpe positif. La même remarque s'applique aussi pour les stratégies de trading de type moyenne mobile simple et exponentielle. De plus, utiliser une pondération exponentielle (à savoir donner plus de poids aux données récentes par rapport aux données lointaines) offre des ratios plus élevées par rapport à une pondération équitable (moyenne mobile simple). La stratégie Achat et Conservation a fourni quant à elle un ratio de Sharpe négatif égale à $-0,0094$. On note par ailleurs que la meilleure

stratégie se trouve dans la catégorie RCA(1) avec les paramètres suivants $m = 12$, $n = 20$. Son ratio de Sharpe égal à $0,068$. Dans la classe des stratégies de moyenne mobile simple, la meilleure stratégie est celle consistant à prendre les paramètres suivants $m = 15$, $n = 150$ (ratio $0,0381$) tandis que dans la classe de moyenne mobile exponentielle, les paramètres optimaux proviennent de deux stratégies différentes données par $m = 20$, $n = 100$ et $m = 20$, $n = 200$. Leur ratio est de $0,0547$. Dans la suite on travaille avec une de ces deux stratégies à savoir $m = 20$, $n = 100$ car sa performance par rapport à l'autre stratégie $m = 20$, $n = 200$ a été la même aussi pour les données hors échantillon. L'idée est de comparer les trois meilleures stratégies issues de ces trois classes par rapport à la stratégie de référence préconisée par la théorie de l'efficience de marchés. On essaie de voir, en d'autres termes, si on peut battre le marché en ignorant les coûts de transaction. Dans le tableau ci-dessous, on donne leur performance par rapport aux données hors échantillon.

TABLE 6.3 – Ratio de Sharpe des meilleures stratégies sur l'indice CAC 40.

Out of sample (Oct 2002, Dec 2010)				
	Stratégie	Moyenne	Ecart-type	Ratio Sharpe
1) Achat/Conservation		0,7473	61,480	0,0122
2) RCA(1)	(m=8,n=20)	-0,0911	61,4848	-0,0015
3) Moy. simple	(m=15,n=50)	-1,6464	61,4628	-0,0268
4) Moy. expo	(m=20,n=100)	-3,3049	61,396	-0,0538

Sur les données hors échantillon (567 données), on remarque que seule la stratégie d'achat conservation arrive à fournir des rentabilités positives. La stratégie Achat et Conservation se rattrape de sa phase de sous-performance notée dans la partie apprentissage. Elle arrive ici en première position avec un ratio égal à $0,0122$. Ensuite, vient la stratégie de la classe RCA avec un ratio négatif mais faible de $-0,0015$. Enfin les deux dernières stratégies basées sur la moyenne mobile terminent la marche. De ces résultats, il semble sur l'intervalle global (Janvier 2002- Décembre 2010) qui comporte aussi bien la période avant et un peu après la crise Subprime, que la stratégie issue de RCA(1) soit meilleure. En effet leur moyenne globale respective selon le listing du tableau 6.3 est donnée par : $-0,22$; $2,92$; $1,24$; $1,54$ tandis que leur écart type reste presque identique $58,776$; $58,704$; $58,763$; $58,756$; respectivement. Il est clair que si les transactions sont ignorées, toutes les autres stratégies font mieux que celle consistant à ne détenir que l'indice (Achat et Conservation). Il reste à voir si cette différence est significative ou pas sur le plan statistique. Pour ce faire, on note z_t^1, z_t^2, z_t^3, z_t^4 la série des variations des prix ($\Delta S_t = S_t - S_{t-1}$) générée par les quatre stratégies selon toujours l'ordre du listing au niveau du tableau 6.3. Si

l'hypothèse de normalité et d'indépendance était satisfaite on pouvait appliquer simplement la t-statistique pour le test d'égalité entre deux moyennes en regardant par exemple leur différence $z_t = z_t^j - z_t^1$ que l'on comparera par rapport à la moyenne nulle. Comme dans le cas des rentabilités financières, le tableau ci-dessous montre que la distribution de la variation des prix est très différente d'une distribution normale. L'illustration est faite pour le cas $j = 2$ correspondant à la stratégie appartenant à la classe RCA(1). La même conclusion est obtenue pour les autres stratégies ($j = 3$ et $j = 4$).

TABLE 6.4 – Tests de Normalité pour $z_t = z_t^2 - z_t^1$

Test	Statistique			p -Value
Kolmogorov-Smirnov	D	0,26	Pr > D	<0,0100
Cramer-von Mises	W-Sq	52,42	Pr > W-Sq	<0,0050
Anderson-Darling	A-Sq	228,27	Pr > A-Sq	<0,0050

On voit que, pour les trois types de statistique utilisés, on arrive au même résultat à savoir le rejet de la distribution normale au seuil de signification 1% pour la distribution de la variation des prix. C'est pourquoi, on propose le test de signe qui est plus robuste car ne faisant aucune hypothèse sur la distribution de z_t. Seule l'hypothèse d'indépendance est faite et qui semble raisonnable car il n'y a aucun lien entre les stratégies j et celle consistant à ne détenir que l'actif. Précisément, l'hypothèse nulle est que z_t est issue d'une distribution continue (symétrique ou pas) [27] avec une médiane égale à 0. Ci -dessous, on montre les résultats obtenus :

TABLE 6.5 – Tests d'égalité de médiane pour le CAC, $z_t = z_t^j - z_t^1$

Indice	Statistique		p -Value
j=2	2,0243	Pr > |stat|	0,0429
j=3	0,11	Pr > |stat|	0,91
j=4	0,31	Pr > |stat|	0,75

27. Comme on l'a noté, un des faits stylisés des séries financières est une présence d'asymétrie au niveau de la distribution inconditionnelle. Il est donc préférable de travailler sans l'hypothèse de symétrie.

Le tableau 6.5 montre que toutes les trois statistiques sont positives, ce qui signifie que leur médianne est supérieure à la stratégie Achat et Conservation. Par contre, cette différence n'est statistiquement significative au seuil de 5% que pour la stratégie $(m = 12, n = 20)$ issue de la classe RCA(1). Pour les deux autres meilleures stratégies issues des classes moyennes mobiles, l'information disponible ne permet pas de rejeter de l'hypothèse nulle.

Ce que l'on a fait, jusqu'à présent, est de comparer **individuellement** des stratégies par rapport à une fixée et considérée comme la référence (Achat et Conservation). Une autre possibilité plus intéressante est de faire une comparaison globale. En d'autres termes, on compare une stratégie fixée par rapport à toutes les autres restantes en une seule fois au lieu que cela soit par deux à deux. L'hypothèse nulle est ici qu'aucune stratégie dans le groupe n'a pu faire mieux que la stratégie de référence. Si elle est rejetée, cela signifie que parmi les stratégies compétitives, il y a eu au moins une qui est arrivée à surperformer par rapport à la stratégie de référence. Le test utilisé est connu sous l'acronyme SPA (Superior Predictive Ability). Il a été d'abord proposé par White [112] et étendu par Hansen [52] pour tenir compte surtout de l'inclusion dans l'alternative de modèles qui ont sous performé. La statistique est définie comme précédemment à partir des variables

$$d_{j,t} = z_t^j - z_t^{ref}, \qquad j \in I = \{1,\ 2,\ 3,\ ,4\}, \quad j \neq ref$$

où ref est un entier appartenant à l'ensemble I et spécifiant le modèle de référence. L'indice j parcourt le restant des autres stratégies. Posons $\mathbf{d}_t = (d_{i_1,t}, \cdots, d_{i_m,t})$ le vecteur regroupant la performance des différentes alternatives par rapport au modèle de référence. Dans cette étude, m est égale à 3. Si $\mu = E(\mathbf{d}_t)$ est bien défini, alors l'hypothèse nulle peut être récrite comme suit

$$H_0 : \quad \mu \leq 0_{\mathbb{R}^m} \quad i.e \quad \mu_k \leq 0, \quad pour \quad k = i_1, \cdots, i_m$$

Si la série $(\mathbf{d}_t)$ est (strictement) stationnaire et de mémoire courte, on a alors d'après le théorème centrale limite

$$n^{1/2}(\bar{\mathbf{d}} - \mu) \longrightarrow N_m(0, \Omega)$$

où $\bar{\mathbf{d}} = n^{-1} \sum_{t=1}^{n} \mathbf{d}_t$ et Ω est l'estimateur asymptotique de la matrice de variance covariance. White [112] propose la statistique suivante pour l'hypothèse nulle donnée ci dessus :

$$T_n^{RC} = n^{1/2} \max(\bar{d}_{i_1}, \cdots, \bar{d}_{i_m})$$

Cette statistique T_n va converger en distribution sous l'hypothèse nulle $\mu = 0_{\mathbb{R}^m}$ (grâce au théorème de continuité) vers $\max(G_{i_1}, \cdots, G_{i_m})$ où $(G_{i_1}, \cdots, G_{i_m}) \sim N_m(0, \Omega)$. Or cette distribution, faisant intervenir un maximum sur des variables gaussiennes, n'est pas connue de façon explicite. On fait donc recours aux méthodes statistiques de types Boostrap pour obtenir une distribution empirique.

L'approche de Hansen [52] a été de travailler, d'une part, sur une statistique de type Student et, d'autre part, de négliger l'impact des stratégies qui ont sous performé. Précisément, sa statistique est donnée par

$$T_n^{SPA} = \max\left\{ \max\left(\frac{n^{1/2}\bar{d}_{i_1}}{\sqrt{var(n^{1/2}\bar{d}_{i_1})}}, \cdots, \frac{n^{1/2}\bar{d}_{i_m}}{\sqrt{var(n^{1/2}\bar{d}_{i_m})}} \right), 0 \right\}$$

Cette statistique est très favorable par rapport à l'alternative, car elle est définie comme un maximum d'un maximum. C'est pourquoi, ces $p - values$ correspondantes sont appelées Lower P-values par l'auteur. De même, il a proposé une statistique qui dépend de la variance des échantillons générés lors du processus de Boostrap. On appellera ces p-values *consistant* P-values. Enfin, les p-values associées à la statistique de White seront dénotées Upper P-values car elles sont plus favorables à l'hypothèse nulle. Nous montrons ci dessous seulement les Upper P-values et Lower P-values. Les résultats sont obtenus pour un nombre de ré-échantillonnages égal à 5000, une longueur de bloc égale à 10 et un type de Boostrap stationnaire.

Table 6.6 – Tests de White et Hansen pour l'indice CAC 40.

Benchmark	long. bloc	SPA_L	SPA_U
Achat/Conservation	10	0,0606	0,0606
RCA, (12, 20)	10	0,6006	0,955
Moy. mob. simpl (15,150)	10	0,1712	0,2492
Moy. mob. expo (20,100)	10	0,2108	0,3242

Comme noté, la statistique de White est très conservatrice et souvent l'hypothèse nulle est rarement rejetée.
Concernant la statistique proposée par Hansen, on note que si les autres stratégies alternatives sont les références, on n'arrive pas à rejeter l'hypothèse nulle sous le seuil de signification $\alpha = 10\%$. Par contre, si c'est la stratégie acheter/conserver qui est la référence, alors l'hypothèse

nulle est rejetée pour le même seuil $\alpha = 10\%$. Ce qui veut dire qu'il existe au moins une stratégie parmi les alternatives qui lui est supérieure. Vraisemblablement, cette stratégie est celle provenant de la classe RCA, vu que sa médiane a été notée supérieure à celle de la stratégie Achat et Conservation. Nous avons varié plusieurs paramètres tels que le nombre de simulations, la longueur du bloc, les conclusions restent les mêmes. Pour économiser de l'espace, ces tableaux ne sont pas montrés.

Nous terminons par l'indice américain $S\&P$ 500 qui est une pondération des 500 plus grandes entreprises américaines. Cet indice est souvent une référence pour des études liées à l'efficience de marché. L'objectif est de voir, si les résultats satisfaisants obtenus dans l'indice CAC 40 vont persister sur cet indice. Le tableau qui suit montre la performance de toutes les stratégies (au total 49) dans l'intervalle d'étude réservé à l'apprentissage.

TABLE 6.7 – Ratio de Sharpe des différentes stratégies sur le S & P 500.

In sample (Janvier 2002, Oct 2008)				
RCA(1)				
Stratégie(m,n)	20	24	32	40
4	0,04821015	0,05903548	0,03193214	0,05826279
8	0,04447684	0,05747905	0,04043064	0,05832051
12	0,05448836	0,05527822	0,01898549	0,05513401
16	0,06555362	0,07722972	0,04895556	0,08431328
Achat/Conservation			$2,96 \times 10^{-4}$	
Moyenne mobile simple				
Stratégie(m,n)	50	100	150	200
5	0,00529631	-0,00249832	0,00178239	-0,00249832
10	-0,00121095	0,00752571	-0,00126096	0,00752571
15	-0,00080437	0,01275191	-0,01091218	0,01275191
20	0,01220257	0,00209544	-0,01431107	0,00209544
Moyenne mobile exponentielle				
Stratégie(m,n)	50	100	150	200
5	0,02094178	0,00785823	0,00825559	0,00785823
10	0,01161063	0,0105259	0,01236654	0,0105259
15	0,01176163	0,0052148	0,02192882	0,0052148
20	0,00998865	0,01090569	0,01807831	0,01090569

Une première remarque, à propos du tableau ci-dessus, est que beaucoup de stratégies ont fourni des rentabilités négatives dans la classe de la moyenne mobile simple. Par contre, les deux autres classes (RCA, moyenne mobile exponentielle) continuent toujours de générer des rentabilités positives comme dans le cas de l'indice CAC 40. On note aussi que l'indice S&P a fourni une rentabilité positive contrairement à l'indice CAC. Le ratio de Sharpe le plus élevé provient de la classe RCA avec, cependant, des paramètres optimaux différents $m = 16$, $n = 40$ de l'indice français. Dans la suite, on regarde seulement la performance des meilleures stratégies pour chaque classe :

TABLE 6.8 – Ratio de Sharpe des meilleures stratégies sur l'indice S&P 500.

Out of sample (Oct 2002, Dec 2010)				
	Stratégie	Moyenne	Ecart-type	Ratio Sharp
1) Achat/Conservation		0,1703	18,344	0,0093
2) RCA(1)	(m=16,n=40)	-0,8442	18,34	-0,0461
3) Moy. simple	(m=15,n=100)	-0,4158	18,34	-0,0227
4) Moy. expo	(m=15,n=150)	0,4523	18,34	0,0247

Dans ce tableau, la stratégie issue de RCA donne les résultats les moins satisfaisants malgré sa performance sur le domaine d'apprentissage. La meilleure stratégie provient de la moyenne mobile exponentielle suivie ensuite de de la stratégie d'achat / conservation. Sur l'échantillon globale, les rentabilités de ces quatre stratégies sont données par $0,0454$; $0,5894$; $0,0175$; 0.3221. La rentabilité assez élevée de la stratégie RCA s'explique par le fait qu'elle a capté la bonne direction concernant les larges mouvements (valeurs extrêmes) mais pas pour les petites variations. Ce qui va se matérialiser par une médiane faible. Précisément, les médianes sur l'échantillon global des différentes stratégies, sont données respectivement par $0,8450$; $0,2200$; $0,5350$; $0,8250$ en respectant l'ordre de listing du tableau 6.8.
La médiane de la stratégie achat/conservation est donc la plus élevée. Par contre, sa différence par rapport aux autres médianes n'a pas été trouvé significative au seuil $\alpha = 5\%$, voir le tableau 6.9.

TABLE 6.9 – Tests d'égalité de médiane pour le S&P, $z_t = z_t^j - z_t^1$

Indice	Statistique		p -Value
j=2	-1,8452	Pr > |stat|	0,0650
j=3	-1,2814	Pr > |stat|	0,20
j=4	-0,136	Pr > |stat|	0,89

On termine cette partie de prédiction de la direction des prix, en revenant sur la variation des prix définis par les variables z_t^j, $j = 1, \cdots, 4$, $t = 1, \cdots, N$ où N est le nombre total de données. L'idée est de voir s'il existe une stratégie qui n'est pas dominée par les autres alternatives dans un test collectif. Nous montrons dans le tableau 6.10 les P-values SPA_L, SPA_U associées respectivement aux statistiques de Hansen et de White.

TABLE 6.10 – Tests de White et Hansen pour le S&P.

Benchmark	long. bloc	SPA_L	SPA_U
Achat/Conservation	5	0,1528	0,1574
RCA, (12, 20)	5	0,559	0,9176
Moy. mob. simpl (15,150)	5	0,1152	0,1152
Moy. mob. expo (20,100)	5	0,3008	0,4232

On remarque que dans le cas du S& P, on n'arrive plus à rejeter l'hypothèse nulle si la stratégie acheter/conserver est la référence. Cette même remarque s'applique aussi pour les autres stratégies même si celle utilisant une moyenne mobile simple est assez limite. [28] Dans les deux indices, on notera que la p-value associée à la stratégie RCA demeure la plus élevée. Cela est dû au fait que ces statistiques sont basées sur une moyenne des rentabilités et non sur la médiane. Or, on a vu que dans les deux cas, la rentabilité moyenne de la meilleure stratégie était issue de la classe

28. Pour les deux indices $S\&P$ et CAC, j'ai collecté toutes les 49 stratégies pour un test de supériorité. En d'autres termes, chaque stratégie de référence est comparée par rapport aux 48 stratégies restantes. Les résultats obtenus sont similaires à notre ensemble réduit de 4 stratégies.

RCA. La conclusion générale que l'on tire de cette première application est la confirmation de notre postulat d'efficience de marché en moyenne. Comme on l'a vu, dans le cas simplifié où toutes les transactions sont ignorées, une différence significative au niveau de la médiane des rentabilités n'a pas pu être obtenue au niveau de l'indice $S\&P$. Pour l'indice CAC 40, cette différence significative a été notée mais avec les coûts de transactions, le profit pourrait être nul à la fin. Ces résultats montrent aussi qu'il est plus difficile de battre le marché $S\&P$ que celui du CAC 40, du moins sur la période d'étude considérée.

La deuxième application consistera à proposer un modèle économétrique pour prédire non plus la direction des rentabilités, mais leur volatilité conditionnelle. Là également, pour tenir compte de la complexité de la dynamique des prix, on a proposé un modèle qui se comporte comme un modèle GARCH en moyenne.

Chapitre 7

Prédiction de la volatilité.

7.1 Introduction

Comprendre l'évolution de la volatilité des marchés a été toujours un challenge pour les investisseurs financiers. Cette information peut être utile dans différentes activités telles que le management de risque, le développement de stratégies d'investissement ou l'évaluation (pricing) de produits dérivés. Un fait stylisé que l'on retrouve dans les séries financières est que les fortes rentabilités en valeur absolue ont plus de chance d'être suivies par de fortes rentabilités. La même remarque s'applique aussi pour les petites rentabilités, voir la figure 4.2 du chapitre 4. Ce point a motivé le travail de Engle [35] qui propose le modèle ARCH(P) où la variance conditionnelle dépend linéairement des carrées des innovations passées. Dans les applications financières, pour obtenir une bonne approximation, on a besoin que le paramètre P soit grand et par conséquent un nombre élévé de paramètres doit être estimé. Dans l'optique d'obtenir une paramétrisation plus adéquate, Bollerslev [14] introduit le modèle Generalized ARCH(P) dénoté simplement par GARCH(P,Q). Ce modèle a la capacité de suffisamment bien modéliser la dynamique des prix d'actifs financiers même avec des entiers petits pour P et Q. Vers les années 1990, une autre extension apparaît pour le modèle GARCH qui intègre le fait que les innovations négatives et positives ont des impacts différents sur les volatilités futures. Précisément, les études empiriques ont montré que les chocs négatifs ont un impact plus fort que ceux positifs (effets leviers). Aujourd'hui, il y a plusieurs modèles qui ont été développés et qui tiennent en compte cette asymétrie d'effets, voir le modèle EGARCH proposé par Nelson [89], le modèle TARCH de Zakoian [115] ou le modèle AGARCH de Glosten, Jagannathan et Runkle [45] (GJR pour faire court).

Une troisième extension considère plusieurs régimes de modèles GARCH univariés au lieu d'un seul pour pouvoir générer assez de skewness et kurtosis comme noté dans les rentabilités financières. Les différents régimes peuvent être indépendants et chacun apparaissant avec une certaine probabilité donnée. On parle dans ce cas, d'un mélange de modèles GARCH, voir Haas, Mittnik et Paolella [51] pour une revue. La seconde possibilité est de choisir les régimes à travers une chaîne de Markov (Markov regime switching GARCH models). Quelques papiers relatifs à cette seconde possibilité sont donnés par Klaassen [66], Marcucci [84] et les références qui y sont citées. Tous ces modèles, basés sur plusieurs régimes, ont pour objectif de capturer le fait que l'état de la volatilité (faible, élevée) ne persiste pas longtemps dans le temps. En guise d'illustration, on peut voir dans ces modèles, un petit choc suivi par un grand choc et inversement. Leur avantage par rapport au régime GARCH univarié est donc leur flexibilité. Il est cependant connu que les modèles avec plusieurs paramètres en entrée peuvent avoir des problèmes de robustesse dans le

cas d'une évaluation hors échantillon. Toutes ces remarques m'ont motivées à proposer un modèle mathématique qui est basé sur un modèle GARCH avec un seul régime où l'asymétrie est prise aussi en compte. La particularité de ce modèle par rapport aux autres est qu'il est très proche du modèle GARCH standard. Cette approche a été choisie car il a été montré que le modèle GARCH standard a la capacité de modéliser de façon satisfaisante les séries financières, voir par exemple Andersen et Bollerslev [5] ou Hansen et Lunde [53]. On s'est donc posé la question suivante : Peut on trouver un modèle AGARCH qui partage des similarités avec le modèle GARCH(P,Q) et avec le moins de paramètres possibles ? Si la réponse est positive, le mieux que l'on puisse faire est de trouver un modèle avec $P + Q + 2$ paramètres où le paramètre additionnel par rapport au modèle GARCH standard GARCH(P,Q) devrait gérer l'effet d'asymétrie.[29] Ce modèle est obtenu en considérant un modèle économétrique qui se comporte en moyenne comme un modèle GARCH. On l'appellera dans la suite SUGARCH pour Stochastic Unit GARCH. Il est défini comme un modèle GARCH standard où cependant quelques coefficients du GARCH sont multipliés par un facteur stochastique prévisible d'espérance un. Il y a plusieurs façons de construire de tels modèles puisque le facteur stochastique peut être associé à chaque coefficient du modèle GARCH(P,Q). Dans la section suivante, on présente d'abord la formulation de notre modèle économétrique SUGARCH suivi d'une description des autres modèles compétitifs (EGARCH, GJR, GARCH). La comparaison sera basée sur la capacité des modèles à prédire la volatilité des marchés financiers.

7.2 Quelques modèles GARCH asymétriques.

7.2.1 La classe SUGARCH.

L'idée de la classe SUGARCH est de capturer quelques propriétés du modèle GARCH standard tout en intégrant l'effet asymétrique entre les chocs positifs et négatifs. Précisément, nous la définissons comme un modèle GARCH avec certains coefficients multipliés par un facteur stochastique disons v_t qui est tel que $E(v_t) = 1$. On admettra de plus que la distribution conditionnelle des innovations appartient à la classe des distributions symétriques. Cette condition n'est pas si restrictive que cela vu qu'elle peut englober des lois de type Student qui est à queues épaisses par rapport à la loi normale.

Contrairement aux autres modèles asymétriques, la classe SUGARCH(P,Q,O) ne peut inclure un entier plus grand que 1 pour son troisième paramètre O. En d'autres termes, un seul paramètre ($O = 1$) est utilisé pour gérer l'effet levier. Pour simplifier la notation, on remplacera SUGARCH(P,Q,1) dans la suite par SUGARCH(P,Q) où P, Q sont des entiers positifs. Dans cette étude, on s'intéresse à la prédiction de la volatilité sur des horizons à court terme et où les données seront journalières. On peut donc négliger la moyenne conditionnelle et la supposer constante comme dans les études de Klassen [66] ou de Marcucci [84]. Notons $r_t = \ln S_t - \ln S_{t-1}$ la rentabilité à l'instant t d'un actif donné et μ sa moyenne conditionnelle, alors la classe SUGARCH(P,Q) est définie par :

29. Un modèle GARCH (P,Q) a au total P+Q+1 paramètres.

$$r_t = \mu + \epsilon_t = \mu + \sigma_t \eta_t \ , \quad (\epsilon_t = \sigma_t \eta_t) \tag{7.1a}$$

$$\sigma_t^2 = \alpha_{0,t} + \sum_{i=1}^{P} \alpha_{i,t} \epsilon_{t-i}^2 + \sum_{j=1}^{Q} \beta_{j,t} \sigma_{t-j}^2, \tag{7.1b}$$

$$\eta_t \sim IID \ f(0,1), \tag{7.1c}$$

où $\alpha_{0,t}$, $\alpha_{i,t}$, $\beta_{0,t}$ peuvent être constants ($\alpha_{0,t} = \alpha_0$, $\alpha_{i,t} = \alpha_i$, $\beta_{j,t} = \beta_j$) ou stochastiques. Dans le dernier cas, leur expression est donnée par :

$$\alpha_{0,t} = \alpha_0 * v_t, \quad \alpha_{i,t} = \alpha_i * v_t; \quad \beta_{j,t} = \beta_j * v_t; \qquad v_t = 1 - \gamma \epsilon_{t-1}. \tag{7.2}$$

Le terme $f(0,1)$ représente une fonction de densité symétrique ayant une moyenne zéro et une variance unitaire. Le paramètre γ modélise l'effet asymétrique qui existe entre les innovations positives et négatives sur la volatilité future. Remarquons d'après l'équation (7.2) que $E(v_t) = 1$ en se basant sur les équations (7.1a) et (7.1c). Notons aussi qu'il y a $2^{P+Q+1} - 1$ formulations possibles de modèles asymétriques dans la classe SUGARCH.[30] Dans la suite, on travaillera avec le modèle général de la classe SUGARCH où chaque coefficient est stochastique.[31] Dans ce cas, la volatilité conditionnelle (7.1b) peut être récrite comme suit :

$$\sigma_t^2 = \left(\alpha_0 + \sum_{i=1}^{P} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{Q} \beta_j \sigma_{t-j}^2 \right) - \gamma \epsilon_{t-1} \left(\alpha_0 + \sum_{i=1}^{P} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{Q} \beta_j \sigma_{t-j}^2 \right). \tag{7.3}$$

Pour obtenir une variance inconditionnelle similaire au modèle GARCH, on a besoin que l'espérance du second terme de l'équation (7.3) soit égale à zéro. Pour cela, on suppose que la distribution de η_t est symétrique (ses moments d'ordre impair sont donc nuls) et de variance égale à 1. Dans un tel cadre, on a grâce à la loi itérée des espérances :

$$E(\epsilon_{t-1}) \ = \ E[(\epsilon_{t-1} \cdot \epsilon_{t-i}^2] \ = \ E[(\epsilon_{t-1} \cdot \sigma_{t-j}^2] = 0, \quad \forall i \geq 2, \ \ j \geq 1. \tag{7.4}$$

Le fait que η_t soit une distribution symétrique entraîne aussi que

30. Puisque chaque paramètre peut prendre deux formes, le nombre total de choix possibles est donc 2^{P+Q+1}. Le moins 1 correspond au modèle GARCH standard (tous les coefficients constants) qui a été exclu car n'intégrant pas l'asymétrie.

31. Le même raisonnement peut s'appliquer pour les autres modèles de la classe SUGARCH

$$E\Big[(\epsilon_{t-1})^{2n+1}\Big] = 0 \quad \forall n \in \mathbb{N}. \tag{7.5}$$

Ces deux équations impliquent que :

$$E\left[\gamma\epsilon_{t-1}\left(\alpha_0 + \sum_{i=1}^{P}\alpha_i\epsilon_{t-i}^2 + \sum_{j=1}^{Q}\beta_j\sigma_{t-j}^2\right)\right] = 0.$$

Par conséquent, l'équation (7.3) devient

$$E(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{P}\alpha_i E(\epsilon_{t-i}^2) + \sum_{j=1}^{Q}\beta_j E(\sigma_{t-j}^2),$$

Par suite, l'équation (7.6) est obtenue en notant que $E(\sigma_{t-j}^2) = E(\epsilon_{t-i}^2)$ pour tout entier i, j.

$$E(\sigma_t^2) \;=\; \bar{\sigma} = \frac{\alpha_0}{1 - \sum_{i=1}^{P}\alpha_i - \sum_{i=1}^{Q}\beta_i}. \tag{7.6}$$

Vu que la volatilité est toujours positive et pour obtenir un modèle stationnaire, les contraintes suivantes sont faites

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad i = 1, \cdots, P; \quad \beta_j \;\geq\; 0, \quad j = 1 \cdots, Q; \quad \sum_{i=1}^{P}\alpha_i \;+ \sum_{j=1}^{Q}\beta_j \;< 1 \tag{7.7a}$$

$$|\gamma| < \frac{1}{max\{|\epsilon_t|, t = 1, \cdots, N\}}\;, \quad \epsilon_t = r_t - \mu, \tag{7.7b}$$

où N est le nombre total d'observations au niveau des rentabilités.

L'équation (7.7a) donne les mêmes contraintes qu'un modèle GARCH standard tandis que l'équation (7.7b) gère la contrainte de positivité du facteur stochastique $v_t = 1 - \gamma\epsilon_{t-1}$. Une borne générale peut être prise pour les innovations et donc pour le paramètre γ. Nous avons préféré laisser cette borne dépendre des données. L'idée de cette approche est de permettre un intervalle assez grand pour γ de sorte que l'effet asymétrique soit bien pris en compte. En pratique, vu que l'on travaille sur des observations (intra) journalières, la moyenne conditionnelle μ dans l'équation (7.1a) est souvent négligeable et donc les innovations peuvent être approximées par les rentabilités financières r_t, voir l'équation (7.7b). [32] Il est aussi attendu que γ sera positif pour

32. Dans l'implémentation, on a utilisé $|\gamma| < \frac{0.9}{max\{|r_t|, t=1,\cdots,T\}}$.

intégrer le fait que les chocs négatifs impactent plus les volatilités futures.

D'un autre côté, quand un choc apparaît dans les innovations, la volatilité ne peut pas persister dans le temps. Cela est dû au facteur multiplicatif v_t qui permet alternativement de générer des mouvements importants et faibles au niveau de la volatilité puisque (ϵ_t) se comporte comme un jeu équitable, voir la définition 1.3.2 du chapitre 1. En d'autres termes, on peut avoir un petit choc dans une période de forte volatilité ou un choc important dans une période de faible volatilité. Ce point est aussi partagé par les modèles GARCH avec plusieurs régimes. Pour voir formellement le lien, notons que si on a deux régimes comme c'est souvent le cas dans les applications financières, les coefficients du modèle GARCH prennent chacun 2 différentes valeurs. Ici, pour plus de flexibilité, la classe SUGARCH permet aux paramètres de prendre une infinité de valeurs car v_t est à valeurs réelles. [33]
Même si les coefficients sont aléatoires, la classe SUGARCH offre des propriétés presque similaires au modèle GARCH. En effet, on a vu que leur variance inconditionnelle était identique. La différence apparaît seulement au niveau local où le facteur v_t génère des effets d'asymétrie pour la variance conditionnelle. Ces mouvements entraînent par conséquent des queues plus épaisses au niveau de la distribution inconditionnelle des rentabilités. Nous illustrons ce fait dans le cas gaussien $\eta_t \sim N(0,1)$ avec comme paramètres $P = 1$ et $Q = 1$. Précisément, dans la classe SURGARCH(1,1), on considère le modèle défini par la volatilité conditionnelle suivante :

$$\sigma_t^2 = \alpha_0 v_t + \alpha_1 v_t \epsilon_{t-1}^2 + \beta_1 v_t \sigma_{t-1}^2 \;=\; \Big(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2\Big) v_t, \qquad v_t = 1 - \gamma \epsilon_{t-1}.$$

On a alors,

$$\begin{aligned} E(\sigma_t^4) &= E\Big[(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2 (1 - \gamma \epsilon_{t-1})^2\Big] \\ &= E\Big[(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2 (1 - 2\gamma \epsilon_{t-1} + \gamma^2 \epsilon_{t-1}^2)\Big] \\ &= E\Big[(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2 (1 + \gamma^2 \epsilon_{t-1}^2)\Big]. \end{aligned}$$

La dernière équation résulte du fait que $E\Big[\epsilon_{t-1}(\alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2)^2\Big] = 0$. Pour les raisons, on peut voir les équations (7.4), (7.5) qui entraînent que

$$E(\epsilon_{t-1}) = E(\epsilon_{t-1}^3) = E(\epsilon_{t-1}^5) = E(\epsilon_{t-1}\sigma_{t-1}^2) = E(\epsilon_{t-1}\sigma_{t-1}^4) = E(\epsilon_{t-1}^3\sigma_{t-1}^2) = 0.$$

33. Nous avons essayé d'autres formulations pour la classe SUGARCH qui sont proches des modèles à 2 régimes. Par exemple, on a défini $v_t = 1 - \gamma * sign(\epsilon_{t-1})$ qui est à valeur binomiale ($1-\gamma$ ou $1+\gamma$). Cette formulation donne également la même variance inconditionnelle que le modèle GARCH standard ainsi qu'une formule fermée pour le moment d'ordre 4. Cependant, les résultats obtenus ne sont pas si satisfaisants. C'est pourquoi, on a préféré travailler avec $v_t = 1 - \gamma * \epsilon_{t-1}$.

D'où,

$$E(\sigma_t^4) = E\big(\alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2\big)^2 \ + \ \gamma^2 E\Big[\epsilon_{t-1}^2(\alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2)^2\Big].$$

En développant la dernière égalité et en utilisant les points suivants basés sur l'hypothèse $\eta_t \sim N(0,1)$

$$E(\epsilon_{t-1}^2) = E(\sigma_{t-1}^2), \qquad E(\epsilon_{t-1}^4) = 3E(\sigma_{t-1}^4), \qquad E(\epsilon_{t-1}^2\sigma_{t-1}^2) = E(\sigma_{t-1}^4)$$

$$E(\epsilon_{t-1}^6) = 15E(\sigma_{t-1}^6), \quad E(\epsilon_{t-1}^2\sigma_{t-1}^4) = E(\sigma_{t-1}^6), \quad E(\epsilon_{t-1}^4\sigma_{t-1}^2) = 3E(\sigma_{t-1}^6),$$

on obtient l'équation suivante[34]

$$E(\sigma_t^4) = \frac{\alpha_0^2 + E(\sigma_t^2)\Big[2\alpha_0\alpha_1 + 2\alpha_0\beta_1 + \gamma^2\alpha_0^2\Big] + \gamma^2 E(\sigma_t^6)\Big[15\alpha_1^2 + \beta_1^2 + 6\alpha_1\beta_1\Big]}{1 - 3\alpha_1^2 - \beta_1^2 - 2\beta_1\alpha_1 - 2\gamma^2(3\alpha_0\alpha_1 + \alpha_0\beta_1)}. \tag{7.8}$$

Le kurtosis est ensuite déduit de la formule

$$\kappa_{SUGARH} \ = \ \frac{E(r_t^4)}{E(r_t^2)^2} \ = \ \frac{3E(\sigma_t^4)}{E(\sigma_t^2)^2}.$$

L'expression ci-dessus n'est pas explicite. En effet, la variance inconditionnelle $E(\sigma_t^2)$ est connue, voir (7.6) mais pas le terme $E(\sigma_t^4)$ qui dépend du moment inconnu d'ordre six $E(\sigma_t^6)$. Malgré tout, pour $\gamma \neq 0$, on peut affirmer que le kurtosis de SUGARCH(1,1) est plus grand que celui du GARCH(1,1) standard correspondant à $\gamma = 0$. L'explication vient du fait qu'ils ont le même dénominateur tandis que $E(\sigma_t^4)$ de l'Eq. (7.8) est une fonction croissante de γ^2. Si on veut obtenir une formule explicite pour le moment d'ordre 4, la classe SUGARCH peut être redéfinie (simplifiée) en posant par exemple $v_t = 1 - \gamma \cdot sign(\epsilon_{t-1})$ au lieu de $v_t = 1 - \gamma \cdot \epsilon_{t-1}$ en perdant en contrepartie de la précision dans l'aspect prédiction. On peut aussi donner une formulation moins complexe comme celle de l'équation (7.10a).
Concernant l'aspect prédiction, toutes les formulations de la classe SUGARCH(1,1) offrent des formules explicites comme dans le cas du modèle GARCH(1,1) standard. En effet, pour tout entier $h \in \mathbb{N}^*$, on a :

$$E_t(\sigma_{t+h}^2) \ = \ \frac{\alpha_0\Big[1 - (\alpha_1 + \beta_1)^{h-1}\Big]}{1 - \alpha_1 - \beta_1} \ + \ (\alpha_1 + \beta_1)^{h-1}\sigma_{t+1}^2 \ , \tag{7.9}$$

où σ_{t+1}^2 est la volatilité conditionnelle définie dans l'équation (7.1b) i.e

34. Sous la condition que le moment d'ordre six existe.

$$\sigma_{t+1}^2 \;=\; \alpha_{0,t} \;+\; \alpha_{1,t}\epsilon_t^2 \;+\; \beta_{1,t}\sigma_t^2$$

On ne considère que trois modèles parmi les sept que compte la classe SUGARCH(1,1). Ils sont donnés par les expressions suivantes :

$$\sigma_{t+1}^2 = \alpha_0 * (1-\gamma\epsilon_t) \;+\; \alpha_1\epsilon_t^2 \;+\; \beta_1\sigma_t^2 \tag{7.10a}$$
$$\sigma_{t+1}^2 = \alpha_0 \;+\; \alpha_1 * (1-\gamma\epsilon_t)\epsilon_t^2 \;+\; \beta_1\sigma_t^2 \tag{7.10b}$$
$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1\epsilon_t^2 \;+\; \beta_1 * (1-\gamma\epsilon_t) * \sigma_t^2 \;. \tag{7.10c}$$

Les quatre autres modèles restants utilisent au moins deux paramètres stochastiques et on a trouvé qu'ils n'offrent pas de différences significatives par rapport aux formulations (7.10a), (7.10b), (7.10c). Une autre raison du choix de ces trois modèles est qu'ils ont chacun une façon différente d'introduire les effets asymétriques au niveau de leur volatilité conditionnelle. Ce qui nous permettra de savoir quelle est la meilleure façon de capturer l'effet d'asymétrie.

Dans la section suivante, on présente deux modèles asymétriques (EGARCH, GJR) proposés dans la littérature. Additionnellement, on a inclus le modèle symétrique du GARCH standard qui peut aussi donner de bons résultats surtout dans la partie hors-échantillon. Puisque la vraie variance conditionnelle est non observable, les rentabilités au carré ont été utilisées comme proxy. L'avantage d'utiliser le carré des rentabilités est qu'il permet de conserver l'ordre de classement pour différentes fonctions de perte telles que la MSE et QLIKE définies dans les équations (7.14a), (7.14b), respectivement. Les autres fonctions de perte telles que la moyenne absolue prise sur les erreurs de la variance, la moyenne prise sur les erreurs au carré des écarts types, peuvent générer des biais, voir A. Patton [94] pour plus d'informations.

7.2.2 Modèles de Comparaison.

Le modèle EGARCH : De son acronyme Exponential GARCH (EGARCH), ce modèle a été proposé par Nelson [89]. La variable d'intérêt est, ici, le logarithme de la variance conditionnelle. Le modèle est défini par

$$r_t = \mu + \epsilon_t; \qquad \epsilon_t = \sigma_t\eta_t \tag{7.11a}$$
$$\ln\sigma_t^2 = \alpha_0 \;+\; \sum_{i=1}^{P} g(\eta_{t-i}) + \sum_{j=1}^{Q} \beta_j \ln\sigma_{t-j}^2 \tag{7.11b}$$
$$g(\eta_{t-i}) = \alpha_i\Big[|\eta_{t-i}| - E(|\eta_{t-i}|)\Big] \;+\; \gamma_i\eta_{t-i}, \tag{7.11c}$$

où η_t est donné par l'équation (7.1c). L'asymétrie entre chocs positifs et négatifs est prise en compte par les paramètres $(\alpha_i, \gamma_i),\ i = 1, \cdots, P$. Notons que, si on est intéressé par des objectifs

de prédiction, la variance conditionnelle du modèle EGARCH doit dépendre de la distribution de η_t à travers le terme $E(|\eta_{t-i}|)$. Puisque η_t est symétrique, on a proposé deux distributions normalisées. La première est la distribution normale :

$$f_{\eta_t}(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}), \qquad E(|\eta_t|) = \sqrt{\frac{2}{\pi}} \ .$$

La loi de Student est la seconde distribution qui a une queue plus épaisse que celle normale. Sa densité est de la forme

$$f_{\eta_t}(x,\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\nu/2)} \frac{1}{\sqrt{\pi(\nu-2)}} \left(1 + \frac{z^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \qquad E(|\eta_t|) = \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \ .$$

A la date date t, l'espérance conditionnelle de l'équation (7.11b) est égale pour $P = Q = 1$ à

$$E_t(\ln \sigma_{t+1}^2) = \ln \sigma_{t+1}^2 = \alpha_0 + \gamma_1 \eta_t \ + \ \alpha_1 \Big(|\eta_t| - E(|\eta_t|)\Big) + \beta_1 \ln \sigma_t^2 \ .$$

La prédiction de la volatilité pour un horizon d'un jour est donc obtenue par

$$\ln \hat{\sigma}_{t+1}^2 = \hat{\alpha}_0 + \hat{\gamma}_1 \eta_t \ + \ \hat{\alpha}_1 \Big(|\eta_t| - E(|\eta_t|)\Big) + \hat{\beta}_1 \ln \sigma_t^2 \ . \tag{7.12}$$

Pour des horizons de deux jours ou plus, la prédiction est faite de façon récursive, à partir de $E_t(\ln \sigma_{t+h+1}^2) = \alpha_0 + \beta_1 E_t(\ln \sigma_{t+h}^2)$, obtenue grâce à l'équation (7.11c). Ce qui nous donne

$$\hat{\sigma}_{t+h+1}^2 = \exp\left[\hat{\alpha}_0 \ + \ \hat{\beta}_1 \ln \hat{\sigma}_{t+h}^2\right],$$

où la condition initiale est fournie par l'équation (7.12). Tous les paramètres sont estimés par la méthode de la log vraisemblance.

Le modèle GJR GARCH : Il a été proposé par Glosten, Jagannathan et Runkle [45]. La variance conditionnelle de ce modèle est définie par

$$r_t = \mu + \epsilon_t; \qquad \epsilon_t = \sigma_t \eta_t$$

$$\sigma_t^2 = \alpha_0 \ + \ \sum_{i=1}^{P} \alpha_i \epsilon_{t-i}^2 + \sum_{k=1}^{O} \gamma_k \epsilon_{t-k}^2 I_{\left\{\epsilon_{t-k}<0\right\}} \ + \ \sum_{j=1}^{Q} \beta_j \sigma_{t-j}^2$$

où $I_{\left\{\epsilon_{t-k}<0\right\}}$ est une fonction indicatrice qui prend la valeur 1 si $\epsilon_{t-k} < 0$ et 0 sinon.

Pour que la volatilité soit positive, des contraintes doivent être ajoutées au niveau des paramètres. Concernant, le modèle le plus simple GJR GARCH(1,1,1), on a $\alpha_0 > 0$, $\alpha_1 \geq 0$ $\beta_1 \geq 0$, $\alpha_1+\gamma \geq 0$.

Puisque la distribution de η_t est symétrique, la prédiction de la volatilité pour des horizons de plus d'un jour est faite de façon récursive (pas de formule fermée) à partir de l'expression suivante $E_t(\sigma^2_{t+h+1}) = \alpha_0 + (\alpha_1 + \frac{\gamma_1}{2} + \beta_1)E_t(\sigma^2_{t+h})$. Ce qui donne à l'instant t

$$\hat{\sigma}^2_{t+h+1} = \hat{\alpha}_0 + (\hat{\alpha}_1 + \frac{\hat{\gamma}_1}{2} + \hat{\beta}_1)\hat{\sigma}^2_{t+h}; \tag{7.13a}$$

$$\hat{\sigma}^2_{t+1} = \hat{\alpha}_0 \ + \ \hat{\alpha}_1\epsilon^2_t \ + \ \hat{\gamma}_1\epsilon^2_t I_{\left\{\epsilon_{t-k}<0\right\}} \ + \ \hat{\beta}_1\sigma^2_t. \tag{7.13b}$$

Le modèle GARCH : Il a été proposé par Bollerslev [14] comme une extension du modèle ARCH de Engle [35]. Sa variance inconditionnelle est donnée par l'équation (7.6) et la formule de prédiction est obtenue des équations (7.9) et (7.10a) en posant $\gamma = 0$.

Dans la littérature des modèles GARCH asymétriques, plusieurs paramètres peuvent intervenir pour gérer l'aspect d'asymétrie entre les chocs sans pour autant affecter de façon significative la structure du modèle. Ceci n'est pas le cas pour la classe SUGARCH où le paramètre O est contrainte d'être égal un. La raison en est que si le facteur stochastique v_t contient plus de paramètres ϵ_{t-2}, $\epsilon_{t-3}, \cdots$, alors la formule explicite obtenue pour la la variance inconditionnelle est perdue. Cela est dû au fait que le second terme de l'équation (7.3) peut impliquer des expressions telles que $E(\epsilon_{t-i}\epsilon^2_{t-j})$, $E(\epsilon_{t-i}\sigma^2_{t-j})$ avec $j < i$ et ces termes seront différents de zéro.

7.3 Données et Méthodologie.

On considère les données journalières du S&P 500 (S_t), $t = 1, \cdots, N$ pour évaluer la performance des différents modèles présentés ci-dessus. L'échantillon des données va du 2 Janvier 2002 au 31 Décembre 2010 correspondant à un nombre total de données égal à $N = 2267$. C'est la même série de données utilisée dans la première application qui consistait à prédire la direction des prix avec la classe $RCA(p)$. Pour économiser de l'espace, on ne va montrer que les résultats issus du S&P car ceux issus du CAC 40 lui sont similaires. La table 7.1 donne les statistiques descriptives de cet indice par rapport à la variable rentabilité définie comme suit $r_t = \ln S_t - \ln S_{t-1}$.

TABLE 7.1 – Un résumé de la série (r_t) de S & P 500.

Mean	Std. deviation	Min	Max	Skewness
0,0038	1,3791	-9,469	10,957	-0,139
Kurtosis	JB-stat	LB(20)-stat	LM(10)-stat	
11,854	7,410	97,171	666	
	($<$ 0,001)	(4,0210^{-12})	(0)	

JB-stat représente la statistique de test de normalité proposée par Jarque-Bera. LB(20)-stat correspond à la statistique de test de Ljung-Box basée sur l'auto corrélation des résidus avec des retards allant jusqu'à l'ordre 20 lags. LM(10)-stat est une statistique qui examine la présence d'effets ARCH où 10 retards ont été utilisés pour le carré des résidus. Entre les parenthèses, correspondent les p-values des différentes statistiques.

La table montre que la moyenne des rentabilités de l'indice S&P 500 est positive et petite. L'écart type est aussi petit (1.38%). Le maximum (minimum) des rentabilités est donné par 10.957% (-9.469 %). Les mouvements extrêmes apparaissent plus fréquamment puisque l'hypothèse nulle d'une distribution normale pour les rentabilités est fortement rejectée même au seuil de significativité 1%. La même conclusion est notée aussi pour le test d'absence d'auto-corrélation ainsi que celui de l'absence d'effets ARCH. Leur hypothèse nulle est rejectée avec des p-values proches de zero. Ces résultats issus de la table 7.1 suggèrent d'utiliser des modèles de type GARCH pour prendre en compte l'excès de kurtosis et la présence d'hétérogéneité. Dans la section précédente, on a fait une revue de quelques uns de ces modèles. Pour faire de la prédiction, les paramètres de chaque modèle doivent être estimés au préalable. Pour ce faire, l'approche du Maximum de vraisemblance a été choisie où la distribution conditionnelle des innovations est soit une loi normale soit une loi de Student. Ensuite, les volatilités futures sont prédites et des quantiles (Value-at-Risk) sont également déterminés. L'horizon de temps h appartient ici à l'ensemble $\{1, 2, 5, 10\}$, correspondant à 1 jour, 2 jours, une semaine et deux semaines de prédiction, respectivement. La littérature compare généralement la performance des modèles à travers des fonctions de pertes. Ici, on a considéré qu'une seule fonction de perte comme critère de comparaison qui est donnée par l'équation (7.14b). Une autre fonction de perte importante dans la littérature est la Mean Square Errors (Eq. 7.14a). Ces deux fonctions partagent toutes la qualité de robustesse au cas où un estimateur sans biais de la variance conditionnelle est utilisée, voir Patton [94], Hansen & Lunde [53].

$$MSE = \frac{1}{N-h-T+1} \sum_{t=T}^{N-h} \left(\epsilon_{t+h}^2 - \hat{\sigma}_{t+h|t}^2 \right)^2 \tag{7.14a}$$

$$QLIKE = \frac{1}{N-h-T+1} \sum_{t=T}^{N-h} \left(\log \epsilon_{t+h}^2 + \frac{\epsilon_{t+h}^2}{\hat{\sigma}_{t+h|t}^2} \right), \tag{7.14b}$$

où T et N représentent le nombre de données de l'échantillon in sample et de l'échantillon global,

respectivement. Les données in sample couvrent la période du 2 Janvier 2002 au 2 Octobre 2008. Ce qui correspond à un nombre d'observations égal à $T = 1699$ ($\sim 0.75 * N$) et les données restantes (le quart des données) sont utilisées comme hors échantillon. Le paramètre h donne l'horizon de prédiction. Ici, on se focalise seulement sur la fonction de distance QLIKE pour des raisons fournies par Brownlees et al. [18]. Ces derniers notent que QLIKE peut être récrit sans perte de généralité comme suit :

$$QLIKE = \frac{1}{N-h-T+1} \sum_{t=T}^{N-h} \Big(-\log \frac{\epsilon_{t+h}^2}{\hat{\sigma}_{t+h|t}^2} + \frac{\epsilon_{t+h}^2}{\hat{\sigma}_{t+h|t}^2} \Big).$$

Cette fonction de perte fait donc intervenir une combinaison de termes i.i.d (η_t) ; ce qui n'est pas le cas pour la fonction MSE (7.14a). Une autre raison est que MSE ne pénalise pas les petites volatilités prédites (proche de zéro) alors que QLIKE le fait.

Même si prendre un critère basé sur une métrique est important, avoir des tests statistiques pour évaluer si la la différence entre les fonctions de pertes est significative ou pas l'est encore plus. Pour ce faire, on considère le test EPA (Equal Predictive Ability) proposé par Diebold et Mariano [34], DM pour faire court. Posons d_t comme étant la différence de perte entre deux modèles de prédiction i et j i.e

$$d_t = g(\sigma_t, \hat{\sigma}_{t,t|t-h}^i) - g(\sigma_t, \hat{\sigma}_{t|t-h}^j),$$

où g est donnée par l'équation (7.14b). Les termes $\hat{\sigma}_{t|t-h}^i$, $\hat{\sigma}_{t|t-h}^j$ ci-dessus représentent les prédictions de la volatilité σ_t des modèles i et j faites à la date $t-h$. Le Test asymptotique de DM est basé sur la moyenne de la série temporelle (d_t), $t = T+h, \cdots, N$ qui est définie ci-dessous. Si la série est stationnaire et de mémoire courte, les auteurs montrent que

$$\sqrt{n_h}(\bar{d} - \mu) \xrightarrow{d} \mathcal{N}(0, V(\bar{d})), \qquad \bar{d} = \frac{1}{n_h} \sum_{t=T+h}^{N} d_t, \qquad n_h = N - h - T + 1,$$

où μ représente la moyenne de la différence des fonctions de pertes. La variance est estimée de la manière suivante $\hat{V}(\bar{d}) = \frac{1}{n_h}(\hat{\gamma}_0 + 2\sum_{k=1}^{q} \omega_k \hat{\gamma}_k)$. Ici, $\hat{\gamma}_k$ représente une estimation de l'auto corrélation d'ordre k de la série (d_t), ω_k est son poids associé et q précise le lieu de la troncature. L'approche de Marcucci[84] est suivie. On a posé $q = 4 * (\frac{N}{100})^{2/9}$, $w_k = 1 - \frac{k}{q+1}$. Le test E.P.A de Dielbold et Mariano sous l'hypothèse nulle $\mu = 0$ est asymptotiquement donné par

$$\frac{\bar{d}}{\sqrt{V(\bar{d})/n_h}} \sim \mathcal{N}(0,1). \tag{7.15}$$

Après avoir évalué les fonctions de pertes des différents modèles et analysé leur signification statistique, on investit leur capacité de prédire l'état de la volatilité. Il est supposé que la volatilité

ne peut prendre que deux états (faible ou forte volatilité). Ce point est important pour les investisseurs financiers car leur permettant de créer des stratégies d'investissement. Par exemple, si pour un jour donné, la volatilité est prédite d'être forte (faible), une option peut être achetée (vendue) le jour suivant. Cela suppose d'avoir une idée de la direction du marché. Si ce n'est pas le cas (vue neutre), l'option peut être remplacée par un straddle ou un strangle qui combine à la fois des positions sur un Call et un Put. Plus un modèle a la capacité de donner des prédictions précises sur l'état de la volatilité, plus il peut générer des rentabilités élevées. La précision de la prédiction est ici mesurée à travers deux critères : le Success Ratio (SR) et la statistique de test D.A pour Directional Accuracy de Pesaran et Timmermann [95]. Puisque tous les modèles font intervenir la variance conditionnelle σ_t^2, le Success Ratio est défini par rapport à ce paramètre. Précisément, on a

$$SR = \frac{1}{n_h} \sum_{j=0}^{n_h} I_{\left\{\epsilon^2_{d,t+h} \cdot \hat{\sigma}^2_{d,t+h|t} > 0\right\}}, \tag{7.16}$$

où la notation $(a_{d,t}), t \geq 0$ est faite pour la série centrée de (a_t) i.e $a_{d,t} = a_t - mean(a_t),\ t \geq 0$. Les séries temporelles (ϵ^2_{t+h}), $(\hat{\sigma}^2_{t+h|t})$ représentent le proxy de la variance courante et de la variance prédite, respectivement. Si on dénote par P $(\hat{P})$ la fraction de temps que $\epsilon^2_{d,t+h}$ $(\hat{\sigma}^2_{d,t+h|t})$ est positive, alors le test D.A s'écrit comme suit :

$$DA \ = \ \frac{SR - SRI}{\sqrt{Var(SR) - Var(SRI)}}, \tag{7.17}$$

où $SRI = P\hat{P} + (1-P)(1-\hat{P})$. Le terme SR est défini dans l'équation (7.16) et la notation $Var(X)$ est faite pour la variance de X.

Enfin, on termine par comparer les différents modèles en termes de management de risque. Pour ce faire, la Value-at-Risk est déterminée à partir de la distribution de perte des rentabilités prédites pour un horizon pré-déterminé (h) et pour un certain degré de confiance (α). Précisément, considérons l'équation (7.1a) et posons $F_{t,h} = P(r_t \leq r|\mathcal{F}_{t-h})$, $G_{t,h} = P(\eta_t \leq \eta|\mathcal{F}_{t-h})$ où $F_{t,h}$ et $G_{t,h}$ représentent la fonction de répartition de r_t, (η_t), respectivement, étant donné l'ensemble d'information $\mathcal{F}_{t-h}$. La VaR pour un seuil $\alpha \in]0,1[$ et un horizon h, dénoté $VaR_t(\alpha, h)$, est calculée à la date $t-h$ par :

$$VaR_t(\alpha, h) \ = \ F^{-1}_{t,h}(\alpha) \ = \ \mu + \sigma_t G^{-1}_{t,h} \ = \mu + \sigma_t G^{-1}(\alpha) \ .$$

La dernière égalité est expliquée par le fait que le processus (η_t) est i.i.d, voir 7.1c et que $G^{-1}_{t,h} = G^{-1}$. La différence entre les modèles apparaît sur comment σ_t est prédite à partir de la date $t-h$. Après avoir obtenu les estimations de $VaR_t(\alpha, h)$, $t = T+h, \cdots, N$ pour chaque modèle, on définit ces deux fonctions de pertes suivantes pour des investisseurs ayant des positions longues.

$$\hat{\alpha} = \frac{1}{N-T-h+1} \sum_{t=T+h}^{N} 1_{\left(r_t < VaR_t(\alpha,h)\right)} \tag{7.18a}$$

$$\hat{Q}(\alpha) = \frac{1}{N-h-T+1} \sum_{t=T+h}^{N} \left(\alpha - 1_{\{r_t < VaR_t(\alpha,h)\}}\right)\left(r_t - VaR_t(\alpha,h)\right). \tag{7.18b}$$

Le meilleur modèle, par rapport à la fonction de perte (7.18a), est celui qui minimise la distance $d(\hat{\alpha}) = |\hat{\alpha} - \alpha|$ même si il est préférable d'avoir des valeurs satisfaisant en plus $\hat{\alpha} \leq \alpha$. Pour des tests statistiques utilisant l'expression $\hat{\alpha}$, on peut voir Christoffersen [22]. La seconde fonction de perte a été proposée par Koenker et Bassett [67], dénotée KB pour faire court. C'est une fonction qui pénalise fortement les observations pour lesquelles il y a une violation des contraintes de la VaR. Pour être complet, on a évalué la performance des différents modèles par rapport à des investisseurs ayant des positions courtes. [35]

7.4 Résultats empiriques.

Les résultats sont basés sur les données journalières de l'indice S& P 500 ajustées des dividendes. Pour des raisons de robustesse, le même traitement a été effectué aussi pour l'indice CAC 40 index avec des conclusions similaires. On rappelle que tous les paramètres ont été estimés par la méthode de la log vraisemblance avec des innovations η_t ayant des distributions conditionnelles de type Gaussien ou Student. L'estimation a été faite sur l'échantillon de données allant du 2 Janvier 2002 au 2 Octobre 2008 correspondant à 1699 observations journalières. Les labels ASUG, BSUG et CSUG dans la table suivante correspond à la variance conditionnelle donnée par les équations (7.10a), (7.10b), (7.10c), respectivement.

35. La fonction de perte devient dans ce cas $\frac{1}{N-h-T+1} \sum_{t=T+h}^{N} \left(1_{\{r_t > VaR_t(1-\alpha,h)\}} - \alpha\right)\left(r_t - VaR_t(1-\alpha,h)\right)$

TABLE 7.2 – Paramètres estimés sur les différents modèles.

Models	$\hat{\mu}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\gamma}$	Dof
ASUG_N	0,019	0,008	0,062	0,931	0,082	–
	(0,013)	(0,005)	(0,010)	(0,012)	(0,059)	–
ASUG_T	0,019	0,005	0,061	0,935	0,082	10,318
	(0,028)	(0,003)	(0,014)	(0,015)	(0,075)	(0,026)
BSUG_N	0,019	0,008	0,061	0,933	0,082	–
	(0,018)	(0,005)	(0,012)	(0,016)	(0,028)	–
BSUG_T	0,019	0,005	0,060	0,936	0,082	10,897
	(0,018)	(0,003)	(0,012)	(0,013)	(0,042)	(0,013)
CSUG_N	0,019	0,005	0,035	0,962	0,082	–
	(0,018)	(0,003)	(0,010)	(0,011)	(0,009)	–
CSUG_T	0,019	0,003	0,032	0,966	0,082	12,229
	(0,021)	(0,002)	(0,009)	(0,009)	(0,010)	(0,023)
EGARCH_N	0,007	0,000	0,078	0,986	-0,113	–
	(0,020)	(0,003)	(0,015)	(0,002)	(0,012)	–
EGARCH_T	0,017	-0,003	0,071	0,990	-0,111	13,461
	(0,019)	(0,003)	(0,017)	(0,003)	(0,013)	(2,409)
GJR_N	0,002	0,009	0,000	0,938	0,107	–
	(0,020)	(0,002)	(0,011)	(0,011)	(0,017)	–
GJR_T	0,018	0,006	0,000	0,941	0,108	11,716
	(0,019)	(0,002)	(0,014)	(0,012)	(0,020)	(2,236)
GARCH_N	0,034	0,008	0,063	0,932	–	–
	(0,020)	(0,002)	(0,009)	(0,010)	–	–
GARCH_T	0,044	0,004	0,063	0,937	–	9,623
	(0,019)	(0,003)	(0,011)	(0,011)	–	(1,753)

Note : Chaque modèle a été estimé par le Maximum de vraisemblance avec des distributions de type Normal (_N) et Student(_T) pour η_t. L'écart type asymptotique des estimations est mis entre parenthèses.

Comme attendu, la valeur estimée pour γ (paramètre d'asymétrie) est telle que les chocs négatifs impactent plus les volatilités futures que ceux positifs. Par conséquent, il est positif pour les modèles SUGARCH et GJR et négatif pour le modèle EGARCH. La persistance des chocs sur la volatilité $(\alpha_1 + \beta_1)$ est aussi importante (> 0.99) pour tous les modèles comme c'est le cas souvent pour les séries financières. On note aussi qu'introduire des effets d'asymétrie au niveau de la constante du modèle GARCH(ASUG) ou sur son paramètre ARCH (BSUG) ne donne pas de différence significative par rapport aux estimations fournies par le modèle GARCH standard. Cependant, si l'asymétrie entre les chocs positifs et négatifs est modélisé à travers le paramètre GARCH (CSUG), la différence devient claire puisque $\hat{\beta}_1 > 0.96$ et $\hat{\alpha}_1 < 0.04$, voir figure 7.1.

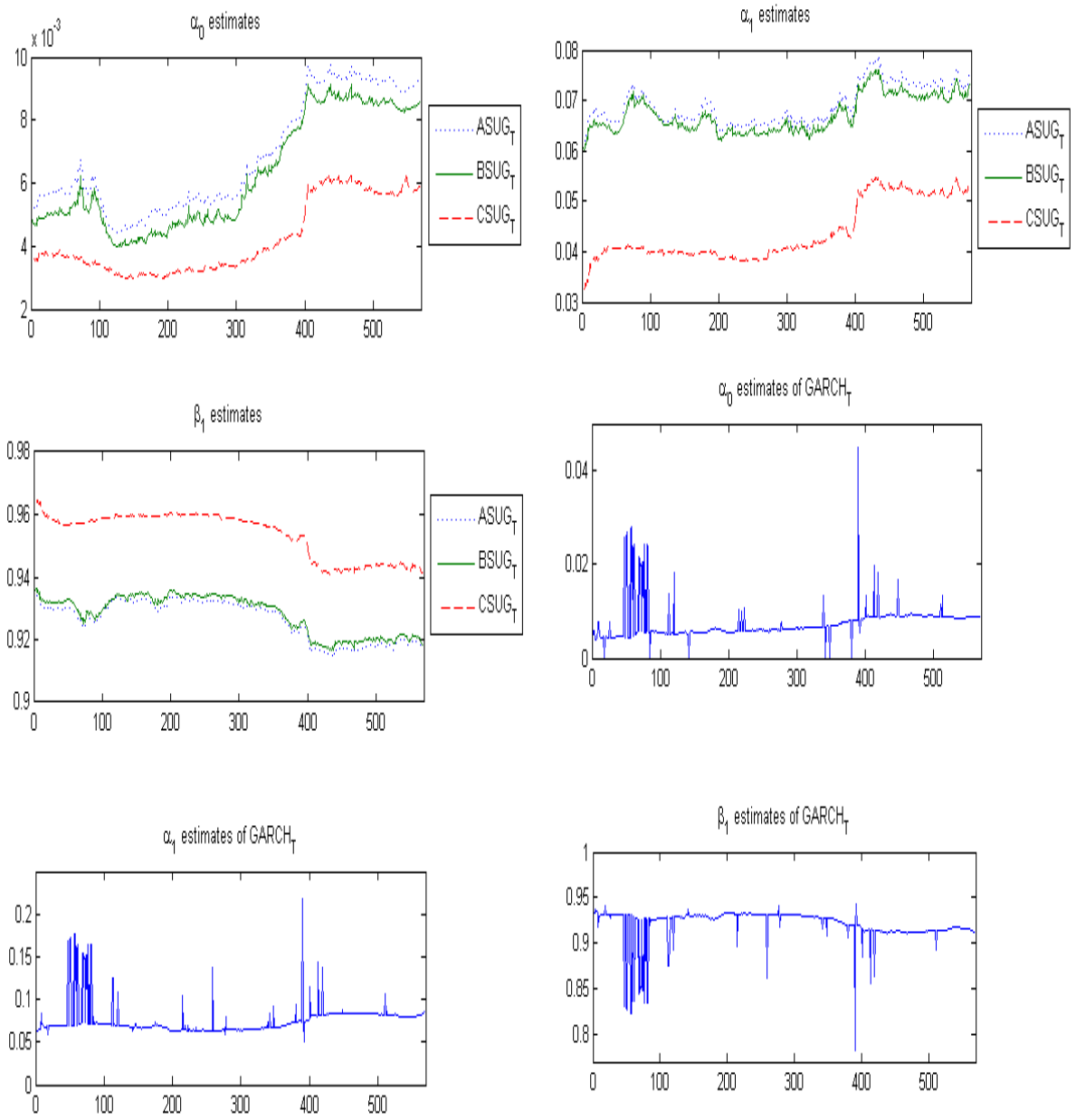

FIGURE 7.1 – Comparaison des paramètres estimés au cours du temps de quelques modèles ($SUGARH_T$ vs $GARCH_T$). Le nombre d'observation est de 568 allant de $T = 1699$ à $N = 2266$.

La figure ci-dessus illustre aussi la régularité des paramètres estimés au cours du temps pour la classe SUGARCH. Pour le montrer clairement, la figure 7.2 représente l'estimation de $(\alpha_0,\ \alpha_1,\ \beta_1)$ pour chaque modèle pour une période allant de T à N. On note que pour les autres modèles compétitifs, quelques oscillations apparaissent dans leur courbe alors que ce n'est pas le cas pour les modèles de la classe SUGARCH. Une explication possible est que, comme mentionné précédemment, la classe SUGARCH offre une certaine flexibilité et peut prendre en compte des changements structurels au niveau des données. [36]

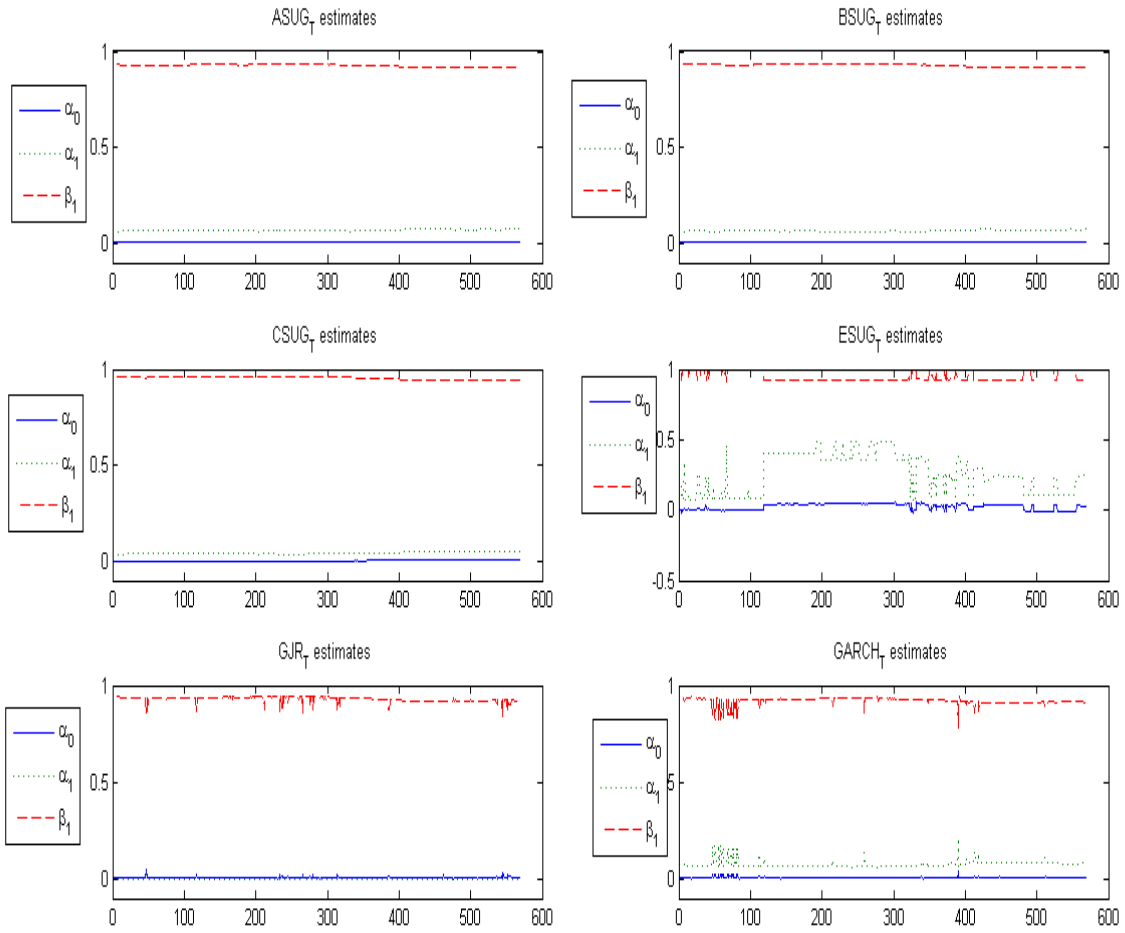

FIGURE 7.2 – Comparaison des trajectoires de $\alpha_0,\ \alpha_1,\ \beta_1$ $(t = T,\ T+1, \cdots,\ N)$ pour chaque modèle. η_t est supposée suivre une distribution de type Student.

36. Dans le code Matlab, nous avons utilisé, à l'itération courante, les paramètres optimaux trouvés à l'itération précédente comme condition initiale pour le programme d'optimisation. Cela peut expliquer aussi le fait que les courbes soient lisses.

Dans la table suivante, on compare la performance de prédiction des différents modèles sur l'échantillon in sample par rapport à la fonction de perte QLIKE mais aussi par rapport à d'autres critères standards tels que la log vraisemblance (LLF), l'$AIC = -2log(LLF)/T + 2k/T$, et le $BIC = -2log(LLF)/T + (k/T)log(T)$ où k représente le nombre de paramètres du modèle et T le nombre de données utilisées dans la phase d'estimation.

TABLE 7.3 – Comparaison de la précision des différents modèles (in sample).

Models	LLF	AIC	BIC	QLIKE
ASUG_N	-2287,243	2,698	2,714	2466834,969
ASUG_T	-2268,494	2,677	2,697	2470301,972
BSUG_N	-2279,304	2,689	2,705	2439856,628
BSUG_T	-2261,692	2,669	2,689	2443977,731
CSUG_N	-2256,800	2,663	2,679	2363387,826
CSUG_T	**-2240,184**	**2,6441**	**2,6633**	**2366316,219**
EGARCH_N	-2258,825	2,665	2,681	2370270,553
EGARCH_T	-2240,275	2,6442	2,6634	2375141,984
GJR_N	-2256,050	2,662	2,678	2360839,063
GJR_T	-2240,896	2,645	2,664	2367111,968
GARCH_N	-2287,273	2,697	2,710	2466936,088
GARCH_T	-2267,389	2,675	2,691	2473089,834

Note : Cette table présente les valeurs suivantes : la log vraisemblance (LLF), le critère d'information Akaike(AIC), le critère de Schwarz (BIC). QLIKE est la fonction de perte définie dans l'équation (7.14b). Les nombres en gras indiquent les meilleures valeurs.

La table 7.3 montre premièrement que la distribution de Student pour η_t donne de meilleurs résultats que la distribution normale. Les trois meilleurs modèles sont donnés respectivement par $CSUG_T$, $EGARCH_T$, GJR_T. Ce point illustre les résultats empiriques montrant que les rentabilités normalisées ($\eta_t = (r_t - \mu)/\sigma_t$) ont encore une distribution à queue épaisse, cependant, moindre par rapport à la rentabilité inconditionnelle ($r_t - \mu$).
D'un autre côté, introduire l'asymétrie à travers β_1 (modèle CSUG) au lieu de le faire sur α_0 ou α_1 (ASUG, BSUG) améliore la précision de la calibration dans la classe SUGARCH. Puisque le modèle CSUG a des similarités avec les autres modèles (A)GARCH dans le sens que la volatilité d'aujourd'hui dépend de celles passées (paramètres GARCH) et des innovations passées (paramètres ARCH), on peut espérer que la même technique marchera aussi pour ces modèles pour bien modéliser les données financières. La prochaine étape est de voir si cette approche de procéder (modification de β_1 au lieu de le faire sur α_1) donne de bons résultats pour les données hors échantillon.

Pour ce faire, on a comparé la prédiction de la volatilité faite par les différents modèles de trois manières. La première se base sur une métrique (fonction de perte), la seconde est basée sur la capacité de prédire la direction de la volatilité du marché alors que le troisième critère se focalise sur le management de risque. Les données hors échantillon couvrent la période du 3 Octobre 2008 au 31 Décembre 2010 et représentent approximativement 26 mois de données (567 observations). La table 7.4 montre cette comparaison pour le premier critère basé sur la fonction de perte QLIKE pour les horizons $h = 1,\ 2,\ 5,\ 10$. La vraie variance pour un horizon $h = m > 1$ est approximée par une somme de m rentabilités journalières élevées au carré. Une autre mesure (pas reportée) est donnée par la rentabilité élevée au carré définie sur un horizon $h = m$ i.e $(\ln S_{t+h} - \ln S_t)^2$.[37]

37. On a trouvé que ces deux mesures donnent un même ordonnancement entre les modèles compétitifs.

TABLE 7.4 – Evaluation de la prédiction de la volatilité sur les données hors échantillon pour $h = 1,\ 2,\ 5,\ 10$.

Models	h	QLIKE	h	QLIKE	h	QLIKE	h	QLIKE
ASUG_N	1	558631,2	2	872646,7	5	1827775,0	10	3446899,3
ASUG_T	1	559348,4	2	868371,1	5	1806952,0	10	3392459,9
BSUG_N	1	554388,8	2	865091,2	5	1810459,4	10	3427409,4
BSUG_T	1	555129,1	2	**859537,0**	5	**1782899,2**	10	3353046,9
CSUG_N	1	548193,3	2	913711,2	5	2014551,7	10	3909633,9
CSUG_T	1	549577,3	2	920363,9	5	2039330,2	10	3974431,8
EGARCH_N	1	560478,8	2	886325,6	5	1849211,3	10	3518372,6
EGARCH_T	1	561467,7	2	891629,5	5	1901609,6	10	3831994,2
GJR_N	1	**545049,6**	2	878955,3	5	1897807,8	10	3671101,7
GJR_T	1	545968,6	2	874280,9	5	1864692,4	10	3585347,4
GARCH_N	1	556849,4	2	872712,8	5	1838520,5	10	3485934,9
GARCH_T	1	557215,9	2	866013,9	5	1792292,3	10	**3346591,4**

Note : Cette table présente les valeurs de QLIKE voir (7.14b). Les nombres en gras donnent les meilleures valeurs.

Il est à noter qu'aucun modèle ne surperforme par rapport aux autres. Par exemple, pour les prédictions de la volatilité journalière ($h = 1$), le modèle GJR_N fournit les résultats les plus satisfaisants tandis que pour l'horizon $h = 10$, $GARCH_T$ est le meilleur modèle. Pour des horizons intermédiaires $h = 2,\ 5$, le modèle $BSUG_T$ est celui qui se comporte le mieux. Un autre point est que l'hypothèse de la distribution normale pour les innovations génère des résultats satisfaisants spécifiquement pour l'horizon journalier. Cependant, si l'horizon devient élevé, les valeurs minimales de la fonction QLIKE sont obtenus avec la distribution de Student. On note aussi qu'une performance au niveau des données in sample n'entraîne pas forcément de bons résultats au niveau des données hors échantillon. En effet, nous avons vu précédemment que CSUG_T était le meilleur modèle pour les données in sample alors qu'il fournit des résultats non satisfaisants pour les données hors échantillon et pour des horizons élevés. Ce point ne peut pas être expliqué par un phénomène de sur-apprentissage puisque tous les modèles asymétriques ont le même nombre de paramètres.[38] Notons que pour $h = 1$, le modèle CSUG_N continue à donner des résultats satisfaisants puisqu'il est le deuxième meilleur modèle après celui de GJR.

Le test de Diebold et Mariano, voir l'équation (7.17), est maintenant utilisé pour examiner la performance entre deux modèles compétitifs i et j sur le plan statistique plutôt que métrique. Les résultats sont disponibles à travers toutes les comparaisons deux à deux possibles et pour tous les horizons. Dans le tableau suivant, on présente que les résultats obtenus à travers les deux modèles $BSUG_T$ et $GARCH_T$ considérés comme benchmark et pour les horizons $h = 2,\ 10$, respectivement.[39]

38. Une raison peut être due au fait que le carré des rentabilités est un estimateur sans bias de la variance conditionnelle mais contenant beaucoup de bruits.

39. Pour des contraintes d'espace, tous les résultats ne sont pas inclus. Le restant peut être obtenu à partir du lien https ://sites.google.com/site/makonte/

TABLE 7.5 – Test de Diebold-Mariano avec GARCH_T et GJR_T pris comme référence.

Bench. BSUG_T	horizon	QLIKE	Bench. GARCH_T	horizon	QLIKE
ASUG_N	5	-2,71**	ASUG_N	10	-3,27**
P-values		(0,007)	P-values		(0,001)
ASUG_T	5	-1,37	ASUG_T	10	-1,43
P-values		(0,17)	P-values		(0,15)
BSUG_N	5	-5,80**	BSUG_N	10	-2,20*
P-values		(0,00)	P-values		(0,028)
CSUG_N	5	-4,96**	BSUG_T	10	-0,16
P-values		(0,00)	P-values		(0,87)
CSUG_T	5	-4,93**	CSUG_N	10	-5,09**
P-values		(0,00)	P-values		(0,00)
EGARCH_N	5	-0,64	CSUG_T	10	-4,99**
P-values		(0,52)	P-values		(0,00)
EGARCH_T	5	-1,07	EGARCH_N	10	-0,70
P-values		(0,29)	P-values		(0,49)
GJR_N	5	-4,29**	EGARCH_T	10	-1,81
P-values		(0,00)	P-values		(0,07)
GJR_T	5	-2,82**	GJR_N	10	-4,95**
P-values		(0,005)	P-values		(0,00)
GARCH_N	5	-3,58**	GJR_T	10	-3,22**
P-values		(0,00)	P-values		(0,00)
GARCH_T	5	-0,19	GARCH_N	10	-6,36**
P-values		(0,85)	P-values		(0,00)

Note : Bench. est mis pour Benchmark. * et ** représente les valeurs de la statistique DM pour qui l'hypothèse nulle est rejeté au seuil de signification 5% et 1%, respectivement.

Comme attendu, le test de Diebold et Mariano (DM) confirme les résultats obtenus avec la table 7.4 précédente. Si on considère le modèle GARCH_T comme benchmark, il est noté que les valeurs de toutes les statistiques sont négatives montrant que sa fonction de perte est la plus petite et donc la meilleure. D'un autre côté, la table 7.5 montre que l'hypothèse nulle d'une performance égale est rejetée pour les modèles compétitifs suivants : ASUG_N, BSUG_N, CSUG, GJR. Pour le restant des modèles, la différence entre les fonctions de perte n'est pas significative au seuil de 5%.
Si le modèle BSUG_T est maintenant considéré comme benchmark, des résultats similaires sont aussi obtenus. Le modèle performe en termes de prédiction de la volatilité par rapport aux autres modèles. Par exemple, l'hypothèse nulle est rejetée au seuil de signification 1% si la comparaison est faite avec les modèles GJR, *GARCH_N*, *ASUG_N*, *BSUG_N*, *CSUG*.
Le modèle EGARCH n'a jamais fait parti des meilleurs modèles mais il arrive à se positionner convenablement. On note que les meilleurs modèles pour un horizon donné n'ont pas donné une différence significative par rapport à ce modèle.

On investit maintenant l'aptitude des modèles à prédire l'état de la volatilité du marché. Il est supposé qu'à tout instant t, la variance est soit dans un niveau faible soit dans un niveau fort et leur frontière est donnée par la moyenne de la variance. La table suivante montre les résultats obtenus pour les horizons $h = 1$ et $h = 2$.

TABLE 7.6 – Valeurs pour D.A (Directional accuracy) et SR pour les différents modèles.

Models	horizon	DA	SR	horizon	DA	SR
ASUG_N	1	10,21**	0,79	10	17,13**	0,89
ASUG_T	1	10,21**	0,79	10	17,22**	0,89
BSUG_N	1	10,40**	0,80	10	17,75**	0,90
BSUG_T	1	10,54**	0,80	10	17,94**	0,91
CSUG_N	1	10,07**	0,78	10	17,84**	0,90
CSUG_T	1	10,00**	0,78	10	17,58**	0,90
EGARCH_N	1	8,65 **	0,76	10	14,80**	0,85
EGARCH_T	1	9,05 **	0,76	10	16,95**	0,90
GJR_N	1	9,98**	0,79	10	18,84**	0,92
GJR_T	1	9,81 **	0,79	10	18,74**	0,92
GARCH_N	1	10,47**	0,80	10	17,94**	0,91
GARCH_T	1	9,67 **	0,79	10	17,48	0,90

DA est la statistique donnée par l'Eq. (7.17) et est asymptotiquement normale. ** représente le rejet de l'hypothèse nulle au seuil de signification 1% et SR est défini dans l'Eq. (7.16).

La table 7.6 montre que tous les modèles ont la capacité de prédire l'état de la volatilité du marché. L'hypothèse nulle d'absence de pouvoir prédictif est rejeté au seuil de signification $\alpha = 1\%$. Il est à noter que la prédiction de l'état de la volatilité est plus précise pour un horizon de temps élevé. Pour $h = 10$, le ratio de réussite (SR) est au moins égal à 85% alors que pour $h = 1$, la valeur la plus élevée est égale à 80%. Des résultats satisfaisants sont obtenus en général des modèles GJR, BSUG et GARCH. Etant donné la capacité de prédire l'état de la volatilité du marché, on peut se demander si on peut en tirer un gain économique. Cette question, relative à l'efficience de marché ne se sera pas investie, voir par exemple Noh et al. [91] et les références qui y sont citées.

Enfin, on termine cette étude en comparant la performance des différents modèles en termes de prédiction de certains indicateurs relatifs au management de risque. Ici, les fonctions de pertes sont données respectivement par les équations (7.18a), (7.18b). Pour la détermination du quantile VaR, on a pris $\alpha = 0.01$ avec une distribution de η_t Normale ou de type Student. Le résultat général est que tous les modèles ont des problèmes pour prédire la VaR quand l'horizon n'est pas proche (une ou 2 semaines). Si par contre l'horizon est rapproché, par exemple un ou 2 jours, les résultats sont satisfaisants. Pour l'objet du management de risque, nous analysons un des deux cotés (pertes) de la distribution conditionnelle des rentabilités car si un investisseur prend une position longue (courte), il n'y aura que la gestion des rentabilités extrêmes négatives (positives) qui va l'intéresser.

TABLE 7.7 – Critère d'évaluation sur données hors échantillon : 99 % VaR, Position longue.

Models	horizon	PF	KB	horizon	PF	KB
ASUG_N	1	2,4691E-02	4,6547E-02	2	5,300E-02	9,259E-02
ASUG_T	1	5,2910E-03	4,9157E-02	2	3,004E-02	6,844E-02
BSUG_N	1	2,2928E-02	**4,6344E-02**	2	4,947E-02	8,840E-02
BSUG_T	1	3,5273E-03	4,9032E-02	2	3,004E-02	**6,714E-02**
CSUG_N	1	3,5273E-02	5,7917E-02	2	6,360E-02	1,165E-01
CSUG_T	1	1,4109E-02	5,2039E-02	2	4,770E-02	9,325E-02
EGARCH_N	1	3,1746E-02	5,8107E-02	2	6,184E-02	9,952E-02
EGARCH_T	1	7,0547E-03	5,8940E-02	2	**2,120E-02**	7,341E-02
GJR_N	1	2,6455E-02	4,7136E-02	2	5,477E-02	9,031E-02
GJR_T	1	**8,8183E-03**	4,8898E-02	2	3,180E-02	6,835E-02
GARCH_N	1	2,9982E-02	4,7154E-02	2	5,477E-02	9,296E-02
GARCH_T	1	5,2910E-03	4,9534E-02	2	2,827E-02	6,793E-02

Note : Cette table présente le pourcentage d'échecs (PF) et la fonction de perte de Koenker et Bassett (KB) pour la VaR au seuil de 99 % et pour les horizons $h = 1$, $h = 2$. Les nombres en gras montrent les meilleurs valeurs.

La table 7.7 montre que pour des positions longues et $h = 1$, le modèle GJR_T donne la meilleure estimation de la valeur théorique pour α. EGARCH_T est le second meilleur modèle. Ce dernier devient le meilleur pour l'horizon $h = 2$ jours suivi par le modèle GARCH_T. Puisque le pourcentage d'échecs (PF) ne tient pas en compte la magnitude de la violation de la VaR (même poids), il est plus convenable de travailler avec la fonction de perte KB pour remédier à cet inconvénient. Dans ce cas, les violations de la VaR sont pondérées par $1 - \alpha$ tandis que les non violations le sont par α. Vu que α est très petit en pratique de l'ordre de 1% à 5%, la fonction de pertes KB pénalise lourdement les mauvaises prédictions. Pour ce nouveau critère, on a trouvé que le meilleur modèle est maintenant BSUG_N suivi par ASUG_N pour $h = 1$ tandis que pour $h = 2$ on trouve les modèles BSUG_N et GARCH_T. Il est à noter que pour $h = 2$, tous les modèles génèrent un nombre de violations supérieur à la valeur théorique 1%. ce qui signifie qu'il est préférable d'utiliser ces modèles pour des horizons journaliers ($h = 1$). La dernière table montre les résultats obtenus pour une position courte.

TABLE 7.8 – Critère d'évaluation sur données hors échantillon : 99 % VaR, Position courte.

Models	horizon	PF	KB		PF	KB
ASUG_N	1	1,23E-02	4,79E-02	2	3,18E-02	6,05E-02
ASUG_T	1	3,53E-03	4,94E-02	2	**1,06E-02**	5,12E-02
BSUG_N	1	**1,06E-02**	4,50E-02	2	2,65E-02	5,36E-02
BSUG_T	1	0	4,80E-02	2	7,07E-03	5,11E-02
CSUG_N	1	7,05E-03	**4,03E-02**	2	2,65E-02	4,61E-02
CSUG_T	1	5,29E-03	4,29E-02	2	8,83E-03	**4,44E-02**
EGARCH_N	1	1,41E-02	4,60E-02	2	2,65E-02	5,52E-02
EGARCH_T	1	1,76E-03	5,26E-02	2	3,53E-03	5,29E-02
GJR_N	1	1,41E-02	4,37E-02	2	2,47E-02	4,98E-02
GJR_T	1	3,53E-03	4,79E-02	2	7,07E-03	4,86E-02
GARCH_N	1	**1,06E-02**	4,54E-02	2	3,00E-02	5,59E-02
GARCH_T	1	0	4,93E-02	2	7,07E-03	5,09E-02

Note : Cette table présente le pourcentage d'échecs (PF) et la fonction de perte de Koenker et Bassett (KB) pour la VaR au seuil de 99 % et pour les horizons $h = 1$, $h = 2$. Les nombres en gras montrent les meilleurs valeurs.

Pour les prédictions journalières de la VaR, les meilleurs résultats pour le critère PF sont donnés par les modèles BSUG_N et GARCH_N. Pour le même horizon toujours avec maintenant la fonction de perte KB, CSUG_N devient le modèle le plus performant. Maintenant pour l'horizon $h = 2$, on trouve que CSUG_T est le modèle qui minimise la fonction de perte KB. Précédemment, ce modèle avait donné aussi des résultats satisfaisants pour les données in sample. Pour le critère PF, le meilleur modèle est ASUG_T suivi par GJR_T. En résumé, on note que la classe SUGARCH a donné des résultats intéressants aussi bien pour le critère PF que celui de KB pour une position courte considérée.

Chapitre 8

Conclusion et discussion

Dans ce document, on est revenu sur la formation des prix d'équilibre des actifs financiers. D'abord, on a rappelé de façon brève les fondements de la finance néo-classique et celle comportementale. Bien qu'une différence claire apparaisse au niveau micro dans la formulation de ces deux théories, on a du mal à cerner la question de l'efficience au niveau macro-économique. Ce problème est souvent attribué à l'hypothèse de la loi jointe. Tester l'efficience de marché, c'est spécifier également un modèle d'équilibre de sorte quand l'hypothèse nulle est rejetée, on se retrouve dans un dilemme. La difficulté peut être due aussi aux approches méthodologiques. On a illustré que sur de simples transformations (choix de la fréquence des données, choix de la période d'études, choix du jour pour une fréquence donnée), on pouvait faire basculer les conclusions d'une étude d'un camp à l'autre.

Notre lecture de cette difficulté a été que les prix des actifs intègrent à la fois les deux types d'arguments. C'est pourquoi il est problématique de dire quel paradigme a raison. Pour le prouver, on a proposé un modèle théorique qui simule une violation de la loi du prix unique pour éviter les questions de l'hypothèse jointe et de méthodologie. Précisément, on se donne une diffusion continue quelconque pour un premier actif A et une diffusion similaire pour l'autre actif B à la quelle on rajoute une partie saut d'espérance nulle. On montre, dans ce cadre, qu'il est difficile de dire encore si le marché est efficient ou pas bien qu'ayant écarté les points à discussion. Ce résultat nous renseigne en même temps que les modèles de diffusion avec une partie saut d'espérance nulle sont convenables pour modéliser les prix des actifs car réconciliant les deux mouvances.

En deuxième lieu, on a proposé un modèle à agents qui intègre aussi bien les arguments de la finance néo-classique que comportementale, puis on a déterminé les prix d'équilibre sous certaines simplifications. La dynamique en log des prix appartient à la classe des modèles auto-régressifs à coefficients aléatoires RCA(p) et ayant une racine unitaire en moyenne. Cette dynamique a été reliée aussi aux processus de type ARCH dans le cas simple où la volatilité courante est approximée par la valeur absolue de la dernière rentabilité. Les simulations ont montré que les principaux faits stylisés des séries financières sont reproduits indépendamment du paramètre p. Bien que ce point soit important, ce qui intéresse le plus les opérateurs économiques est de pouvoir prédire la direction des prix. C'est pourquoi, on s'est intéressé à un algorithme utilisant le modèle RCA(1) et qui puisse fournir une performance supérieure à celle de ne détenir que l'indice. On a proposé une méthode d'estimation (méthode des moindres carrées) dynamique qui n'utilise que peu de données pour éviter d'accorder de l'importance aux données anciennes. Précisément, une fenêtre glissante de n données avec $4 \leq n \leq 40$ a été prise dans la phase de l'estimation puis une prédiction d'un pas est faite. Les résultats obtenus sont assez intéressants dans le cas de l'indice CAC 40 où on arrive à faire mieux que la stratégie Achat et Conservation. Pour le S&P, la stratégie de la classe RCA a donné aussi la meilleure rentabilité avec cependant une médiane faible indiquant que l'on a capté plutôt la direction des forts mouvements de prix. Un test d'égalité de médiane à montrer qu'aucune stratégie (incluant les techniques de type moyennes mobiles) n'a pu faire que l'indice américain. Ce résultat conforte l'idée de l'efficience en moyenne où on peut battre tantôt le marché mais sans pourtant conserver une persistance dans la durée.

Une deuxième application a été d'étendre le modèle GARCH standard en permettant ses coefficients d'être stochastiques pour tenir en compte la complexité de la dynamique des prix d'actifs financiers. Le nouveau modèle qui y résulte devait être proche du modèle GARCH initial qui a l'avantage d'être simple et d'avoir un pouvoir prédictif assez intéressant d'après les conclusions de beaucoup d'études empiriques. Ceci nous a amené à la classe SUGARCH(P,Q) qui contient des modèles se comportant comme un GARCH standard 'en moyenne'. Les résultats obtenus sont satisfaisant dans l'ensemble avec trois critères différents : l'un utilisant la métrique QLIKE, l'autre un test statistique d'égalité EPA et enfin un dernier qui s'intéresse à des fonctions de pertes liées au management de risque.

Bibliographie

[1] D. Abreu and M.K. Brunnermeier. Synchronization risk and delayed arbitrage. *Journal of Financial Economics*, 66 :341–360, 2002.

[2] S. Alfarano. *An agent-based stochastic volatility model.* PhD thesis, Department of Economics, Christian-Albrechts University of Kiel, Germany, 2006.

[3] S. Alfarano and T. Lux. A minimal noise trader model with realistic time series. Economics working papers 317, Christian-Albrechts-University of Kiel, Departement of Economics, 2006.

[4] M. Allais. Le comportement de l'homme rationnel devant le risque : Critique des postulats et axiomes de l'école américaine. *Econometrica*, 22 :503–546, 1953.

[5] T. G. Andersen and T. Bollerslev. Answering the skeptics : Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39 :885–905, 1998.

[6] W. B. Arthur, J. H. Holland, B., R. Palmer, and P. Talyer. Asset pricing under endogenous expectations in an artificial stock market. *Economic Notes*, 26 :297–330, 1997.

[7] L. Bachelier. Théorie de la spéculation. *Annales Scientifiques de L'Ecole Normale Supérieure*, 17 :21–88, 1900.

[8] I. Bajeux-Besnaino and R. Portait. The numeraire portfolio : A new perspective on financial theory. *The European Journal of Finance*, 3 :291–309, 1997.

[9] R. Banz. Investment performance of common stocks in relation top their price earnings ratios : A test of market efficiency. *Journal of Finance*, 32 :663–682, 1981.

[10] N. Barberis, M. Huang, and T. Santos. Prospect theory and asset prices. *The quarterly journal of economics*, 116 :1–53, 2001.

[11] N. Barberis, A. Shleifer, and R. W. Vishny. A model of investor sentiment. *The Journal of Financial Economics*, 49 :307–343, 1998.

[12] D. Bernoulli. Exposition of a new theory on the measurement of risk, (original 1738). *Econometrica*, 22 :23–36, 1954.

[13] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Political Economy*, 81 :637–654, 1973.

[14] T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31 :307–327, 1986.

[15] De Bondt and R. Thaler. *Financial decision-making in markets : a behavioral perspective» in R. Jarrow, V. Maksimovic and W. Ziemba.* Finance, Handbooks in Operations Research and Management Science, 9 , Amsterdam, 1995.

[16] W.F.M. De Bondt and R. Thaler. Does the stock market overreact ? *Journal of Finance*, 40 :793–805, 1985.

[17] W. A. Brock and C. H. Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22 :1235–1274, 1998.

[18] C. Brownlees, R. Engle, and B. Kelly. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1) :246–256, 2010.

[19] T. Chauveau and N. Nalpas. *A theory of Disappointment.* working paper, 2009.

[20] S. H. Chen and C. H. Yeh. Genetic programming and the efficient market hypothesis. In John R. Koza, David E. Goldberg, David B. Fogel, and Rick L. Riolo, editors, *Genetic Programming 1996 : Proceedings of the First Annual Conference*, pages 45–53, Stanford University, CA, USA, 28–31 July 1996. MIT Press.

[21] S. H. Chen and C. H. Yeh. Evolving traders and the business school with genetic programming : A new architecture of the agent-based artificial stock market. *Journal of Economic Dynamics and Control*, 25 :363–394, 2001.

[22] P. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39(4) :841–862, 1998.

[23] J. H. Cochrane. *Asset Pricing(Revised)*. Princeton University Press., 2005.

[24] R. Cont. *Volatility Clustering in Financial Markets : Empirical Facts and Agent-Based Models*. Long memory in economics, A Kirman & G Teyssiere (eds), Springer, 2005.

[25] M. Cooper and H. Gulen. Is time-series-based predictability evident in real time? *Journal of Business*, 79 :1263–1292, 2006.

[26] J. M. Courtault, F. Delbaen, Y. Kabanov, and C. Stricker. On the law of one price. *Finance and Stochastics*, 5 :525–530, 2004.

[27] J.C. Cox, J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2) :385–407, 1985.

[28] R.C. Dalang, A. Morton, and A. Willinger. Equivalent martingale measures and no-arbitrage criteria in stochastic securities market model. *Stochastics and Stochastic Reports*, 29 :185–201, 1990.

[29] R. Damasio. *Descartes' Error : Emotion, Reason, and the Human Brain*. Harper Perennial, 1994.

[30] K. Daniel, D. Hirshleifer, and A. Subrahmanyam. Investor psychology and security market under- and overreactions. *Journal of Finance*, 53(6) :1839–1885, 1998.

[31] P. Dechow, R. Sloan, and M. Soliman. Implied equity duration : a new measure of equity security risk. *Review of Accounting Studies*, 9 :197–228, 2004.

[32] F. Delbaen and W. Schachermayer. The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312 :215–250, 1998.

[33] F. Delbaen and W. Schachermayer. *The Mathematics of Arbitrage*. Springer Finance. 1 edition, 2006.

[34] F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3) :253–263, 1995.

[35] R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of u.k. inflation. *Econometrica*, 50 :987–1008, 1982.

[36] E. F. Fama. Random walks in stock market prices. *Financial Analysts Journal*, 21 :55–60, 1965.

[37] E.F. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49 :283–306, 1998.

[38] E.F. Fama and S. French. The cross-section of expected stock returns. *Journal of Finance*, 47 :427–465, 1992.

[39] K. French. Stock returns and the weekend effect. *Journal of Financial Economics*, 8(1) :55–69, 1980.

[40] M. Friedman. *The case for flexible exchange Rates, Essays in positive Economics*. University of Chicago Press, 157-203, 1953.

[41] K. Froot and E. Dabora. How are stock prices affected by the location of trade. *Journal of Financial Economics*, 53 :189–216, 1999.

[42] Nicolas Gaussel. *Problèmes choisis pour l'intégration et la transmission d'information par les marchés financiers*. PhD thesis, Université Paris Panthéon Sorbonne (Paris 1), 2002.

[43] F. Ghoulmie, R. Cont, and J.-P. Nadal. Heterogeneity and feedback in an agent based market model. *Journal of Physics : Condensed Matter*, (17) :1259–1268, 2005.

[44] I. Gilboa. *Theory of Decision under Uncertainty*. Lectures notes for a graduate-level course, 2008.

[45] L. Glosten, R. Jagannathan, and D. Runkle. On the relationship between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance.*, 48 :347–370, 1993.

[46] D. K. Gode and S. Sunder. Allocative efficiency of markets with zero intelligence traders. *Journal of Political Economy*, 101 :119–137, 1993.

[47] W. Goetzmann and N. Peles. Cognitive dissonance and mutual fund investors. *Journal of financial Research*, 20 :145–158, 1997.

[48] C. W Granger and N. R. Swanson. An introduction to stochastic unit-root processes. *Journal of Econometrics*, 80 :35–62, 1997.

[49] S. J Grossman and J. E.Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review, American Economic Association*, 70(3) :393–408, 1980.

[50] R. Grothmann. *Multi-agent market modeling based on neural networks*. PhD thesis, Faculty of Economics, University of Bremen, 2002.

[51] M. Haas, S. Mittnik, and M. S. Paolella. Mixed normal conditional heteroskedasticity.

Journal of Financial Econometrics, 2 :211–250, 2004.

[52] P. R. Hansen. A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23(4) :365–380, 2005.

[53] P. R. Hansen and A. Lunde. A forecast comparison of volatility models : Does anything beat a garch(1,1). *Journal of Applied Econometrics*, 20(7) :873–889, 2005.

[54] J. M. Harrison and D.M. Kreps. Martingales and arbitrage in multiperiod security markets. *Journal of Economic Theory*, 20 :381–408, 1979.

[55] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2) :327–343, 1993.

[56] D. Hirshleifer and T. Shumway. Good day shunshine : stock returns and the weather. *Journal of Finance*, 58(3) :1009–1033, 2003.

[57] A. Hoffmann, W. Jager, and J. von Eije. Combining empirical data with multi agent social simulation : Investigating micro - macro links in stock markets. `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=898826`, 2006.

[58] G. Iori. A micro-simulation traders ' activity in the stock market : the rule of heterogeneity, agents ' interactions and trade friction. *Journal of Economic Behavior and Organisation*, (49) :269–285, 2002.

[59] N. Jegadeesh and S. Titman. Returns to buying winners and selling losers : implication for stock market efficiency. *Journal of Finance*, 48 :65–91, 1993.

[60] N. Jensen. An introduction to bernoullian utility theory : (1) utility functions. *Swedish Journal of Economics*, 69 :163–183, 1967.

[61] G. Jiang, M. C. Lee, and Y. Zhang. Information uncertainty and expected returns. *Review of Accounting Studies*, 10 :185–221, 2005.

[62] E. Jouini and C. Napp. Consensus consumer and intertemporal asset pricing with heterogeneous beliefs. *The Review of economic Studies*, 74(4) :1149–1174, 2007.

[63] D. Kahneman and A. Tversky. Judgment under uncertainty : Heuristics and biases. *Science*, 185 :1124 – 1131, 1974.

[64] D. Kahneman and A. Tversky. Prospect theory : An analysis of decision under risk. *Econometrica*, 47 :263–291, 1979.

[65] A. Kamara and T. Miller. Daily and intradaily tests of european put call parity. *Journal of Financial and Quantitative Analysis*, 30 :519–539, 1995.

[66] F. Klaassen. Improving garch volatility forecasts. *Empirical Economics*, 27 :363–394, 2002.

[67] R. Koenker and G. Bassett. Regression quantiles. *Econometrica*, 46 :33–50, 1978.

[68] M. Konte. A link between random coefficient autoregressive models and some agent based models. *Journal of Economic Interaction and Coordination*, 6(1) :83–92.

[69] M. Konte. Behavioural finance and efficent markets : Is the joint hypothesis really the problem ? *IUP Journal of Behavioral Finance*, 7 :9–19, 2010.

[70] D.M Kreps. Arbitrage and equilibirum in economies with infinitely commodities. *Journal of Mathematical Economics*, 8 :15–35, 1981.

[71] M. Kurz and M. Moteleze. Endogenous uncertainty and market volatility. *Economic Theory*, 17 :497–544, 2001.

[72] O. Lamont and R. Thaler. Can the market add and subtract ? mispricing in tech stock carveouts. *Journal of Political Economy*, 111 :227–268, 2003.

[73] L.Deville and F. Riva. Liquidity and arbitrage in options markets : A survival analysis approach. *Review of Finance*, 11(3), 2007.

[74] B. LeBaron. Agent-based computational finance. Handbook of Computational Economics, Volume 2 : Agent-Based Computational Economics. Elsevier Science B.V, 2005.

[75] B. LeBaron, W. B. Arthur, and R. Palmer. Time series properties of an artificial stock market. *Journal of Economic Dynamics and Control*, 23 :1487–1516, 1999.

[76] S. L. LeRoy. Risk aversion and the martingale property of stock prices. *International Economic Review*, 14(2) :436–446, 1973.

[77] H. Levy, M. Levy, and S. Solomon. *Microscopic Simulation of Financial Markets : From Investor Behavior to Market Phenomena*. Academic Press, 2000., 2000.

[78] M. Levy, H. Levy, and S. Solomon. A microscopic model of the stock market : Cycles, booms, and crashes. *Economic Letters*, 94(103) :103–111, 1994.

[79] A. Lo. Reconciling efficient markets with behavioral finance : The adaptive markets hypothesis. *Journal of Investment Consulting*, 7(2) :21–44, 2005.

[80] A. Lo and A. C. MacKinlay. Stock market prices do not follow random walks : Evidence from a simple specification test. *Review of Financial Studies*, 1(1) :41–66, 1988.

[81] J. B De Long, A. M. Shleifer, L. H. Summers, and R. J. Waldmann. Noise trader risk in financial markets. *Journal of Political Economy*, 98 :703–738, 1990.

[82] J. M. Machina. Expected utility theory without the independence axiom. *Econometrica*, 50 :277–323, 1982.

[83] M. Mangot. *Psychologie de l'investisseur et des marchés financiers.* 2008.

[84] J. Marcucci. Forecasting stock market volatility with regime-switching garch models. *Studies in Nonlinear Dynamics & Econometrics*, 9(4) :1–55, 2005.

[85] S. Martinez-Jaramillo. *Artificial Financial Markets : An Agent Based Approach to Reproduce Stylized Facts and to study the Red Queen effect.* PhD thesis, Centre for Computational Finance and Economic Agents, UNIVERSITY OF ESSEX, 2007.

[86] A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory.* New York : Oxford Press, 1995.

[87] M. Mitchell, T. Pulvino, and E. Stafford. Limited arbitrage in equity markets. *Journal of Finance*, 57 :551–584, 2002.

[88] D. Nagakura. Testing for coefficient stability of ar(1) model when the null is an integrated or a stationary process. *Journal of statistical planning and inference*, 139(8) :2731–2745, 2009.

[89] D. Nelson. Conditional heteroskedasticity in asset returns : a new approach. *Econometrica*, 59 :347–370, 1991.

[90] D. F Nicholls and B. G. Quinn. *Random Coefficient Autoregressive Models : An Introduction,.* Lecture Notes in Statistics 11, New York : Springer-Verlag, 1982.

[91] J. Noh, R. F. Engle, and A. Kane. Forecasting volatility and option prices of the s & p 500 index. *Journal of Derivatives*, pages 17–30, 1994.

[92] T. Odean. Do investors trade too much. *American economic review*, 89 :1279–1298, 1999.

[93] E. Ofek, M. Richardson, and R. Whitelaw. Limited arbitrage and short sales restrictions : evidence from the options markets. *Journal of Financial Economics*, 74 :305–342, 2004.

[94] A. Patton. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1) :246–256, 2010.

[95] M. H. Pesaran and A. Timmermann. A simple nonparameteric test of predictive performance. *Journal of Business and Economic Statistics*, 10(4) :461–465, 1992.

[96] D. Prelec. The probability weighting function. *Econometrica*, 66 :497 – 527, 1998.

[97] S. A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13 :341–60, 1976.

[98] S. A. Ross. Return, risk, and arbitrage. *in I. Friend and J. Bicksler, eds., Risk and Return in Finance 1, Cambridge, Mass Ballinger*, pages 189–218, 1976.

[99] B. Routledge. Adaptive learning in financial markets. *Review of Financial Studies*, 12(5) :1165–1202, 1999.

[100] E Samanidou, E Zschischang, D Stauffer, and T Lux. Agent-based models of financial markets. *Reports on progress in physics*, (70) :409–450, 2007.

[101] P. A. Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6 :41–49, 1965.

[102] A.-H. Sato and M. Takayasu. Derivation of arch(1) process from market price changes based on deterministic microscopic multiagent. *Proceedings Empirical science of financial fluctuations, Springer(Tokyo), Ed. by H. Takayasu*, pages 172–178, 2002.

[103] L. J. Savage. *The foundation of Statistics.* New York : John Wiley and Sons, 1954.

[104] A. Shleifer and L. Summer. The noise trader approach to finance. *Journal of Economic Perspectives*, (4) :19–33, 1990.

[105] A. Shleifer and R. Vishny. The limits of arbitrage. *Journal of Finance*, 52 :35–55, 1997.

[106] R. Sollis, S.J. Leybourne, and P. Newbold. Stochastic unit roots modelling of stock price indices. *Applied financial economics*, 10 :311–315, 2000.

[107] R. Sullivan, A. Timmermann, and H. White. Data-snopping , technical trading rule performance, and the boostrap. *The journal of finance*, 54 :1647–1691, 1999.

[108] H. Takayasu and M. Takayasu. Critical fluctuations of demand and supply. *Physica A : Statistical Mechanics and its Applications*, 269 :24–29, 1999.

[109] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *Science*, 211 :253–458, 1981.

[110] A. Tversky and D. Kahneman. Advances in prospect theory : Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4) :297 – 323, 1992.

[111] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, 2nd edition, 1947.

[112] H. White. A reality check for data snooping. *Econometrica*, 68 :1097–1126, 2000.

[113] G. Yoon. A simple model that generates stylized facts of returns. Department of Economics, UCSD. Paper 2003-04., 2003.

[114] K. Yuan, L. Zhenz, and Q. Zhu. Are investors moon struck ? - lunar phases and stock returns. *Journal of Empirical Finance*, 13(1) :1–23, 2006.

[115] J. M. Zakoian. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18 :931–955, 1994.

[116] X. Zhang. information uncertainty and stock returns. *The Journal of finance*, 61(1) :105–137, 2006.