



HAL
open science

Le phénomène leptokurtique sur les marchés financiers

Christian Walter

► **To cite this version:**

Christian Walter. Le phénomène leptokurtique sur les marchés financiers. Finance, 2009, 23 (2), pp.15-68. hal-04530006

HAL Id: hal-04530006

<https://hal.science/hal-04530006>

Submitted on 2 Apr 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Le phénomène leptokurtique sur les marchés financiers

Christian Walter

PRICEWATERHOUSECOOPERS – UNIVERSITÉ D'EVRY
32, rue Guersant – 75017 Paris
tél. 01 56 57 85 96
fax 01 56 57 24 09
e-mail : christian.walter@fr.pwcglobal.com

version revue et corrigée¹
11 avril 2009

Résumé

La non normalité des distributions des variations boursières est devenue un fait d'expérience omniprésent dans la finance de marché. Le but de cet article est de présenter un état de la question en mettant en évidence certains aspects de la non normalité encore peu explicités. On montre que le phénomène leptokurtique a toujours été présent dans les travaux de recherche sur les fluctuations boursières, mais qu'il n'a acquis que récemment une reconnaissance institutionnelle. On fait apparaître la perte de pertinence de la notion de moyenne dans les pratiques professionnelles de couverture et de gestion du risque, en présentant un nouveau type de risque de modèle. On montre la relation existant entre la forme des processus d'investissement dans la gestion d'actifs avec la forme des distributions des rentabilités, en présentant une conjecture sur la revalorisation de l'importance des choix de titres comme conséquence du phénomène leptokurtique. On résume enfin les différentes compréhensions actuelles de la non normalité en proposant un point de vue distinguant l'information exogène de l'information endogène, pour introduire des considérations sur la microstructure des marchés et la mesure du temps.

¹Je remercie les rapporteurs pour les commentaires très éclairants qui ont été faits et les conseils qui ont été donnés sur les deux premières versions de l'article, ainsi que Monique Jeanblanc et Jacques Lévy-Véhel pour des discussions fructueuses concernant certaines des questions développées ici. La clarté finale de la présentation doit beaucoup à Georges Gallais-Hamonno, dont les très fécondes suggestions ont rendu la lecture du texte plus agréable et l'enchaînement des parties plus logique.

Table des matières

1	Introduction	4
2	Mise en évidence de la leptokurticité : naissance d'un fait expérimental	7
2.1	Historicité du problème leptokurtique	7
2.1.1	Ancienneté du phénomène leptokurtique	7
2.1.2	Permanence du phénomène leptokurtique	8
2.2	Reconnaissance du problème leptokurtique	10
2.2.1	Evolution de l'environnement professionnel	10
2.2.2	Apport des données de haute fréquence	10
2.3	Modélisation du phénomène leptokurtique	11
2.3.1	Termes du débat sur la modélisation	11
2.3.2	Caractère parétien des queues de distribution et indice de queue	12
2.3.3	La quête du processus des rentabilités	15
3	Importance pratique du phénomène leptokurtique : la nécessaire gestion des queues de distribution	17
3.1	Queues de distribution, modèles de risque et risque de modèle	17
3.1.1	Poids des grandes variations et perte de pertinence des moyennes	17
3.1.2	Risque de modèle dû aux queues de distribution parétiennes	21
3.2	Impact des queues de distribution sur la VaR	23
3.2.1	La VaR et les grandes variations	23
3.2.2	Exemple de calcul de VaR avec des distributions α -stables	25
3.3	Impact des queues de distribution sur la gestion d'actifs	28
3.3.1	La concentration de la performance	28
3.3.2	La revalorisation de l'importance des choix de titres	31
4	Interprétations théoriques du phénomène leptokurtique : développement de nouveaux modèles	34
4.1	Leptokurticité, information et mimétisme	36
4.1.1	Leptokurticité d'origine externe au marché : l'économie des extrêmes	36
4.1.2	Leptokurticité d'origine interne au marché : l'interaction extrême	37
4.2	Leptokurticité, durée du temps, et valeur du temps	38
5	Conclusion	41
A	Annexe : calcul numérique de la fonction de répartition d'une loi α-stable	42
B	Annexe : tabulations de la fonction de répartition d'une loi α-stable	44

Liste des tableaux

1	Les travaux précurseurs de Mandelbrot et Fama sur la leptokurticité	8
2	Permanence de la leptokurticité sur le marché des actions	8
3	Permanence de la leptokurticité sur le marché des changes	9
4	Décroissance du coefficient de kurtosis K	9
5	Quantiles élevés q_C^α en fonction de α	27
6	Construction de la courbe de Gini-Lorenz de concentration de la performance	29

7	La profession de la gestion d'actifs et le phénomène leptokurtique.	35
8	Valeurs critiques x_c en fonction de α	43
9	Fonction de répartition F_α en fonction de α	45
10	Probabilité de perte : $\Pr(X > x) = 1 - F_\alpha(x)$ en fonction de x et α	46
11	Durée de retour $d_\alpha(x) = 1/(1 - F_\alpha(x))$ en fonction de x et α	47

Table des figures

1	Variations positives de l'indice CAC 40	19
2	Variations négatives de l'indice CAC 40	19
3	Densités des lois α -stables.	25
4	VaR en fonction de α	28
5	Concentration de la performance positive du CAC 40	29
6	Concentration de la performance négative du CAC 40	29
7	Concentration de la performance totale du CAC 40	30

1 Introduction

Soit $\{S(t), t \geq 0\}$ le processus d'évolution du cours d'un actif S . Le *modèle standard* de la finance de marché fait l'hypothèse que la dynamique du cours $S(t)$ peut être décrite par l'équation différentielle stochastique (EDS) linéaire

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t) \quad (1)$$

où $W(t)$ est un mouvement brownien, avec $W(0) = 0$, $E(W_1) = 0$ et $E(W_1^2) = 1$, équation qui a pour solution

$$S(t) = S(0) \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W(t) \right\}$$

et qui définit un mouvement brownien géométrique sur les cours des actifs.

Associé à une interprétation économique utilisant la notion de transmission d'information dans et par les cours, le mouvement brownien géométrique devint le coeur probabiliste du paradigme qui a longtemps été généralement admis en finance, largement véhiculé par un très vaste ensemble de pratiques professionnelles et de théories successives², et qui est une description du comportement des marchés au moyen d'un concept fondateur et structurant, celui de l'efficacité informationnelle. Comme le soulignent Jacquillat et Solnik [1997, p.48], l'efficacité informationnelle constitue le fondement de la théorie financière moderne. Concept aux contours fluctuants, pressenti par Cowles, esquissé par Hayek, l'efficacité informationnelle dans sa forme moderne a été successivement définie par Fama [1970] pour sa forme initiale (information disponible), à nouveau par Fama [1976] pour une seconde forme (information disponible et pertinente), puis par Jensen [1978] qui introduisit les coûts de transactions, pour être enfin précisée dynamiquement dans les travaux de Grossman et Stiglitz [1980]. L'association établie dès l'origine entre ce processus gaussien et le concept d'efficacité informationnelle a conduit à une forme particulière d'efficacité, que l'on a proposé d'appeler (Walter [1994]) "efficacité gaussienne".

Très tôt, c'est-à-dire dès les années soixante, un grand nombre de difficultés sont apparues quand il s'est agi de mettre en évidence la validité de la forme gaussienne de l'efficacité dans l'industrie et dans la recherche. Ces difficultés ont été de deux ordres : pratiques et théoriques. D'une part, des problèmes pratiques se sont posés avec les tentatives d'application directe de cette forme d'efficacité sur les outils de gestion des portefeuilles et de positions de marché. Dans l'industrie financière, presque tous les métiers de la finance de marché furent concernés par ces difficultés opérationnelles. D'autre part, les tests statistiques ont mis en évidence la violation quasi-universelle de la normalité des distributions des variations des cours pour des hautes fréquences, et la violation partielle de la normalité pour des variations de basse fréquence. On observe un trop grand nombre de grandes et de petites variations et un nombre trop faibles de variations moyennes sur les marchés. Les distributions réelles empiriques des variations boursières relevées en fréquence fixe (quotidienne, hebdomadaire, mensuelle etc.), c'est-à-dire en temps "physique", ou temps "calendaire", présentent un coefficient d'aplatissement (kurtosis) trop élevé, ont une "forme" non normale, ce qui caractérise les marchés par une propriété que l'on peut appeler le "phénomène leptokurtique". Intuitivement, cette caractéristique statistique correspond à cette impression souvent vérifiée par les professionnels des marchés, selon laquelle "le marché

²Sur l'interaction entre théorie financière et pratiques professionnelles, voir par exemple Longin et Walter [1994].

est très calme, sauf quand il bouge beaucoup”. Alors que la distribution normale (gaussienne) reflète plutôt un marché ni très calme, ni très agité, mais qui reste en permanence, “moyen”.

De manière plus précise, soit $\{r(t), t \geq 0\}$ le processus du taux de rentabilité continu de S , tel que

$$r(t) = \ln S(t) - \ln S(0)$$

et l’on notera

Notation 1 $\Delta r(t, \tau) = r(t) - r(t - \tau) = \ln S(t) - \ln S(t - \tau)$

la rentabilité continue périodique sur l’intervalle de durée τ .

Le modèle standard défini par l’équation (1) implique que les rentabilités $\Delta r(t, 1) = (\mu - \sigma^2/2) + \sigma(W_t - W_{t-1})$ soient iid-normales de variance σ^2 , et que les rentabilités $\Delta r(t, \tau) = (\mu - \sigma^2/2)\tau + \sigma(W_t - W_{t-\tau})$ soient iid-normales de variance $\sigma^2\tau$. En pratique, on examine le comportement d’un marché sur une période $[0, T]$, subdivisée en n sous-périodes de longueur τ , ($T = n\tau$), avec des longueurs τ éventuellement différentes (par exemple, 1 jour, 1 semaine, 1 mois, ou autre) qui déterminent donc la taille n de l’échantillon. On note, pour $t = k\tau$, $\Delta r(t, \tau) = \Delta r_k = r_k - r_{k-1}$. L’analyse du marché entre les dates 0 et T conduit à l’obtention d’un échantillon de taille n , $\{\Delta r_1, \dots, \Delta r_n\}$.

La forme d’une distribution est en général appréciée par la valeur des paramètres, ou coefficients, de Pearson et Fisher. Le coefficient de Pearson d’aplatissement (ou “kurtosis”) K d’une variable aléatoire X de fonction de répartition F et de moyenne $\mu = EX = \int_0^{+\infty} x dF(x)$ est défini comme

$$K_X = \frac{E(X - \mu)^4}{(E(X - \mu)^2)^2}$$

Pour une distribution normale, $K = 3$. Une distribution est dite “leptokurtique” si $K > 3$. Le coefficient de Fisher normalise celui de Pearson en posant $K' = K - 3$, et les résultats se comparent alors directement à 0.

Définition 1 *Ce qui est appelé le “phénomène leptokurtique” est l’obtention systématique, sur tous les échantillons $\{\Delta r_1, \dots, \Delta r_n\}$, d’une valeur de K supérieure à 3.*

Il est à noter que cela ne signifie pas nécessairement la présence de queues de distributions épaisses, dans la mesure où une kurtosis $K > 3$ peut être obtenue avec des distributions à queue fine, comme par exemple dans le cas d’un mélange de lois normales. La théorie des valeurs extrêmes montre qu’une distribution à queue épaisse correspond à une distribution des valeurs extrêmes avec un indice de queue différent de zéro. En d’autres termes, la présence d’une kurtosis excessive n’implique pas obligatoirement l’existence de grandes variations³. Cependant, sur les marchés financiers, le terme “leptokurtique” a été identifié à la présence de ces grandes variations, en complément (et au détriment) des autres aspects de la leptokurticité. C’est la raison pour laquelle on a choisi de désigner sous le nom de “phénomène leptokurtique” l’existence

³Je remercie François Longin d’avoir attiré mon attention sur ce point.

institutionnellement reconnue sur les marchés financiers, de variations boursières à l'amplitude excessive par rapport à l'hypothèse centenaire⁴ de normalité de ces variations.

Donnons quatre raisons (inspirées de Fama [1965]) qui présentent l'importance de la forme (et donc de la courbure) de la distribution des variations boursières :

Nature du risque réel La forme de la distribution définit et quantifie la nature du risque pris au moment de l'achat d'un actif financier, à travers les possibilités de plus ou moins grandes variations de valeur de cet actif.

Outils statistiques La forme de la distribution influence le choix et la définition des outils statistiques qui seront le plus adaptés à l'analyse du comportement des marchés.

Choix de modèles de risque La forme de la distribution détermine le choix des modélisations du risque pris par les acteurs des marchés, et donc la capacité d'utilisation des outils d'aide à la décision et des techniques de gestion qui reposent sur des modélisations données.

Processus des rentabilités La forme de la distribution fournit une information sur la nature du processus générateur des variations des prix.

L'importance du phénomène leptokurtique pour le développement comme pour la sécurité des opérations financières est donc claire.

L'objet de cet article est de présenter un état de la question et de mettre en évidence certains aspects de la non normalité peu explicités jusqu'à présent. Dans une première partie, on rappelle la permanence du phénomène leptokurtique, à travers un parcours historique des observations faites sur différents marchés depuis 1960, en faisant apparaître sa récente émergence en tant que "fait" d'expérience des marchés, sous le double effet d'une modification réglementaire (la démarche prudentielle du comité de Bâle) et d'un changement technologique (les systèmes de cotations automatisés des bourses). Les termes du débat sur la modélisation sont alors introduits, avec l'outil de base de l'analyse des queues de distribution, la théorie des valeurs extrêmes et la distribution de Pareto généralisée. La seconde partie traite de l'importance pratique de la non normalité en finance, en considérant ses implications dans trois questions professionnelles d'actualité : le risque de modèle, le calcul des indicateurs de perte maximale probable, et la gestion des portefeuilles. On introduit la notion de hiérarchisation des variations boursières pour faire apparaître le poids des grandes variations, on définit une fonctionnelle de value-at-risk en lois α -stables de Lévy, et on présente une typologie particulière de l'industrie de la gestion d'actifs, typologie construite spécifiquement pour faire apparaître l'impact de l'hypothèse gaussienne sur les processus d'investissement. Enfin, on donne un résultat nouveau sur la concentration de la performance des marchés. Dans la troisième partie, on aborde la question des explications du phénomène leptokurtique. On présente une grille de lecture simple des principaux courants d'explication, en proposant un point de vue qui utilise comme variable différenciante la nature de l'information : par la structure de l'économie (information exogène), par le comportement des agents (information endogène), ces deux voies conduisant finalement à l'étude de la microstructure des marchés et au réexamen de la question du temps boursier. En conclusion sont indiquées des voies de recherches pour des développements futurs.

⁴Car introduite en 1900 par Louis Bachelier.

2 Mise en évidence de la leptokurticité : naissance d'un fait expérimental

2.1 Historicité du problème leptokurtique

2.1.1 Ancienneté du phénomène leptokurtique

Le phénomène leptokurtique n'est pas nouveau, c'est sa reconnaissance institutionnelle qui l'est : il est aujourd'hui clairement acquis (au sens d'un "common knowledge") que les marchés boursiers ne possèdent pas un comportement statistique modélisable par une densité gaussienne lorsque les variations des cours sont relevées en fréquence fixe. La non-normalité des distributions empiriques est acceptée comme un "fait" de base des marchés. Il est intéressant de noter qu'il aura fallu près de quarante ans pour ce "fait" observable devienne un "fait" observé, la non-normalité entrant dans le champ de la recherche scientifique.

En effet, il est apparu assez rapidement (par exemple Osborne [1959], Alexander [1961]) que les queues de distribution des rentabilités empiriques contenaient trop de points pour pouvoir être ajustées par une loi de Laplace-Gauss. Mais ce phénomène n'était pas considéré comme "fait" significatif pour la bonne compréhension des variations du marché étudié et la modélisation des mouvements des marchés.

Benoît Mandelbrot, le premier, prit en 1962 l'option opposée, et attira l'attention sur l'importance financière du caractère leptokurtique des distributions observées, ainsi que sur la nécessité de ne pas chercher à ignorer ce phénomène. Il fallait selon lui, tout au contraire, s'intéresser particulièrement à ces points de queues. Mandelbrot considérait en effet que les queues de distributions, loin d'être un refuge de "points aberrants", contenaient des informations essentielles sur le fonctionnement du marché, pour la compréhension des risques potentiels des marchés. Selon Mandelbrot, l'information importante ne se perd pas dans la somme (les moyennes des distributions), mais se réfugie dans les queues. La position de Mandelbrot représentait donc une modification radicale du regard porté sur les variations boursières. Jusqu'alors, on s'était intéressé aux fluctuations des sommes partielles des variables aléatoires représentant les rentabilités boursières, et au comportement asymptotique des moyennes empiriques. Mandelbrot suggérait de s'intéresser aux fluctuations des extrêmes.

Malgré ces premiers travaux de Mandelbrot [1962, 1963, 1967], et de Fama [1965], le phénomène "événement rare" resta longtemps considéré comme négligeable ou non significatif : la vision du monde des années soixante n'incluait pas les queues de distribution dans l'analyse des fluctuations boursières. Dix ans après les premières mises en évidence, certains travaux suggéraient même, en cohérence avec cette conception intellectuelle, de tronquer les queues de distribution pour en retirer les grandes valeurs, assimilées à des "points aberrants" non significatifs pour la modélisation : le "fait" disparaissait purement et simplement du phénomène observé. Granger et Orr [1972] représentent une illustration de cette position, qui valide de manière emblématique la conception de Kuhn [1983] sur la nature subjective des "faits" d'expérience dans les paradigmes⁵.

⁵Pour un développement sur l'utilisation de la notion de paradigme au sens de Kuhn à la conception normale-gaussienne des variations boursières, voir Walter [1994, ch.8], et une discussion de l'applicabilité de la notion de

auteurs	année	test	période	fréquence
Mandelbrot	1962	cours du coton	1880-1958	quotidienne, mensuelle
Mandelbrot	1963	cours du blé	1883-1936	quot., hebdo., mensuelle
Mandelbrot	1963	4 actions de sociétés de chemins de fer	1883-1936	quot., hebdo., mensuelle
Mandelbrot	1963	taux de change USD/STG	1857-1936	quot., hebdo., mensuelle
Fama	1965	30 actions du DJIA	1957-1962	quotidienne
Mandelbrot	1967	écart de taux CT à 6%	1857-1936	mensuelle

TAB. 1 – Les travaux précurseurs de Mandelbrot et Fama sur la leptokurticité

auteurs	année	test	période	fréquence
Press	1967	10 actions du DJIA	1926-1960	mensuelle
Teichmoller	1971	30 actions du NYSE	1962-1967	1 à 10 jours
Officer	1972	39 actions du NYSE	1926-1968	mensuelle
Officer	1972	50 actions du NYSE	1926-1968	quotidienne
Fielitz et Smith	1972	200 actions du NYSE	1963-1968	quotidienne
Praetz	1972	17 indices de la bourse de Sydney	1958-1966	mensuelle
Barnea et Downes	1973	81 actions du NYSE	1968-1969	quotidienne
Hsu <i>et al.</i>	1974	Boeing, ATT, General Motors, RCA	1963-1970	mensuelle
Brenner	1974	indice SP 500	1928-1971	quotidienne
Blattberg et Gonedes	1974	indice DJIA	1957-1962	1 à 5 jours
Zajdenweber	1974	24 actions du RM de Paris	1970-1973	quotidienne
Hagerman	1978	805 actions du NYSE	1962-1976	quotidienne
Jorion	1988	indice CRSP pondéré	1973-1985	hebdo., mensuelle
Mantegna	1991	indice MIB	1975-1989	quotidienne
Mantegna	1991	15 indices sectoriels du MIB	1980-1989	quotidienne
Walter	1994	21 indices MSCI	1980-1993	mensuelle
Walter	1994	indice SP 500	1926-1991	mensuelle

TAB. 2 – Permanence de la leptokurticité sur le marché des actions

2.1.2 Permanence du phénomène leptokurtique

Historiquement, c'est sur les marchés d'actions que s'est d'abord concentrée l'étude de la leptokurticité. Progressivement, est apparue une deuxième voie de recherche, sur les marchés des changes ; enfin, les marchés de matières premières et de taux d'intérêt, puis les marchés à terme, ont aussi été examinés. Aujourd'hui, tous les compartiments des marchés ont, à peu près, fait l'objet d'investigations statistiques de leurs distributions empiriques. Les tableaux 2 et 3 présentent un échantillon de cette évolution phénoménologique. Cette chronologie relève quelques dates, ou époques, sans aucunement chercher à prétendre être, ni exhaustive, ni complète (ce qui supposerait une ampleur trop importante). Nous choisissons quelques exemples dans le très grand nombre d'études consacrées à l'analyse des variations boursières, afin d'illustrer la permanence de la question leptokurtique dans la finance de marché : cette question n'a, de fait, jamais été absente des investigations entreprises pour caractériser au mieux les distributions réelles observées.

Certains travaux mettent en évidence une diminution de la leptokurticité avec l'augmentation de l'ouverture du pas de temps, en temps calendaire. Le tableau 4 donne un échantillon de ces investigations.

paradigme en économie dans Coats [1969].

auteurs	année	test	période	fréquence
Westerfield	1977	USD/GBP-DEM-CAD-CHF-NLG	1970-1975	hebdomadaire
Rogalski et Vinso	1978	USD/GBP-DEM-CAD-CHF-NLG	1970-1975	hebdomadaire
Cornell et Dietrich	1978	USD/GBP-DEM-JPY-CHF	1975-1979	quotidienne
McFarland <i>et al.</i>	1982	USD/GBP-DEM-JPY-CAD	1973-1979	quot., hebdo., mens.
Friedmann et Vandersteel	1982	USD/9 devises	1973-1979	quotidienne
Hsieh	1988	USD/GBP-DEM-JPY-CAD-CHF	1974-1983	
Jorion	1988	USD/DEM	1973-1985	hebdo., mensuelle
So	1987	USD/GBP-DEM-JPY-CAD-CHF	1974-1982	quotidienne
Atlan <i>et al.</i>	1990	FRF/8 devises	1970-1991	hebdo., mens., trim.

TAB. 3 – Permanence de la leptokurticité sur le marché des changes

auteurs	année	test	période	fréquence
Blattberg et Gonedes	1974	30 actions du NYSE	1957-1962	1,3,5,8 jours
Hagerman	1978	805 actions du NYSE	1962-1976	1-35 jours
Hagerman	1978	286 actions de l'AMEX	1962-1976	1-35 jours
Upton et Shannon	1979	50 actions du NYSE	1956-1975	1,3,6,12 mois
Fielitz et Rozelle	1983	50 actions du NYSE	1962-1969	1-30 jours
Fielitz et Rozelle	1983	indice SP 500	1928-1973	1-30 jours
Wasserfallen et Zimmermann	1985	USD/CHF	1978-1980	1-10 minutes
Boothe et Glassman	1987	USD/GBP-DEM-JPY-CAD	1973-1984	1,5 jours, 1 mois, 1 trim.
Bouchaud et Potters	1997	contrat sur SP500	1991-1995	15 minutes-5 jours
Bouchaud et Potters	1997	contrat sur USD/DEM	1991-1995	15 minutes-5 jours
Bouchaud et Potters	1997	contrat sur Bund	1991-1995	15 minutes-5 jours

TAB. 4 – Décroissance du coefficient de kurtosis K

L'interprétation de la diminution de la leptokurticité (décroissance de K) reste cependant une question ouverte. Par exemple, parmi d'autres, Hall *et al.* [1989] y voient la trace d'une structure de dépendance complexe des variations des cours. Cette introduction d'une dépendance pour rendre compte de la leptokurticité, qui contrevient donc à la forme classique d'accroissements i.i.d. dans le modèle standard (1), permet d'envisager la question de l'agrégation temporelle des variations des cours à partir de l'existence d'horizons de temps différents selon les intervenants des marchés, et ouvre la voie à une interprétation de la décroissance de K au moyen de modèles d'interactions entre agents. De la même manière, Bouchaud et Potters [1997] analysent cette décroissance comme un effet de "respiration" de l'échelle des fluctuations boursières avec le temps, en introduisant des considérations sur l'activité du marché. Ainsi, de manière plus générale, on peut estimer que cette décroissance de K est liée à un certain type d'interaction entre acteurs des marchés, et que des modélisations de comportements d'agents pourraient représenter une voie de recherche pertinente pour interpréter efficacement la convergence à la normalité des variations boursières au fur et à mesure que la fréquence d'observation diminue. Nous abordons cet aspect de la leptokurticité dans la dernière partie de l'article.

2.2 Reconnaissance du problème leptokurtique

2.2.1 Evolution de l'environnement professionnel

On pouvait encore écrire en 1993 que “étudier les mouvements extrêmes est une approche originale en finance”⁶. Ce temps n’est plus.

En effet, parallèlement à un changement de vision dans la communauté scientifique, est apparue dans la communauté des professionnels une prise de conscience grandissante de la non-normalité et de ses conséquences. Un suivi de la presse professionnelle sur les deux dernières années l’illustre : entre 1998 et 2000, pour une revue comme *Risk*, on recense plus d’une quinzaine d’articles consacrés au problème de la non-normalité. La leptokurticité est même présentée comme faisant partie de la liste des dix problèmes principaux à aborder par l’industrie financière pour les prochaines années⁷, et la gestion des “événements rares” devient la question clé à résoudre dans la gestion du risque⁸.

Cet intérêt professionnel nouveau vient de ce que, dans le champ de l’industrie financière se déploie, depuis une dizaine d’année, une approche institutionnelle des risques de marché, approche dite “prudentielle”, qui se traduit par l’existence de groupes de réflexion internationaux dont l’objectif est de promouvoir des règles de quantification des risques universellement applicables par les établissements financiers dans tous les domaines de leur activité⁹. Ce souci prudentiel est étroitement corrélé avec l’augmentation extrêmement rapide des encours d’instruments financiers dérivés, dont l’instabilité et la difficulté de gestion constituent des sources grandissantes de préoccupation pour les autorités de tutelle des marchés¹⁰. Il est apparu de plus en plus clairement aux instances représentatives que constituent ces groupes de réflexion, que le phénomène leptokurtique venait affaiblir la fiabilité des outils et instruments tous fondés sur la distribution normale-gaussienne.

2.2.2 Apport des données de haute fréquence

Un autre changement, aux conséquences expérimentales importantes, est apparu dans l’environnement informatique des marchés boursiers : la puissance de calcul des machines et de la

⁶Longin [1993], p. 25.

⁷“The 10 challenges left to tackle”, *Risk*, avril 2000, pp 34-39.

⁸“Tailor-made for tails”, *Risk*, Février 2000, pp 95-98.

⁹Par exemple, le “Groupe des 30” (G30), le “Comité de Bâle”, ou les Commissions nationales des opérations de bourse, fédérées internationalement (International Organization of Securities Commissions, IOSCO), qui ont conduit depuis quelques années une réflexion sur la mise en place de cadres d’analyse appropriés pour l’évaluation et l’intégration comptable de ces risques : voir les références Basle Committee on Banking Supervision [1994, 1996, 1998] et IOSCO [1994].

¹⁰Un échantillon représentatif de soixante établissements financiers actifs sur les marchés dérivés, appartenant aux groupe des dix pays les plus riches (G10), fait apparaître en 1999 un encours nominal sur ces instruments de 130 000 milliards de dollars. L’exposition au risque des établissements concernés (rapport entre les encours nominaux et les actifs sous-jacents) peut atteindre le coefficient de 34, ce qui veut dire que l’établissement est exposé pour 34 fois la valeur des actifs de base sur lesquels se négocient les produits dérivés (source : Joint Report by the Basle Committee on Banking Supervision and the Technical Committee of the IOSCO, décembre 1999, p.4).

taille des bases de données disponibles, changement qui se traduit par une abondance toujours plus grande des données de cours disponibles. Avec des systèmes automatisés de cotation tels le système SuperCAC de la bourse de Paris ou SuperDOT de la bourse de New York, apparaît la possibilité de disposer de la totalité des cours cotés entre deux dates, ainsi que de la composition des carnets d'ordres. Ce nouveau gisement de données dilate démesurément le champ des possibilités d'analyse des distributions empiriques des rentabilités boursières. Ainsi, l'ordre de grandeur de la taille des échantillons a augmenté de plusieurs puissances de dix. Les tests statistiques réalisés sur les marchés dans les années soixante utilisaient environ 1000 données. Ceux du début des années quatre-vingt dix traitent environ 100 000 données. Les plus récentes investigations examinent environ 1 million de données. Ce qui représente une progression de 10^3 à 10^5 puis 10^7 .

Des conférences annuelles sur le traitement des données boursières de haute fréquence sont apparues à la suite de ces nouvelles possibilités, suivant l'initiative de la société suisse Olsen & Associates qui publie régulièrement les résultats de ses études (Olsen [2000]), et les travaux sur données de haute fréquence se sont ainsi multipliés, permettant un approfondissement des investigations du phénomène leptokurtique au moyen de la modélisation des durées entre échanges boursiers et de certaines caractéristiques de ces échanges (sens, volume, volatilité); ce qui permet de repenser la question du temps des marchés en considérant différents types de temps, selon que l'on suit le temps physique de l'horloge ou le temps opérationnel des échanges.

Ces études ouvrent de plus la voie à une analyse fine des marchés financiers à l'échelle la plus élémentaire, celle de la formation d'*un* prix particulier à partir de la rencontre entre *une* offre et *une* demande spécifiques, détectées grâce à l'analyse précise du carnet d'ordres, analyse permettant de tester les théories de la microstructure des marchés apparues à la suite des travaux de Kyle [1985]. Cette nouvelle sorte d'investigation des marchés permet de chercher une explication au phénomène leptokurtique à partir des comportements d'agents. Nous abordons cet aspect en dernière partie de l'article.

En conclusion de cette revue rapide, on peut dire que, de manière générale, le phénomène leptokurtique est nettement présent sur de nombreux marchés, et sur de nombreuses périodes d'observation.

2.3 Modélisation du phénomène leptokurtique

2.3.1 Termes du débat sur la modélisation

Les recherches relatives à la prise en considération du problème leptokurtique et à la modélisation de la leptokurticité ont eu (et ont toujours) un objectif double :

Descriptif Il s'agit dans cette perspective de parvenir à modéliser le mieux possible les variations des cours réels au moyen d'outils probabilistes ou statistiques, en postulant des classes de processus qui reproduisent au mieux la non-normalité, ceci de manière phénoménologique, et sans chercher à expliquer l'origine de ces variations, c'est-à-dire sans référence à des comportements d'agents. C'est une approche minimaliste, mais qui a une utilité directe et pratique dans l'évaluation des actifs financiers et la gestion des positions de marché.

Elle apparaît dès les années soixante, et fait l'objet aujourd'hui de recherches intenses, renouvelées par les données de haute fréquence.

Explicatif Il s'agit dans cette perspective de développer des hypothèses relatives aux comportements des agents et à la microstructure des marchés, ceci pour parvenir, de manière inductive, à des processus de prix pertinents quant à la réalité observée. C'est une démarche plus ambitieuse, donc plus complexe (et d'ailleurs pour cette raison plus récente), mais plus délicate à tester. Son utilité pour les professionnels ne leur apparaîtra peut-être pas immédiate, mais c'est cependant un moyen efficace pour pouvoir choisir entre plusieurs classes de processus candidats, relevant de l'approche descriptive seule.

Ces deux démarches sont cependant dans un mouvement de rapprochement l'une de l'autre, avec les investigations sur données de haute fréquence, dans la mesure où ces données font intervenir d'autres variables que les seuls cours cotés. Elles restent pour le moment encore séparées, mais la jonction entre les travaux sur la microstructure des marchés et les choix de processus de prix peut apparaître comme un objectif atteignable à un horizon raisonnable pour la recherche en finance.

On présente ci-dessous l'outil de base de description des queues de distribution, laissant pour la dernière partie de l'article la question de l'interprétation de la leptokurticité au moyen de modèles explicatifs.

2.3.2 Caractère parétien des queues de distribution et indice de queue

La nécessité de prendre en compte une structure différente d'aléa dans les distributions réelles, et plus particulièrement de s'intéresser aux propriétés des fluctuations des *extrêmes* est à l'origine d'un renouveau d'une partie de la théorie des probabilités jusqu'alors moins explorée par la finance : le comportement des maxima dans les suites de sommes partielles de variables i.i.d., formalisé par la théorie des valeurs extrêmes, qui conduit à la distribution de Pareto généralisée définie en (4) ci-dessous.

Introduite¹¹ dans le champ de l'analyse des marchés financiers par Longin [1993, 1995, 1996], la théorie des valeurs extrêmes connaît depuis un succès grandissant et apparaît comme un outil pertinent pour le traitement adéquat des queues de distribution. A nouveau, bons indicateurs de cet intérêt, les articles professionnels : par exemple, en 1998, il est affirmé que "la théorie des valeurs extrêmes fournit un outillage puissant pour la gestion des risques"¹². On résume ci-dessous les principaux résultats de cette approche pour la gestion du problème leptokurtique, dans le but de faire apparaître le caractère "universel" de la distribution de Pareto dans la mise en évidence de ce phénomène.

Soit $\{X_1, X_2, \dots, X_n\}$ une suite de variables aléatoires i.i.d. de fonction de répartition $F(x) =$

¹¹Longin [1993, p. 21] : "C'est la première fois, à notre connaissance, que la théorie des valeurs extrêmes est appliquée en finance pour caractériser la distribution du prix des actifs. On peut s'étonner du peu d'études sur les variations extrêmes des cours entreprises par le milieu académique sachant que les investisseurs et les gérants de portefeuille sont très sensibles aux fortes variations de leur richesse investie. C'est d'autant plus surprenant que la théorie a été élaborée dans la première partie du XXème siècle et qu'elle a été appliquée depuis à de nombreux domaines".

¹²"History repeating", *Risk*, Janvier 1998, p. 99.

$\Pr(X \leq x)$. On peut considérer par exemple que $X = \Delta S$ ou $X = \Delta r$ (voir définitions dans l'introduction). La gestion des risques de marché s'intéresse à différentes quantités relatives aux queues de distribution de X .

Une première mesure intéressante pour la gestion des risques est l'espérance de X si une amplitude a donnée est dépassée, ou moyenne conditionnelle en a

$$\mu(a) = E(X/X > a) = \frac{1}{1 - F(a)} \int_a^{+\infty} x dF(x) \quad (2)$$

et $\mu(a)$ est une manière commode de percevoir des répartitions *inégalitaires*.

Par exemple, dans le cas où X est un revenu, $\mu(a)$ représente le revenu moyen des individus ayant un revenu supérieur à a . Dans le cas où X est un montant de sinistre en assurance-dommage, $\mu(a)$ représente le montant moyen des sinistres de montant supérieur à a . Dans le cas où $X = \Delta S$ et $a = \text{VaR}_c$, où VaR_c est la "value-at-risk" au seuil de confiance c (voir plus bas la définition (20) de la VaR), $\mu(a)$ représente la perte moyenne si la VaR_c est dépassée, et correspond à la mesure appelée "tail conditional expectation" introduite par Artzner *et al.* [1997] et Bouchaud et Potters [1997].

En écrivant que $E(X/X > a) = a + E(X - a/X > a)$, on fait apparaître un autre indicateur très utile dans la gestion des risques, qui est la notion d'espérance de dépassement de seuil. Plus précisément, on a :

Définition 2 Si X est une variable aléatoire de fonction de répartition F , alors

$$F_a(x) = \Pr(X - a \leq x / X > a) \quad x \in \mathbb{R} \quad (3)$$

est la fonction de répartition de l'excédent de X au-delà du seuil a , et

$$e(a) = E(X - a/X > a)$$

est appelée *espérance de dépassement de X au-delà du seuil a* .

En assurance, $e(a)$ est appelé "excédent de perte" ("excess of loss") au-delà d'un seuil a . En finance, si par exemple $a = \text{VaR}_c$, $e(a)$ est appelé "expected shortfall" au seuil c

$$\text{ES}_c = E(X - \text{VaR}_c/X > \text{VaR}_c)$$

de fonction de répartition $F_{\text{VaR}_c}(x)$.

Ceci fait apparaître le rôle et donc l'importance de la forme de la distribution $F_a(x)$ définie dans (3) pour la gestion des risques. Le résultat suivant est le plus important pour la caractérisation du phénomène leptokurtique.

On montre¹³ que :

¹³Voir Embrechts *et al.* [1997], chapitre 3.

– le *modèle naturel* de la fonction de répartition $F_a(x)$ est la distribution de Pareto généralisée

$$G_{\xi,\sigma}(x) = \begin{cases} 1 - (1 + \xi x/\sigma)^{-1/\xi} & \xi \neq 0 \\ 1 - \exp(-x/\sigma) & \xi = 0 \end{cases} \quad (4)$$

avec

$$\begin{cases} x \geq 0 & \text{si } \xi \geq 0 \\ 0 \leq x \leq -\sigma/\xi & \text{si } \xi < 0 \end{cases}$$

où σ est un paramètre d'échelle, et $\xi > 0$ un paramètre de "forme" de la distribution. Plus précisément, ξ est un indice de poids des queues de distribution. Quand $\xi > 0$, la distribution est à queue épaisse. Plus ξ augmente, plus la queue de distribution s'épaissit. Les cas $\xi \leq 0$ correspondent à des distributions à queue fine. Le cas $\xi = 0$ correspond à la présence de queues de distribution exponentielles, telles que les gaussiennes.

– l'espérance de dépassement de seuil $e(a)$ est une fonction affine de la forme

$$e(a) = \frac{\xi}{1-\xi} a + \frac{\sigma}{1-\xi} \quad \xi < 1, \sigma + a\xi > 0 \quad (5)$$

où la condition $\xi < 1$ correspond à $EX < \infty$.

Dans le cas $\xi \leq 0$, les grandes variations sont très improbables : par exemple, avec une gaussienne, la probabilité qu'une baisse de 10 écarts-types survienne est de l'ordre de 10^{-23} . Dans la pratique, comme il faut prendre en considération la présence de queues de distribution épaisses, le cas $\xi > 0$ (i.e. $x \geq 0$) est le plus utilisé dans la gestion des risques. Dans ce cas, la fonction de répartition $F_a(x)$ pour des seuils a suffisamment élevés est de la forme

$$G_{\xi,\sigma_a}(x) = \begin{cases} 1 - (1 + \xi x/\sigma_a)^{-1/\xi} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

où $\sigma_a > 0$ est un paramètre d'échelle qui dépend du seuil a .

La fonction F_a est de mieux en mieux approchée par une distribution de Pareto au fur et à mesure que a augmente. On peut alors ajuster la fonction de répartition empirique des observations x_i au-delà de a par une fonction de répartition de la forme

$$G_{a,\xi,\sigma}(x) = \begin{cases} 1 - (1 + \xi(x-a)/\sigma)^{-1/\xi} & x \geq a \\ 0 & x < a \end{cases} \quad (7)$$

Remarquons que, en posant $\sigma = a\xi$, quand $\xi > 0$, et $\xi = 1/\alpha$, (7) devient

$$G_\alpha(x) = \begin{cases} 1 - (a/x)^\alpha & x \geq a \\ 0 & x < a \end{cases} \quad (8)$$

et l'on retrouve la forme classique de la distribution de Pareto avec α qui est l'exposant quantifiant le poids des queues de distribution. Cette distribution de Pareto définit une catégorie importante de comportements de queues de distribution : les comportements en *lois de puissance*. Une densité

de probabilité $P(x)$ d'une variable aléatoire X admet un comportement asymptotique en loi de puissance d'exposant $1 + \alpha$ si

$$P(x) \cong C \frac{1}{|x|^{1+\alpha}} \quad \text{quand} \quad x \rightarrow \pm\infty \quad (9)$$

où C dépend de α .

De (8), il vient, pour $x \geq a$

$$\Pr(X \geq x) = 1 - G_\alpha(x) = \left(\frac{a}{x}\right)^\alpha \quad (10)$$

soit une relation très commode pour la mise en évidence visuelle de lois de puissance

$$\ln(\Pr(X \geq x)) = -\alpha \ln x + C_\alpha \quad (11)$$

Un graphique en coordonnées bilogarithmiques produira dans ce cas une droite de pente $(-\alpha)$, et réciproquement, la mise en évidence d'un tel alignement pourra être interprété comme un signe de la présence d'une loi de puissance sur la queue de distribution. Une telle représentation bilogarithmique permettra d'insister sur l'importance des queues de distribution, en isolant visuellement les quelques grandes valeurs qui produisent l'impact maximum sur le phénomène observé.

2.3.3 La quête du processus des rentabilités

La théorie des valeurs extrêmes répond adéquatement aux trois premiers aspects développés par Fama [1965], cités en introduction. En ce qui concerne le quatrième, le processus des rentabilités, cette théorie ne fournit pas directement d'équation de comportement des cours, c'est-à-dire de variante globale de modélisation, mais donne une indication sur le choix de processus candidats au remplacement du mouvement brownien, grâce à la description précise des queues de distribution. En effet, dans la mesure où la distribution de Pareto généralisée correspond à la présence de valeurs extrêmes, il semble naturel de rechercher des processus dont la loi marginale présente un comportement asymptotique parétien.

C'est par exemple le cas des mouvements α -stables de Lévy, introduits pour la première fois en finance par Mandelbrot [1962]. En effet, les lois α -stables $S_{\alpha,\beta}(\gamma, \delta)$, que l'on définit plus loin (voir définition 5), possèdent la propriété asymptotique suivante.

Propriété 1 Si $X \stackrel{d}{=} S_{\alpha,\beta}(\gamma, \delta)$ avec $\alpha \in]0, 2]$, alors

$$\lim_{x \rightarrow \infty} x^\alpha \Pr(X > x) = C_\alpha \frac{1 + \beta}{2} \gamma^\alpha \quad (12)$$

$$\lim_{x \rightarrow \infty} x^\alpha \Pr(X < -x) = C_\alpha \frac{1 - \beta}{2} \gamma^\alpha \quad (13)$$

où

$$C_\alpha = \left(\int_0^\infty x^{-\alpha} \sin x dx \right)^{-1} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha) \cos(\pi\alpha/2)} & \text{si } \alpha \neq 1 \\ \frac{2}{\pi} & \text{si } \alpha = 1 \end{cases}$$

Cette propriété revient à dire que la queue de distribution d'une loi stable décroît selon une *loi de Pareto*. Les mouvements α -stables apparaissent donc comme un candidat naturel de processus dans un contexte i.i.d. pour des variables aléatoires qui présentent des queues de distribution parétiennes.

De nombreuses études statistiques ont été réalisées sur les marchés depuis les propositions de Mandelbrot [1962, 1963] et de Fama [1965] pour tester l'applicabilité des mouvements α -stables à la modélisation des fluctuations boursières. Si, de manière générale, l'ajustement marginal à une distribution α -stable est vérifié dans l'ensemble des travaux, il n'en est pas de même pour la propriété de stabilité (qui suppose la constance de α lorsque la fréquence d'observation se modifie). Les recherches ont presque toutes fait apparaître une invalidation partielle de cette propriété de stabilité¹⁴, se traduisant par une croissance de l'exposant caractéristique α , croissance qui correspond à une décroissance de la kurtosis K , lorsque la fréquence diminue (voir Walter [1994] pour plus de précision sur la relation entre α et K).

Lévy-Véhel et Walter [2002] proposent de résumer l'état de la controverse sur la stabilité de la manière suivante :

- Même si la stabilité n'est pas validée sur tout le spectre des fréquences d'observation, on peut trouver sur les marchés des bandes d'invariance partielles (par exemple de 1 à 10 jours), qui permettent d'utiliser de façon fiable de telles distributions. On s'accorde alors à considérer que le rapport entre coût de modélisation (lié aux nombres de paramètres à estimer) et réalisme de la modélisation (lié à la capacité à rendre compte des phénomènes réels) est, dans ces cas, à l'avantage des mouvements α -stables, par opposition aux processus plus généraux, comme les processus de Lévy non stables¹⁵.
- Les méthodes d'estimation de l'exposant caractéristique α semblent présenter des biais importants lorsque les rentabilités ne sont pas i.i.d. Embrechts *et al.* [1997] estiment ainsi que "les méthodes d'estimation du paramètre α ne sont pas assez précises pour inférer une conclusion *nette* sur la nature réelle des distributions à partir d'estimations faites sur plusieurs échelles de temps". Il apparaît difficile de conclure de manière définitive sur la question de la fluctuation de α pour cette raison.
- Le critère de décision de l'utilisation ou non de tels processus est donc lié à l'objectif poursuivi. Dans certains cas, cela s'avère pertinent, dans d'autres, non.

La recherche se poursuit sur les processus de Lévy non stables, qui sont des processus de sauts purs, et qui ont été examinés récemment (voir par exemple Eberlein [1999], Carr *et al.* [2000], Raible [2000]). Plus puissants en termes de modélisation, ils sont cependant plus complexes à manipuler, et l'estimation des paramètres est très délicate.

¹⁴Voir Walter [2002] pour une revue de synthèse sur 40 ans de ce problème et des controverses qu'il a soulevé.

¹⁵Un processus de Lévy est un processus à accroissements i.i.d. Les mouvements α -stables sont des processus de Lévy particuliers, qui sont "stables par addition", c'est-à-dire "fractals" au sens de Mandelbrot. Les processus de Lévy non stables sont à accroissements i.i.d., mais non fractals. Voir Mandelbrot [1997].

3 Importance pratique du phénomène leptokurtique : la nécessaire gestion des queues de distribution

Dans cette partie, on s'intéresse aux conséquences pratiques de la leptokurticité (au sens des queues de distribution) pour les outils d'aide à la décision utilisés par les professionnels. Sont abordés successivement : la question du risque de modèle, la mesure appelée "value-at-risk", et la gestion des portefeuilles.

3.1 Queues de distribution, modèles de risque et risque de modèle

3.1.1 Poids des grandes variations et perte de pertinence des moyennes

La période d'analyse $[0, T]$ d'un marché étant subdivisée en N intervalles de longueur τ , on considère les N variables aléatoires, $\{\Delta r_1, \dots, \Delta r_N\}$, dont on fait l'hypothèse qu'elles sont i.i.d. entre les dates 0 et T , dont la densité de probabilité est $P(\Delta r)$ et dont la fonction de répartition est $F_{\Delta r}(x) = \Pr(\Delta r \leq x) = \int_{-\infty}^x P(t)dt$. On s'intéresse aux queues de distribution de Δr . Pour cela, il est utile de séparer les variations positives des variations négatives, et d'utiliser une statistique d'ordre sur chacune de ces variations en rangeant l'échantillon par ordre décroissant. Dans la suite du texte, on considère les n variations de même signe de Δr .

Notation 2 $\Delta r_{(k)}$ désigne la k -ième plus grande valeur de $\{\Delta r_1, \dots, \Delta r_n\}$.

D'où $\Delta r_{(1)} = \max(\Delta r_1, \dots, \Delta r_n)$ et $\Delta r_{(n)} = \min(\Delta r_1, \dots, \Delta r_n)$ et $\Delta r_{(1)} \geq \dots \geq \Delta r_{(n)}$ est l'échantillon ordonné de $\{\Delta r_1, \dots, \Delta r_n\}$.

Le rapport de hiérarchisation Une particularité importante des distributions à queue étirée est la très forte *hiérarchisation* des grandes variations. Pour préciser cette relation entre les Δr , on introduit la notion de rapport de hiérarchisation entre deux variations rangées successives.

Définition 3 Pour un échantillon de taille n , $\{\Delta r_1, \dots, \Delta r_n\}$, le rapport de hiérarchisation de rang k est défini par

$$\frac{\Delta r_{(k)}}{\Delta r_{(k+1)}} \quad k = 1, \dots, n-1$$

où $\Delta r_{(k)}$ est la k -ième plus grande valeur de l'échantillon ordonné $\Delta r_{(1)} \geq \dots \geq \Delta r_{(n)}$.

On cherche à présent à caractériser l'ordre de grandeur de $\Delta r_{(k)}$ lorsque les Δr suivent une loi de puissance.

On rappelle que la fonction de répartition empirique $F_n(x)$ d'une variable aléatoire X est définie par

$$F_n(x) = \frac{1}{n} \text{card}\{i : 1 \leq i \leq n, X_i \leq x\} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}, \quad x \in \mathbb{R}$$

où $\mathbf{1}_A$ désigne la fonction indicatrice de l'événement A .

On a donc

$$F_n(\Delta r_{(k)}) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\Delta r_i \leq \Delta r_{(k)}\}} = 1 - \frac{k}{n}$$

c'est-à-dire

$$\text{Fr}(\Delta r_i \geq \Delta r_{(k)}) = 1 - F_n(\Delta r_{(k)}) = \frac{k}{n} \quad (14)$$

où Fr est la fréquence cumulée des variations Δr .

Ainsi, si les $\Delta r_{(k)}$ sont répartis en lois de puissance, en vertu de (10) et (14) on a

$$\Delta r_{(k)} = a \left(\frac{n}{k} \right)^{1/\alpha} \quad (15)$$

qui est la relation recherchée, et qui donne l'ordre de grandeur de la k -ième valeur rencontrée.

On en déduit

$$\frac{\Delta r_{(k)}}{\Delta r_{(k+1)}} = \left(\frac{k+1}{k} \right)^{\frac{1}{\alpha}} \quad (16)$$

qui fournit une relation très importante pour la gestion des risques.

Considérons par exemple le rapport de hiérarchisation entre la plus grande variation et la seconde plus grande. D'après (16), il vient

$$\frac{\Delta r_{(1)}}{\Delta r_{(2)}} = 2^{\frac{1}{\alpha}}$$

Le rapport entre les deux premières plus grandes variations est de l'ordre de grandeur de $2^{1/\alpha}$. Si $\alpha = 1,5$, on obtient $2^{2/3} \cong 1,6$: la plus grande variation observée est 1,6 fois plus grande que la seconde plus grande. Si $\alpha = 0,5$, on obtient $2^2 = 4$: la plus grande variation est 4 fois plus grande que la seconde plus grande. Plus α diminue, et plus ce rapport, donc la hiérarchisation des variations, augmente. À l'opposé, ce rapport tend vers 1 quand α tend vers l'infini (queues exponentielles). L'existence de queues de distribution en loi de puissance se traduit par la *concentration du risque* sur un très petit nombre d'événements, ou événements rares : les grandes baisses.

En passant en coordonnées bilogarithmiques, (14) et (15) donnent

$$\ln \left(\text{Fr}(\Delta r_i \geq \Delta r_{(k)}) \right) \cong -\alpha \ln \Delta r_{(k)} + C_\alpha \quad (17)$$

qui est historiquement la relation qu'a obtenue Mandelbrot lors de la première mise en évidence de lois de puissance sur les marchés.

FIG. 1 – Variations positives de l'indice CAC 40

FIG. 2 – Variations négatives de l'indice CAC 40

Enfin, en utilisant la relation (14) dans l'équation (17), on obtient

$$\ln k = -\alpha \ln \Delta r_{(k)} + C'_\alpha \quad (18)$$

qui définit une relation entre le rang et l'amplitude d'un événement, qui est une relation rang-taille introduite par Zipf et Estoup. Un graphique de type "rang/amplitude", où l'amplitude $\Delta r_{(k)}$ est figurée en abscisse et le rang k en ordonnée permet donc d'obtenir l'histogramme cumulé de la loi de $P(\Delta r)$, ou "fréquence cumulée" $\text{Fr}(\Delta r_{(k)})$. Les deux graphiques 1 et 2 présentés en exemple illustrent cette propriété. Les données sont les cours de l'indice CAC 40 entre le 1^{er} janvier 1988 et le 30 juin 2001 avec $\tau = 1$ jour, soit $N = 3460$ cours. Sur la période, on observe 1856 variations positives et 1604 variations négatives.

La perte de pertinence des moyennes Cette hiérarchisation des grandes variations a une conséquence importante dans la gestion des risques : l'affaiblissement de la pertinence des valeurs typiques (moyenne, écart-type) des populations statistiques concernées par la gestion de ces risques, et la contribution essentielle des grandes variations au résultat final. Considérons les baisses de marchés ($\Delta r < 0$), et associons la présence d'un risque à l'occurrence d'une baisse. La gestion des risques s'intéresse, comme on l'a vu, à la quantité $\mu(a)$ pour déterminer un niveau de fonds propres nécessaire à la couverture d'une activité de marché sur la variable Δr . Un calcul usuel de couverture de positions de marchés contre des baisses utilisera les valeurs typiques des échantillons $\{\Delta r_1, \dots, \Delta r_n\}$ et, plus généralement, les propriétés limites des fluctuations des *sommes* partielles $r_n = \Delta r_1 + \dots + \Delta r_n$. Mais, si le calcul d'une moyenne empirique μ_n , telle que $r_n = n\mu_n$, est toujours possible, quelle est sa réelle pertinence ? On va montrer que, dans le cas où les variations Δr suivent une loi de puissance, le comportement de r_n est principalement déterminé par le comportement des premiers $\Delta r_{(k)}$, ce qui invalide l'usage direct de μ_n .

Observons tout d'abord que, si $\alpha < 1$, la relation (15) montre que $\Delta r_{(1)}$ est de l'ordre de grandeur de $n^{1/\alpha} \gg n$ et donc $\Delta r_{(1)} \gg r_n$: la plus grande variation serait dans ce cas supérieure à la somme tout entière ! C'est une situation dans laquelle la moyenne μ n'est plus du tout une valeur "typique", et donc n'est plus d'aucune utilité pour un calcul de couverture de risque. Pour l'écart-type de r_n qui est $\sigma_n = \sigma_1 \sqrt{n}$, un raisonnement analogue peut être tenu avec $\alpha < 2$. En effet, dans ce cas, $n^{1/\alpha} > \sqrt{n}$ et $\Delta r_{(1)}$ augmente plus vite que \sqrt{n} : la variance ne sera pas non plus utilisable.

Plus généralement, faisons l'hypothèse que les variations Δr suivent une loi de puissance au-delà d'un certain seuil a , et effectuons le calcul de perte probable $\mu(a)$ définie en (2). Dans le cas d'une distribution de Pareto de forme classique (8), on a

$$\Pr(\Delta r > x / \Delta r > a) = 1 - G_\alpha(x) = \left(\frac{a}{x}\right)^\alpha$$

et donc

$$\mu(a) = E(\Delta r / \Delta r > a) = \int_a^{+\infty} x \alpha \frac{a^\alpha}{x^{\alpha+1}} dx = \frac{\alpha}{\alpha - 1} a \quad (19)$$

soit

$$e(a) = \mu(a) - a = \frac{1}{\alpha - 1} a$$

On voit que la perte probable $\mu(a)$ est égale au produit de a par un terme $\alpha/(\alpha - 1)$. Par exemple, si $\alpha = 1,5$, alors $\mu(a) = 3a$ et $e(a) = 2a$. Ce qui signifie que pour chaque seuil a , la moyenne des pertes supérieures à a est égale à $3a$. En d'autres termes, si a est dépassé, alors la perte moyenne conditionnelle triple et le montant de réserves à constituer ($e(a)$) double. Et plus α diminue, plus $\mu(a)$ et $e(a)$ augmentent¹⁶. Dans cette situation, la valeur moyenne des baisses "normales" ne dit rien de l'amplitude des grandes baisses qui constituent la source principale du risque.

Dans la mesure où $\alpha/(\alpha - 1) \rightarrow \infty$ quand $\alpha \rightarrow 1$, l'espérance de perte supérieure à a peut devenir *infinie* : la question du niveau de couverture $e(a)$ (qui représente, rappelons-le, la quantité "expected shortfall") devient alors extrêmement délicate. Si la couverture est calculée à partir d'une moyenne empirique $\mu_n(a)$, alors, selon la valeur choisie a , elle ne sera pas assez efficace en présence d'événements rares, et trop chère en leur absence.

Plus généralement, la pertinence des moyennes empiriques dépend de l'existence des moments des variables aléatoires. Mais avec des distributions en loi de puissance, les moments n'existent pas toujours. Dans le cas de la distribution de Pareto introduite en (4), $E(X^k)$ est infini pour $k \geq 1/\xi$. Par exemple, quand $\xi = 1/2$, le deuxième moment (la variance) est infini ; quand $\xi = 1/4$, le moment d'ordre 4 est infini. La gestion des risques par des valeurs typiques de type "moyenne" perd alors de sa pertinence, et de son efficacité. Ce phénomène est relevé par les professionnels qui utilisent des moyennes pour gérer leurs positions : comme le titrait un quotidien au moment de crises financières récentes, tout semble se passer "normalement", et puis... "soudain l'été dernier"¹⁷. Dans l'article cité précédemment, il était expliqué que, fondés sur la notion de perte moyenne et de couverture moyenne, "les instruments de limitation et de mesure des risques n'ont pas fonctionné", et on parlait pour cela "d'infraction aux lois de probabilité". Ainsi (suite de l'article), les techniques de gestion fondées sur la moyenne "offrent une *protection illusoire*" (souligné par nous). On peut dire que, sur un marché dont les queues de distribution sont épaisses, la pertinence des notions de "gain moyen", de "perte moyenne", ou de "journée moyenne" de bourse est relative aux observateurs : l'objectivation statistique des moyennes empiriques $\mu_n = r_n/n$, quoique toujours possible, ne possède qu'un sens subjectif (ou limité), logé dans l'esprit de ceux qui font les calculs, mais non un sens "objectif", qui dépend de l'existence des moments. En d'autres termes, les risques réels se jouent des calculs de moyennes.

Une autre conséquence de cette absence éventuelle de moments est la fluctuation des estimations des moments empiriques, à mesure de leur inexistence selon le poids des grandes valeurs. Ce mauvais conditionnement est bien reconnu dans les différents travaux existants. Par exemple, observant les variations permanentes des estimations du bêta du CAPM ou des facteurs de l'APT, Ghysels [1998] attribue ces biais et cette variabilité au fait que les *hypothèses probabilistes* ne sont pas validées.¹⁸ Dans ce cas, la perte de pertinence des moyennes se traduit par des problèmes

¹⁶Observons aussi que, à l'opposé, si $\alpha \rightarrow \infty$, alors $\mu(a)$ tend vers a , ce qui correspond aux queues de distribution exponentielles.

¹⁷"Soudain l'été dernier", *Le Figaro*, 14 mai 1999

¹⁸On peut noter que ces difficultés d'estimation des paramètres des modèles, qui rend leur usage délicat, ne sont d'ailleurs pas propres à la finance : Artus et Muet [1990, p.7] évoquaient, pour l'économie, les problèmes posés par

d'estimation.

En résumé, le phénomène leptokurtique fournit une raison à la défaillance des approches conceptuelle et prudentielle des marchés fondées sur l'utilisation de la notion de "moyenne". Un déplacement conceptuel important s'est ainsi produit au cours des dernières années dans l'esprit des acteurs des marchés financiers, les conduisant à passer des moyennes aux extrêmes.

3.1.2 Risque de modèle dû aux queues de distribution parétiennes

Il reste cependant possible de tenter de compenser la difficulté qu'ont les queues de distribution exponentielles (des distributions gaussiennes) à capter la structure variationnelle des grands risques par le moyen suivant : forcer artificiellement la volatilité implicite des marchés, c'est-à-dire augmenter la variance. En effet, dans les distributions à queue exponentielle, pour lesquelles $\xi = 0$, il vient de (5) que $e(a) = \sigma$ où σ est l'écart-type de la distribution. Cela veut dire que la moyenne de l'excédent de pertes qui adviennent lorsque le seuil a est dépassé est égale à l'écart-type. Ainsi, l'augmentation de σ permet d'augmenter $e(a)$ dans un cas gaussien.

Mais cette augmentation est artificielle si la "vraie" distribution de Δr est de forme parétienne. En considérant que $e(a) = e(a, \xi)$ selon (5), où la valeur de ξ détermine la nature de la distribution, au sens de sa *forme* (et, en ce sens, le paramètre ξ est un paramètre *morphogénétique* du risque réel), on voit que *l'opération de forçage de volatilité revient à chercher à compenser ξ par σ* , c'est-à-dire compenser l'absence d'un réglage de *forme* par une augmentation de la *taille* possible des fluctuations. Précisons cela au moyen de l'équation (5). Supposons que la "vraie" distribution soit parétienne d'indice de queue $\xi_0 \neq 0$, $\xi_0 < 1$, et de paramètre d'échelle σ_0 . Dans ce cas, en vertu de (5), la "vraie" valeur de $e(a)$ (notée arbitrairement M) sera

$$M = \frac{\xi_0}{1 - \xi_0} a + \frac{\sigma_0}{1 - \xi_0}$$

Si l'on augmente la volatilité à un niveau σ_1 pour mieux capter les grandes variations observées, (ce qui revient à chercher à obtenir cette "vraie" valeur de $e(a)$ au moyen de σ_1) dans le cas $\xi = 0$, en vertu de (5), il faudrait que $\sigma_1 = M$, soit

$$\sigma_1 = \frac{\xi_0}{1 - \xi_0} a + \frac{\sigma_0}{1 - \xi_0}$$

ce qui donne la valeur du forçage de volatilité. Par exemple, si $\xi_0 = 2/3$, on aura $M = 2a + 3\sigma_0$, et donc il faudra forcer la volatilité à $\sigma_1 = 2a + 3\sigma_0$.

Mais, ce faisant, on agit sur une distribution de type gaussienne en augmentant simultanément le nombre des variations moyennes, variations qui sont indésirables car d'aucune utilité pour la gestion des risques. En sorte que cette opération de forçage de volatilité accentue l'écart entre la distribution théorique utilisée, qui draine un trop grand nombre de variations moyennes, et la distribution réelle des marchés, avec peu de variations moyennes. On ne peut substituer une forme normale, même dilatée par une opération d'échelle, à une forme non normale de

la "fragilité des modèles dont les coefficients varient lorsque la période d'estimation économétrique est modifiée", et Malinvaud [1981, p.25] voyait dans ces "imprécisions" une des raisons pour lesquelles "des thèses antagonistes peuvent subsister aussi longtemps en économie politique sans que l'observation de la réalité arbitre entre elles".

distribution : il s'agit de deux mondes statistiques différents. Insistons sur ce point : on ne peut remplacer l'absence d'un paramètre de forme par une augmentation d'un paramètre de taille. Les conséquences indésirables d'un tel forçage commencent à être bien perçues par les professionnels : on peut ainsi lire dans une revue professionnelle que la Value-at-Risk utilisant la loi normale "n'est pas utile pour les gérants de portefeuille : elle est trop grande dans la plupart des cas, et trop petite le reste du temps"¹⁹. Ces considérations permettent d'introduire la notion de risque de modèle dans le cas de distributions à queues épaisses²⁰.

Proposition 1 *Il y a risque de modèle en présence d'une queue de distribution épaisse lorsqu'on cherche à compenser l'absence d'un paramètre de queue par l'augmentation d'un paramètre de dispersion.*

Implicitement présente depuis de nombreuses années, la notion de risque de modèle a connu un développement récent²¹. C'est devenu une préoccupation actuelle des établissements financiers, dont la presse professionnelle se fait écho : "Models imperilled by volatility turmoil"²², "Faulty pricing models leave financial institutions at a loss"²³, "Report allays some model risk fears"²⁴. Dans une récente étude sur les modèles d'allocation de fonds propres en présence de risques de marché faite par la cabinet d'audit Deloitte & Touche, 49% des banques interrogées ont indiquées ne pas être satisfaites de leur modèle d'évaluation du risque de marché, et le comportement des modèles en cas de crise s'est révélé défaillant pour 58% des banques.

Ces difficultés de modélisation ont conduit à abandonner progressivement comme critère unique de risque le seul paramètre de volatilité, à chercher à étendre la mesure du risque au-delà de ce paramètre²⁵, et à examiner le risque de modèle de volatilité (Alami et Renault [2000]). Le risque de modèle présenté ici dans la proposition 1 rejoint la recommandation de Gouriéroux et Le Fol [1997, p. 9] mettant en garde les acteurs des marchés contre l'insuffisance de la seule volatilité, et la notion de risque de volatilité défini par Alami et Renault [2000, p. 104] comme résultant de la variabilité de la volatilité et de son insuffisance pour la caractérisation de la dynamique des rentabilités. En effet, l'on a montré que, en présence de queues épaisses, la volatilité est affectée par le poids des queues, et il n'est donc pas certain que "toute la dynamique pertinente (des rentabilités) passe par la volatilité"²⁶, une autre dynamique (propre aux innovations) restant à modéliser. Plus précisément, l'hypothèse proposée ici est que l'on attribue ce risque à l'absence d'un paramètre de forme de queue de distribution dans la définition de la dynamique des rentabilités utilisée dans la gestion usuelle des risques. Gouriéroux et Le Fol [1997] ont proposé de compléter la volatilité par des mesures décomposant des effets de transactions et de volume ; on suggère ici de compléter la volatilité par un paramètre de queue de distribution de type ξ . Le risque est alors bidimensionnel, selon le couple (σ, ξ) , et le risque de modèle vient de la réduction

¹⁹"European pension funds rethink risk", *Risk*, novembre 1999, pp 29-32

²⁰Sur la notion de risque de modèle dans un cas plus général, voir Gibson *et al.* [1999] et Risklab Project in Model Risk [2000].

²¹Voir le compte-rendu de la journée d'étude "Risque de modèle" organisée par le groupe Finance-Assurance de la Société Française de Statistique (SFdS) en collaboration avec la Fédération Française des Actuaire (FFA), 2 mars 2000, publié dans le Journal de la SFdS, tome 141, n°1-2, 2000, pp 71-211.

²²*International Financing Review*, 1300, 11 septembre 1999, p. 85.

²³*International Financing Review*, 1316, 15 janvier 2000, p. 100.

²⁴*International Financing Review*, 1319, 5 février 2000, p. 84.

²⁵Voir une revue dans Gouriéroux et Le Fol [1997].

²⁶Alami et Renault [2000, p. 132]

unidimensionnelle du couple (σ, ξ) sur σ . Une convergence entre ces deux manières de compléter la volatilité serait cependant à examiner, dans la mesure où l'existence de queues de distribution ($\xi > 0$) renvoie à l'analyse de la microstructure des marchés, c'est-à-dire à la prise en compte des effets complémentaires de volume ou d'autres caractéristiques des transactions.

3.2 Impact des queues de distribution sur la VaR

3.2.1 La VaR et les grandes variations

Rappel de la définition de la VaR On sait depuis les travaux de Maurice Allais sur l'attitude face au risque (Allais [1953]), que les investisseurs prudents sont plus sensibles à leur perte maximale éventuelle qu'à leur espérance de gain. Cette sensibilité à la perte trouve un écho dans un outil de mesure de risque de plus en plus utilisé par les professionnels des marchés, la "value-at-risk" (VaR), ou perte maximale probable sur un horizon de temps donné pour un niveau de probabilité fixé²⁷. Plus précisément, on a la définition suivante :

Définition 4 Pour un horizon de temps τ et un niveau de confiance c donnés, la "value-at-risk" d'un actif financier S , notée VaR_c , est la perte qui est dépassée à cet horizon avec la probabilité $1 - c$:

$$\Pr(\Delta S_{t,\tau} \leq -VaR_c) = 1 - c \quad (20)$$

où $\Delta S_{t,\tau} = S(t) - S(t - \tau)$ est la variation de cours de S examinée sur l'intervalle de durée τ .

En utilisant la fonction de répartition $F(\cdot)$, la relation (20) s'écrit simplement :

$$F_{\Delta S}(-VaR_c) = 1 - c$$

Autrement dit, $VaR_c = -F_{\Delta S}^{-1}(1 - c)$ où F^{-1} est l'inverse de F . Si la distribution de $\Delta S_{t,\tau}$ est symétrique (ce qui n'est pas nécessairement vérifié), $-F^{-1}(1 - c) = F^{-1}(c)$ et l'on a simplement $VaR_c = F_{\Delta S}^{-1}(c)$.

Les standards du comité de Bâle précisent que les calculs doivent être effectués pour un seuil de confiance $c = 99\%$ sur un horizon $\tau = 10$ jours. Il est clair que le calcul de la VaR fournira des résultats différents selon le type de fonction de répartition $F(\cdot)$ utilisée.

La VaR dans le modèle standard D'après le modèle standard décrit par l'EDS (1), $\Delta S(t, \tau) \stackrel{d}{=} N(0, S(t - \tau)\sigma\tau)$ avec les hypothèses usuelles du calcul pratique des VaR²⁸.

²⁷Sur la VaR, voir par exemple Jorion [1997] pour une synthèse de la question. La VaR a été critiquée par Artzner *et al.* [1997, 1999] sur deux points importants : a) ce n'est pas une mesure de risque cohérente car elle n'est pas sous-additive ; b) elle ne permet pas d'estimer un montant de perte potentielle au-delà du quantile choisi. Dans le premier cas, on montre que la VaR de deux sous-portefeuilles peut être inférieure à la VaR du portefeuille total. Dans le second cas, on introduit une VaR conditionnelle qui représente une perte probable sachant atteint et dépassé un quantile fixé, définie en (2).

²⁸On pose $\Delta S(t, \tau) = S(t - \tau)\Delta R(t, \tau)$ avec $\Delta r(t, \tau) = \ln(1 + \Delta R(t, \tau))$. Il résulte de l'équation (1) que $\Delta r(t, \tau) = (\mu - \sigma^2/2)\tau + \sigma\Delta W(t, \tau)$, avec $\Delta W(t, \tau) = W(t) - W(t - \tau)$. Pour τ "petit" (hypothèse usuelle dans

Ainsi par exemple, la VaR à 1 jour et 99% d'un actif financier admettant σ_1 comme volatilité quotidienne et cotant $S(0)$ en date $t = 0$, est donnée par

$$\Pr(\Delta S(1) \leq -\text{VaR}_{0,99}(1)) = 0,01$$

où $\Delta S(1) = S(1) - S(0) = S(0)\sigma_1\Delta W(1)$ avec $\Delta W(1) \stackrel{d}{=} N(0,1)$.

D'où

$$\begin{aligned} \text{VaR}_{0,99}(1) &= F_{\Delta S(1)}^{-1}(0,99) \\ &= S(0)\sigma_1 F_{\Delta W(1)}^{-1}(0,99) \\ &= S(0)2,3263\sigma_1 \end{aligned}$$

et plus généralement

$$\text{VaR}_c(1) = S(0)q_c\sigma_1 \tag{21}$$

où q_c est le quantile d'ordre c de la distribution normale centrée réduite.

Pour une VaR à 10 jours, on utilise la loi d'échelle du mouvement brownien telle que $\sigma_{10} = \sigma_1\sqrt{10}$, ce qui donne $\text{VaR}_c(10) = S(0)q_c\sigma_1\sqrt{10}$.

La VaR pour des seuils élevés Grâce aux travaux sur les valeurs extrêmes mentionnés plus haut, et à la prise de conscience récente de l'importance de considérer ces valeurs extrêmes plutôt que les moyennes dans la gestion des risques, il devient aujourd'hui mieux perçu que l'approximation classique de la réalité financière fournie par la modélisation (1) représente un risque de modélisation pour la VaR dans le cas des seuils de confiance de 99% et au-delà. Par exemple, on peut lire dans un article récent intitulé "Beyond Basle", que "cette manière de voir va rapidement devenir *complètement dépassée*"²⁹ (souligné par nous). La nécessité de généraliser la distribution normale, c'est-à-dire de passer d'un cadre gaussien à un cadre parétien pour l'intégration des valeurs extrêmes, apparaît de plus en plus nettement.

Deux modélisations sont en train d'être explorées simultanément pour le calcul de VaR avec grandes variations : une VaR utilisant la théorie des valeurs extrêmes (Longin [1997], McNeil [1999]), et une VaR utilisant les distributions α -stables, soit pures (Khindanova *et al.* [1999], Rachev et Mitnik [2000]), soit tronquées par une exponentielle (Bouchaud et Potters [1997]).

Pour illustrer l'impact des événements rares sur la VaR, on présente ci-dessous un calcul de VaR avec des distributions de probabilités non gaussiennes, et on fait apparaître les écarts à la normalité. Les applications de la théorie des valeurs extrêmes à la VaR étant plus connues que celles des distributions α -stables, on choisit ici un calcul de VaR dans ce dernier cas, avec des lois α -stables symétriques, qui constituent une classe particulièrement simple des lois stables.

le calcul pratique des VaR), $\Delta r(t, \tau) \cong \Delta R(t, \tau)$ et $(\mu - \sigma^2/2)\tau$ est négligeable devant $\sigma\Delta W(t, \tau)$. Ce qui conduit à $\Delta R(t, \tau) = \sigma\Delta W(t, \tau)$, soit à $\Delta S(t, \tau) = S(t - \tau)\sigma\Delta W(t, \tau)$.

²⁹"Beyond Basle", *Risk*, Janvier 1998, p.89.

3.2.2 Exemple de calcul de VaR avec des distributions α -stables

On rappelle³⁰ qu'une distribution stable univariée est caractérisée par quatre paramètres : l'exposant de stabilité α , le paramètre d'asymétrie β , le paramètre d'échelle γ , et le paramètre de localisation δ . Dans le cas où $\alpha = 2$, la distribution stable est gaussienne, $\beta = 0$, γ est proportionnel à l'écart-type, et δ est la moyenne. Plus précisément :

Définition 5 Une variable aléatoire X est appelée α -stable s'il existe des paramètres $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\gamma \geq 0$, et $\delta \in \mathbb{R}$ tels que sa fonction caractéristique s'écrive de la façon suivante :

$$\Phi_X(t) = E(\exp itX) = \exp \{i\delta t - \gamma^\alpha |t|^\alpha (1 - i\beta \operatorname{sgn}(t)W(\alpha, t))\}, t \in \mathbb{R}$$

où

$$W(\alpha, t) = \begin{cases} \tan \frac{\pi\alpha}{2} & \text{si } \alpha \neq 1 \\ -\frac{2}{\pi} \log |t| & \text{si } \alpha = 1 \end{cases}$$

et

$$\operatorname{sgn}(t) = \begin{cases} 1 & \text{si } t > 0 \\ 0 & \text{si } t = 0 \\ -1 & \text{si } t < 0 \end{cases}$$

Les quatre paramètres qui caractérisent une loi stable s'interprètent comme suit :

- L'exposant caractéristique α décrit la forme de la distribution ou son degré de leptokurticité. Les queues de distributions s'étirent de plus en plus en fonction de la décroissance de α .
- β est un paramètre d'asymétrie. Si $\beta = -1$, on dit que la distribution est totalement asymétrique à gauche. Si $\beta = 0$, on dit que la distribution est symétrique. Si $\beta = 1$, la distribution est dite totalement asymétrique à droite.
- γ est un paramètre d'échelle. Quand $\alpha = 2$, ce paramètre est égal à $\sigma/\sqrt{2}$ où σ est l'écart-type.
- Le paramètre de localisation δ est un paramètre de centrage. Quand $\alpha > 1$, il est égal à la moyenne.

La figure 3 montre les diverses formes prises par les densités en faisant varier le paramètre α .

FIG. 3 – Densités des lois α -stables.

Notation 3 $X \stackrel{d}{=} S_{\alpha,\beta}(\gamma, \delta)$ signifie que X est distribuée selon une loi stable de paramètres α , β , γ , et δ .

Une variable aléatoire stable est appelée symétrique si sa distribution est symétrique.

Définition 6 $X \stackrel{d}{=} S_{\alpha,\beta}(\gamma, \delta)$ est symétrique si $\beta = \delta = 0$. X est symétrique par rapport à δ si $\beta = 0$.

³⁰On trouvera des éléments sur les variables aléatoires stables dans, par exemple, Samorodnitsky et Taqqu [1994]

Notation 4 $X \sim S\alpha S$ signifie que X suit une loi stable symétrique $S_{\alpha,0}(\gamma, 0)$

D'après les définitions 5 et 6, on voit que si X est $S\alpha S$, alors sa fonction caractéristique s'écrit sous une forme très simple :

$$\Phi_X(t) = \exp(-\gamma^\alpha |t|^\alpha), \quad t \in \mathbb{R}.$$

De plus, si $\gamma = 1$, alors la variable $S\alpha S$ est appelée *standard*, ou réduite et

$$\Phi_X(t) = \exp(-|t|^\alpha), \quad t \in \mathbb{R}.$$

Dans la suite de ce paragraphe, tous les résultats donnés correspondent à des variables $S\alpha S$ réduites.

Notation 5 F_α désigne la fonction de répartition d'une variable $S\alpha S$ réduite.

Notation 6 q_c^α désigne le quantile d'ordre c d'une variable $S\alpha S$ réduite.

Le calcul de la VaR à 1 jour dans le cas de lois α -stables se fait alors en généralisant l'équation (21) de la manière suivante

$$\text{VaR}_c^\alpha(1) = S(0)q_c^\alpha \gamma_1 \tag{22}$$

où γ_1 est la volatilité quotidienne de l'actif S mesurée avec le paramètre d'échelle des lois α -stables. L'estimation de γ_1 s'effectue au moyen de méthodes statistiques adaptées à ce type de lois³¹.

Pour une VaR à 10 jours, on utilise la loi d'échelle du mouvement α -stable telle que $\gamma_{10} = \gamma_1 10^{1/\alpha}$, ce qui donne $\text{VaR}_c^\alpha(10) = S(0)q_c^\alpha \gamma_1 10^{1/\alpha}$.

L'équation (22) permet de calculer des VaR avec des lois α -stables. Le calcul de la fonction de répartition $F_\alpha(x) = \int_{-\infty}^x f_\alpha(t) dt$ doit s'effectuer numériquement au moyen d'un développement en série, les lois α -stables n'ayant pas de fonction de densité explicite. Le développement en série de ce calcul est donné en annexe. Le tableau 9 en annexe présente une tabulation des fonctions de répartition F_α en fonction de α . Les tableaux 10 et 11 en annexe sont une autre présentation du tableau 9 : le tableau 10 fournit les valeurs de $1 - F_\alpha$, soit la probabilité de perte en fonction de α et x . Le tableau 11 réinterprète le précédent en termes de durée de retour d'un événement selon α et x . La durée de retour pour une valeur x est définie comme $d_\alpha(x) = 1/(1 - F_\alpha(x))$.

Correspondance avec la loi normale Dans ces tableaux, la loi de Gauss est définie comme une loi stable d'exposant caractéristique $\alpha = 2$ et les valeurs indiquées correspondent aux valeurs usuelles de la gaussienne centrée réduite à une constante d'échelle près, qui résulte de l'égalité

³¹Voir Walter [1994] ou Lévy Véhel et Walter [2002] pour une présentation générale de ces méthodes.

c	1	1,2	1,4	1,6	1,8	2	$N(0,1)$
99%		16,1600	9,6576	6,2768	4,2766	3,2899	2,3263
98%	15,8945	9,1635	6,0613	4,3601	3,3915	2,9044	2,0537
97%	10,5784	6,5924	4,6571	3,5780	2,9785	2,6598	1,8808
96%	7,9155	5,2277	3,8778	3,1247	2,7079	2,4758	1,7507
95%	6,3119	4,3687	3,6999	2,8143	2,5049	2,3262	1,6448

TAB. 5 – Quantiles élevés q_c^α en fonction de α .

$\gamma = \sigma/\sqrt{2}$ lorsque $\alpha = 2$. Par exemple, on lit dans la colonne “Gauss” du tableau 9 que $F_2(3) = \Pr(x \leq 3) = 0,9831$. L'équivalence avec la gaussienne centrée réduite est obtenue comme suit :

$$\begin{aligned} \Pr(x_{S_{2,0}(1,0)} \leq 3) &= \Pr(x_{N(0,\sqrt{2})} \leq 3) \\ &= \Pr(x_{N(0,1)} \leq 3/\sqrt{2}) \\ &= \Pr(x_{N(0,1)} \leq 2,12) \end{aligned}$$

et de manière générale,

$$\begin{aligned} \Pr(x_{S_{2,0}(1,0)} \leq t) &= \Pr(x_{N(0,\sqrt{2})} \leq t) \\ &= \Pr(x_{N(0,1)} \leq t/\sqrt{2}) \end{aligned}$$

Distance à la loi normale Les queues de distributions des lois stables s'étirant de plus en plus en fonction de la décroissance de l'exposant caractéristique α , on observe l'augmentation de la valeur des quantiles, donc de la VaR. Pour une valeur de x donnée, on peut comparer directement la probabilité cumulée de chaque loi, et observer comment la décroissance de l'exposant caractéristique α affecte l'allure, et donc la probabilité de la queue de distribution. Par exemple, dans le tableau 9, pour $x = 4$, $F_2(4) = 0,9977$ (gaussienne), tandis que $F_{1,5}(4) = 0,9694$, et $F_1(4) = 0,9220$ pour une loi de Cauchy ($\alpha = 1$). On remarque aussi que, pour les grandes valeurs de x , par exemple $x = 8$, alors que la valeur de la probabilité cumulée est proche de 1 pour une gaussienne, cette valeur est encore éloignée de 1 pour une stable non gaussienne (par exemple 0,9764 pour $\alpha = 1,2$).

VaR avec des lois stables Le tableau 5 donne les valeurs de q_c^α en fonction de α . Par exemple, pour $c = 95\%$, le quantile gaussien usuel est $q_{0,95} = 1,6448$. Le quantile 2-stable équivalent est $q_{0,95}^2 = q_{0,95}\sqrt{2} = 2,3262$, et le quantile 1,4-stable est $q_{0,95}^{1,4} = 3,6999$. Pour $c = 99\%$, on observe une forte croissance sur $q_{0,99}^\alpha$ en fonction de la baisse de α vers la valeur 1 (loi de Cauchy). En particulier, pour $\alpha = 1,6$ qui est une valeur souvent obtenue sur les marchés réels, $q_{0,99}^{1,6} = 6,2768$, soit 1,9 fois plus élevé que $q_{0,99}^2$.

Risque de modèle Pour une probabilité c , le ratio q_c^α/q_c^2 représente une mesure de l'impact de la non normalité sur la VaR. En effet

$$\frac{\text{VaR}_c^\alpha(1)}{\text{VaR}_c^2(1)} = \frac{S(0)q_c^\alpha\gamma_1}{S(0)q_c^2\gamma_1} = \frac{q_c^\alpha}{q_c^2} \quad (23)$$

FIG. 4 – VaR en fonction de α

L'on peut quantifier un *risque de modèle* au sens de la proposition 1 vue plus haut, au moyen de cette mesure.

Facteur 3 des standards de Bâle On peut interpréter le coefficient multiplicatif de 3 à appliquer à la VaR demandé par les standards de Bâle en terme de ratio q_c^α/q_c^2 , et chercher la valeur implicite de α correspondant à cette exigence. En effet, soit VaR^B la VaR demandée par les standards de Bâle : $\text{VaR}_c^B = \text{VaR}_c^2 \times 3$. En posant $\text{VaR}_c^B = \text{VaR}_c^\alpha$, et d'après (23), on doit avoir $q_c^\alpha/q_c^2 = 3$.

Par exemple, pour $c = 99\%$, on veut que $q_{0,99}^\alpha/q_{0,99}^2 = 3$, soit $q_{0,99}^\alpha \cong 9,87$, ce qui conduit à $\alpha \cong 1,4$. Cela veut dire que, pour une probabilité de 99%, en appliquant le coefficient 3, tout se passe comme si l'on faisait l'hypothèse que le processus réel des cours était engendré par un mouvement 1,4-stable. Cette valeur est différente si l'on change le niveau de probabilité. Pour $c = 97\%$, $q_{0,97}^\alpha \cong 7,98$, ce qui conduit à $\alpha \cong 1,1$. On voit que le risque de modèle n'est pas résolu par l'introduction du facteur 3.

Fonctionnelle de VaR On peut, avec des lois stables, remplacer une mesure unique de VaR par une fonctionnelle de VaR selon les valeurs de α . Le graphique 4 illustre cette approche. La représentation graphique de 4 est faite au moyen de la méthode dite "droite de Henry". La courbe d'une fonction de répartition gaussienne se transforme en droite si elle est reportée sur un repère approprié, arithmétique en abscisse et gaussien en ordonnée. L'épaisseur des queues des lois α -stables apparaît, à la mesure de la déviation des points obtenus par rapport à la droite. Pour chaque valeur de la fonction de répartition d'une loi stable, correspondant à une valeur de x , on calcule $F^{-1}(x)$ à partir d'une loi normale centrée : on obtient le quantile correspondant de la loi normale. Considérons par exemple une loi stable d'exposant $\alpha = 1,4$. Pour $x = 3,9$, on a $F_{1,4}(x) = 0,9604$. On cherche alors le quantile x' de la loi normale telle que $F_2(x') = 0,9604$. On trouve $x' = 2,4824$. Le point $(3,9 ; 2,4824)$ est alors positionné dans le graphe. Pour une loi normale, on aurait naturellement trouvé 3,9, et le point aurait été situé sur la droite de Henry.

L'estimation de l'exposant caractéristique α peut être réalisée au moyen de méthodes utilisant la statistique d'ordre ou la fonction caractéristique (méthodes spectrales)³².

3.3 Impact des queues de distribution sur la gestion d'actifs

3.3.1 La concentration de la performance

Dans la gestion d'actifs, la leptokurticité se traduit par un phénomène courant quoique peu documenté : la concentration de la performance des portefeuilles sur les valeurs extrêmes des marchés. Les valeurs extrêmes des variations boursières captent l'essentiel des gains ou des pertes

³²Voir Walter [1994] ou Lévy Véhel et Walter [2002].

valeurs rangées	effectif cumulé (rang)	fréquence cumulée $F_{\Delta r}(\cdot)$	performance cumulée (contribution)	fréquence cumulée $\text{Fr}(\cdot)$
$\Delta r_{(1)}$	1	$1/n$	$\Delta r_{(1)} = r_{(1)}$	$r_{(1)}/r_{(n)}$
$\Delta r_{(2)}$	2	$2/n$	$\Delta r_{(1)} + \Delta r_{(2)} = r_{(2)}$	$r_{(2)}/r_{(n)}$
\vdots	\vdots	\vdots		\vdots
$\Delta r_{(k)}$	k	k/n	$\Delta r_{(1)} + \cdots + \Delta r_{(k)} = r_{(k)}$	$r_{(k)}/r_{(n)}$
\vdots	\vdots	\vdots		\vdots
$\Delta r_{(n)}$	n	1	$\Delta r_{(1)} + \cdots + \Delta r_{(n)} = r_{(n)}$	1

TAB. 6 – Construction de la courbe de Gini-Lorenz de concentration de la performance

FIG. 5 – Concentration de la performance positive du CAC 40

des gestions sur longue période : on observe une très forte hiérarchisation des contributions à la performance totale des journées de bourse, ou de titres particuliers. On précise à présent ces notions, en se limitant cependant à leur description statistique, remettant la modélisation probabiliste du phénomène à un travail ultérieur. On indiquera seulement ici la voie d'une telle modélisation.

Le cadre de modélisation reste celui de la section précédente. La période d'analyse $[0, T]$ de la gestion d'un portefeuille est subdivisée en N intervalles de longueur τ , et le marché de référence est analysé selon la même subdivision. On considère les N variables aléatoires, $\{\Delta r_1, \dots, \Delta r_N\}$, dont on fait l'hypothèse qu'elles sont i.i.d. entre les dates 0 et T , dont la densité de probabilité est $P(\Delta r)$ et dont la fonction de répartition est $F_{\Delta r}(x) = \Pr(\Delta r \leq x) = \int_{-\infty}^x P(t)dt$. On considère les n variations de même signe (par exemple positives) de Δr , rangées par ordre décroissant $\{\Delta r_{(1)}, \dots, \Delta r_{(n)}\}$.

La performance positive totale du marché sur la période $[0, T]$ est $r_{(n)} = \sum_{k=1}^n \Delta r_{(k)}$. On s'intéresse à la concentration de la performance positive. Pour cela, on considère la somme des k plus grandes variations positives, soit $r_{(k)} = \Delta r_{(1)} + \Delta r_{(2)} + \cdots + \Delta r_{(k)}$. Puis on introduit la notion de *contribution* des k plus grandes variations positives à la performance totale, ou fonction de répartition de la masse positive cumulée à la performance positive totale, qui est égale à $r_{(k)}/r_{(n)}$. On construit alors une courbe de Gini-Lorenz de concentration de la performance au moyen du tableau 6.

Le résultat de cette construction pour les variations positives de l'indice CAC 40 entre le 1^{er} janvier 1988 et le 30 juin 2001, avec $\tau = 1$ jour, soit $N = 3460$ cours, dont $n = 1856$ variations positives, est donné sur le graphique 5. Le même calcul est ensuite effectué pour les $n = 1604$ variations négatives de l'indice sur la même période globale : graphique 6. On observe une nette concentration de la performance positive (resp. négative) de l'indice sur les grandes variations positives (resp. négatives) apparues au cours de la période d'analyse.

FIG. 6 – Concentration de la performance négative du CAC 40

FIG. 7 – Concentration de la performance totale du CAC 40

Proposition 2 *La concentration de la performance des portefeuilles est le symptôme de la leptokurticité dans la gestion d’actifs.*

Formellement, pour analyser le problème de la concentration de la performance, on introduit la variable aléatoire

$$T_n(p) = \frac{\Delta r_{(1)} + \dots + \Delta r_{([np])}}{r_{(n)}} \quad \frac{1}{n} < p \leq 1 \quad (24)$$

où $[np]$ est la partie entière de np . La quantité $T_n(p)$ est la contribution des $[np]$ plus grandes variations (par exemple) positives de la période. Par exemple, si $p = 0,2$, et si $T_n(0,2) = 0,8$, cela veut dire que 20% des variations positives contribuent pour 80% de la performance positive totale du marché (c’est la relation dite “loi des 20-80”, bien connue des professionnels, selon laquelle 20% des dossiers contiennent 80% des résultats et réciproquement ; mais aussi connue en économie, par laquelle 20% des pays produisent 80% des richesses etc.³³).

On introduit ensuite la fonction

$$D_F(p) = \frac{1}{\mu_F} \int_{1-p}^p F^{-1}(y) dy, \quad p \in [0, 1] \quad (25)$$

où $F(\cdot)$ est la fonction de répartition d’une variable aléatoire Y , F^{-1} l’inverse de F et

$$\mu_F = EY = \int_0^\infty y dF(y) = \int_0^1 F^{-1}(y) dy$$

avec $EX < \infty$.

On montre alors³⁴ que, lorsque $n \rightarrow \infty$,

$$\sup_{p \in [0,1]} |T_n(p) - D_F(p)| \rightarrow 0 \quad p.s.$$

et $D_F(p)$ est un indice de mesure de concentration, au sens où il quantifie la façon dont les 100p% plus grandes variations positives contribuent à la performance totale positive. Nous laissons pour un travail ultérieur l’application systématique de cette fonction sur un échantillon de portefeuilles réels gérés.

Notons que ces deux concentrations des variations positives et négatives (graphiques 5 et 6) peuvent être regroupées en s’agréant dans un même effet global, celui de l’impact des hausses ou des baisses sur la performance totale des N jours, ceci sans distinction des variations négatives ou positives. De manière empirique, cela revient à évaluer l’effet sur la performance totale des k plus grandes variations positives (resp. négatives), en recalculant cette performance après retrait de ces k variations. Le graphique 7 présente un tel résultat.

Les queues de distribution épaisses ont donc un impact important sur la répartition de la performance des marchés et des portefeuilles. L’on s’intéresse à présent aux conséquences de cet impact sur l’organisation des modes de gestion.

³³On trouvera de nombreux exemples de cette “loi” empirique dans Zajdenweber [2000]

³⁴Voir Embrecht *et al.* [1997], chapitre 8.

3.3.2 La revalorisation de l'importance des choix de titres

Nous voulons ici aborder un deuxième point peu examiné dans la recherche en finance, la forme des processus d'investissement, en dégageant les implications du phénomène de concentration de la performance mis en évidence précédemment.

Cette question de la forme des processus d'investissement fait aujourd'hui l'objet de nombreux débats professionnels, au centre desquels se trouve la compréhension qu'ont les gestionnaires de la notion d'efficience des marchés. Cette présence du concept d'efficience dans les controverses entre gestionnaires apparaît clairement dans les études professionnelles consacrées au sujet. Par exemple, dans un rapport consacré à la présentation de la gestion indicielle, on peut lire : "les arguments théoriques (pour ou contre l'indexation) renvoient à l'hypothèse de l'efficience des marchés" (Bishop [1990], p.7). Dans un autre rapport publié à l'occasion du 25ème anniversaire de la création de la gestion indicielle, on trouve cette assertion : "le débat pour ou contre l'indexation tourne autour des différentes *manières de comprendre* l'efficience des marchés" (PwC [1999], p.3, souligné par nous). Ces manières de *voir* l'efficience renvoient, entre autres, à une manière de *définir un type d'aléa*.

Dans la mesure où le phénomène leptokurtique conduit à modifier les types d'aléas entrant dans la modélisation financière, il a donc nécessairement un impact sur les applications de la notion d'efficience dans la gestion d'actifs. Il a été montré dans un travail précédent³⁵ comment et pour quelle raison les processus d'investissement issus d'une conception gaussienne de l'efficience des marchés ne pouvaient être que du type "top-down", et privilégier la définition des portefeuilles-cibles "stratégiques", au détriment du rôle des choix de titres. On veut montrer ici comment le phénomène leptokurtique peut conduire au contraire à revaloriser la contribution à la performance des choix de titres, et à mettre en place des processus de type inverse, c'est-à-dire "bottom-up".

Queues exponentielles, efficience gaussienne et gestion indicielle Les termes du débat étant liés à la compréhension de la notion d'efficience de marché, on rappelle la forme de cette définition qui est instrumentée par les professionnels, et qui est celle³⁶ de Jensen [1978, p. 96] :

Définition 7 *Un marché boursier est dit "efficent" par rapport à un ensemble d'information donné, s'il est impossible de réaliser un gain boursier en mettant en oeuvre une politique d'intervention ou de gestion fondée sur cet ensemble d'information.*

En d'autres termes, on ne peut pas tirer parti d'un avantage informationnel car, comme le disent les opérateurs des marchés, l'information utile est déjà "dans les cours". Cette définition 7 indique clairement que toute gestion active conduite par des gérants professionnels expérimentés, et aidés par des analystes financiers cherchant à repérer les valeurs à acheter, tout ces efforts ne peuvent pas permettre de réaliser des performances supérieures à celles obtenues par un indice représentatif du marché. C'était déjà la conclusion à laquelle était parvenu Cowles [1933], pour

³⁵Walter [1996, 2000].

³⁶La définition classique de Fama [1970] n'a pas de traduction opérationnelle et conduit au paradoxe de Grossman et Stiglitz [1980]. C'est la raison pour laquelle on utilise plutôt la formulation de Jensen. Sur ces questions liés aux aspects informationnels du concept d'efficience, voir une synthèse dans Roger [1988].

qui la meilleure stratégie d'investissement possible était donc la constitution d'un portefeuille de marché.

De manière plus formelle, soit Δr^P la rentabilité d'un portefeuille P et Δr^B la rentabilité d'un portefeuille de marché, ou "benchmark" B . La définition 7 signifie que l'espérance mathématique de l'écart de gain du portefeuille géré activement par rapport à un indice représentatif du marché, résultant d'une stratégie de gestion fondée sur un ensemble d'information partagé par le marché, est nulle pour un niveau de risque identique, et conduit à la relation du CAPM : $E(\Delta r^P - r^0) - \beta^P E(\Delta r^B - r^0) = 0$, où β^P est le coefficient β du portefeuille P et r^0 le taux sans risque sur la période.

Soit $X_n = (\Delta r_n^P - r^0) - \beta^P (\Delta r_n^B - r^0)$ l'écart de performance observé en date n pour la période $[n-1, n]$ entre un indice de marché B et un portefeuille P géré activement. On observe au fil du temps une suite d'écarts $\{X_n, n > 0\}$. Le problème posé par la compréhension et l'application professionnelles de l'efficacité de Jensen³⁷ peut alors s'analyser comme suit. Etant donné une suite X_1, \dots, X_n de variables aléatoires i.i.d., on se pose la double question de la forme de la distribution limite de leur somme $X_1 + \dots + X_n$ et de leurs maxima, convenablement normalisés, lorsque n tend vers l'infini. La première question renvoie au théorème de la limite centrée; la deuxième question à la théorie des valeurs extrêmes.

La réponse à la première question est la suivante. Lorsque $E(X_i^2)$ est finie, la seule loi limite est la loi de Laplace-Gauss. Dans tous les autres cas, on se trouve en présence de lois α -stables. La réponse à la deuxième question est la suivante. Quand le paramètre ξ défini en (4) est positif, la distribution est à queue épaisse. Dans tous les autres cas, on se trouve en présence de distributions à queues fines.

Seuls les cas $E(X_i^2) < \infty$ et $\xi \leq 0$ ont été considérés dans les débuts de la formation du concept d'efficacité : forme gaussienne, queues fines. Dans ces cas-là, la convergence des moyennes empiriques a lieu "normalement", et les écarts de performance X_n sont limités à du "bruit" autour de zéro. Dans ces cas, les partisans de la gestion indiciaire peuvent faire valoir, à juste titre, qu'il vaut mieux reproduire mécaniquement un indice de marché plutôt que d'engager des coûts importants pour l'entretien d'équipes de gérants qui, en définitive, s'avèrent inutiles dans la mesure où ils ne créent pas de valeur ajoutée (Cowles [1933]); et qui, selon la fameuse et provocatrice boutade que Samuelson lança en 1974, devraient plutôt se recycler comme cadres dans l'industrie³⁸. Plus généralement, les sources de performance résultant d'une démarche CAPM mettent en évidence la contribution principale des choix d'allocations stratégiques par rapport aux choix de titres.

De fait, la profession de la gestion d'actifs, en vertu de ces deux hypothèses ($E(X_i^2) < \infty$ et $\xi \leq 0$) implicites mais réelles, s'est alors lancée en 1970 dans la gestion indiciaire passive avec la création du premier fonds indiciaire sur le New York Stock Exchange par Wells Fargo, pour la société Samsonite³⁹. Cette manière de "voir" l'efficacité (forme gaussienne, queues fines) constitue le *fondement conceptuel de la gestion indiciaire dite "passive"*. L'efficacité des marchés couplée aux hypothèses ($E(X_i^2) < \infty$ et $\xi \leq 0$) apparaît ainsi à premier abord comme l'*ultima*

³⁷On ne considère pas ici la question de la variation des β^P , ni celle des taux sans risque r^0 .

³⁸Voir Samuelson [1974].

³⁹Voir Walter [2000] pour plus de détails sur l'histoire de la création de la gestion indiciaire à la suite de l'hypothèse gaussienne.

ratio de la gestion indiciaire passive. Les processus d'investissement de type "top down" résultent de cette perspective probabiliste, et consistent à considérer comme "bruit" de performance les contributions des choix de titres. Ce déroulement historique précis, avec les enchaînements de causes et d'effets associant efficacité à normalité et queues de distribution fines, est la raison pour laquelle on a proposé⁴⁰ l'hypothèse selon laquelle la gestion indiciaire passive et les formes "top down" des processus d'investissement, sont une conséquence, *non* de l'hypothèse d'efficacité seule, mais de celle de *normalité surajoutée* à l'efficacité.

Queues parétiennes, efficacité non gaussienne et choix de titres On fait à présent l'hypothèse que Δr^B et Δr^P suivent une distribution de Pareto généralisée définie en (4) avec un indice de queue $\xi > 0$. D'où il vient que l'écart de performance X suit une distribution de Pareto généralisée (DPG). D'après la définition de l'efficacité, $E(X) = 0$. Mais si X suit une DPG à queue épaisse, alors on a vu que les grandes valeurs de X peuvent avoir un impact important sur les moyennes. Autrement dit, si *en moyenne*, l'écart de performance entre une gestion active et un indice de marché est nul, il peut cependant exister des écarts temporaires importants. Si, avec un aléa gaussien, le recyclage des gérants suggéré par Samuelson peut (au moins intellectuellement...) trouver un fondement théorique, il n'en est plus de même avec un aléa parétien. Dans cette situation, une espérance nulle d'écart de performance n'est alors plus contradictoire avec la pertinence d'une gestion active.

Par exemple, supposons que les écarts de performance suivent une loi de Cauchy, définie par

$$\Pr(a < X < b) = \int_a^b \frac{dx}{\pi(1+x^2)} = \frac{1}{\pi}(\text{Arc tg } b - \text{Arc tg } a)$$

Pour cette distribution, l'espérance mathématique (l'écart de performance moyen) n'existe pas, car l'intégrale $\int_{-\infty}^{+\infty} \frac{x dx}{1+x^2}$ n'est pas définie. Les moyennes empiriques obtenues par tirages aléatoires de Cauchy ne convergent pas. On se trouve alors devant une configuration probabiliste dans laquelle il pourrait exister pendant de longues périodes un écart de performance important entre un indice de marché et une gestion active de même niveau de risque. Dans ce cas, l'adage selon lequel "la patience réduit le risque" (Hammer [1991]) ne s'appliquera pas, en raison de l'apparition de fortes fluctuations transitoires, dues à l'inexistence des moments. Ou alors il faudra être très patient... Une hypothèse d'efficacité des marchés, couplée avec des distributions de Pareto en remplacement des queues exponentielles, peut donc conduire à justifier, non la reproduction passive ou semi-active d'indices de marchés, mais des processus d'investissement dans lesquels les choix de titres produisent une contribution non négligeable à la performance finale.

Proposition 3 *La concentration de la performance des portefeuilles conduit à la revalorisation des choix de titres dans les processus d'investissement au sein de la profession de la gestion d'actifs.*

Comme indiqué précédemment, nous laissons pour un travail ultérieur l'investigation systématique de la concentration de la performance sur quelques titres particuliers pendant une période donnée de gestion active. En particulier, la question de la différence entre la concentration de la performance des indices et celle des titres spécifiques.

⁴⁰Walter [1996,2000]

Le tableau 7 présente une typologie de la profession de la gestion d'actifs qui fait apparaître les interactions entre processus d'investissement et phénomène leptokurtique.

Réponses de la profession de la gestion d'actifs D'un point de vue théorique, de nouveaux modèles d'allocation d'actifs sont apparus en réponse au phénomène leptokurtique. En approche statique monopériodique, des extensions des modèles de Markowitz et Sharpe pour la gestion des portefeuilles en présence de grandes fluctuations, ont été faites par Bouchaud *et al.* [1998], Gamrovski et Rachev [1999], Belkacem *et al.* [2000]. En approche dynamique, incluant la prise en compte d'une information asymétrique, un courant important de recherche a modélisé les processus d'investissement dans un cadre multipériodique, dont on trouvera un exemple avec Basjeux-Besnaimou et Portait [1998]. Ces modélisations ont en commun la revalorisation des choix de titres et des micro-décisions, par rapport aux allocations stratégiques et aux macro-décisions, dans les contributions des actes de gestion à la performance totale. Enfin, l'apparition de nouvelles formes de gestion non indicées ("alternative management", soit "variantes à la gestion classique") peut être interprété comme le signe d'une défiance des investisseurs pour l'indexation complète (gestion indicielle passive) ou partielle (gestion quantitative référencée à un indice).

D'autre part, dans la pratique, on observe une distorsion de la notion d'allocation stratégique. Bien que les allocations "stratégiques" soient supposées rester stables dans le temps, en général, beaucoup de gérants de portefeuille réévaluent en permanence leurs estimations de facteurs de performance⁴¹, ce qui conduit à réviser périodiquement les allocations stratégiques des portefeuilles, allocations qui, de fait, perdent alors leur caractère "stratégique". On trouve ici l'équivalence du phénomène de l'affaiblissement de la pertinence des valeurs moyennes, décrit plus haut, appliqué à la gestion d'actifs : les allocations stratégiques jouent le rôle de structure-cible "moyenne" du portefeuille, c'est-à-dire devant produire "en moyenne" la performance-cible.

En résumé, les principales conséquences du phénomène leptokurtique pour l'industrie de la gestion d'actifs sont au nombre de trois : concentration de la performance des portefeuilles qui conduit à l'affaiblissement de la pertinence des allocations stratégiques, changement de forme du processus d'investissement, apparition de nouvelles formes de gestion non indicées.

4 Interprétations théoriques du phénomène leptokurtique : développement de nouveaux modèles

Pour rendre compte du phénomène leptokurtique, on peut utiliser de manière commode la grille de lecture suivante : soit on attribue cet effet à des causes externes aux marchés, soit l'on fait des marchés eux-mêmes la cause de l'effet. Pour la première manière de comprendre les fluctuations extrêmes, la non normalité des distributions des variations boursières n'est que la transposition sur les marchés financiers de la non normalité des variables de l'économie "réelle". La propriété d'efficience des marchés assure cette transmission de chocs non normaux de la réalité économique dans les variations des prix. La leptokurticité est dans ce cas externe au marché. La seconde manière de voir les grandes variations prend l'option opposée : le phénomène leptokurtique n'est que le produit d'une amplification par les agents économiques de chocs normaux,

⁴¹Par exemple, voir Ghysels [1998].

	Processus d'investissement de type "top down"	Processus d'investissement de type "bottom up"
Approche quantitative du risque, avec modélisation probabiliste explicite des aléas, et usage de moyens de calculs appropriés	I <ul style="list-style-type: none"> – gestion indicielle ou indicée, passive ou semi-active, privilégiant les grandes décisions de type "macro" (secteurs, pays) sur les décisions de type "micro" (choix de titres), avec des modèles de calibrage de risque – situation majoritaire en Europe aujourd'hui, en déclin aux Etats-Unis 	II <ul style="list-style-type: none"> – gestion active par choix de titres quantifiée par des modélisations paréliennes utilisant la théorie des valeurs extrêmes, permettant de donner un prix au risque spécifique (non rémunéré en approche gaussienne)
Approche qualitative sans modélisation de risque	III <ul style="list-style-type: none"> – incompatibilité entre l'approche qualitative et les processus d'investissement de type "top down" 	IV <ul style="list-style-type: none"> – gestion classique par recherche des bonnes valeurs selon des critères de type "analyse fondamentale" sans considération a priori sur les allocations stratégiques – situation dominante dans la profession jusque dans les années soixante, existe encore aujourd'hui à l'état de survivance

TAB. 7 – La profession de la gestion d'actifs et le phénomène leptokurtique.

Une catégorisation de la profession de la gestion d'actifs est faite selon les deux critères de la modélisation quantitative et de la forme des processus d'investissement. Chronologiquement, l'histoire de la gestion d'actifs a commencé dans le quadrant IV : choix de titres sans modélisation probabiliste du risque. C'est la conception dite "classique" de la gestion, qui s'oppose (encore aujourd'hui chez certains gérants) à l'utilisation de modèles probabilistes. Puis, à partir des années soixante, sous l'influence des travaux de Markowitz et Sharpe, s'est développée une conception quantitative des choix de portefeuilles, conduisant aux processus d'investissement de type "top down", c'est-à-dire instrumentant l'efficience gaussienne en privilégiant les décisions d'investissement de type sectorielles ou géographiques : c'est le quadrant I, considéré par les gestionnaires quantitatifs comme "moderne". L'impact du phénomène leptokurtique devrait conduire l'industrie de la gestion à opérer un déplacement vers des processus d'investissement de type "bottom up", c'est-à-dire à revaloriser les gestions actives avec choix de titres, comme contributions importantes à la performance totale ; mais processus "bottom up" d'un nouveau genre, car quantifiés au moyen de la théorie des valeurs extrêmes et de la distribution de Pareto généralisée : c'est le quadrant II. Le quadrant III ne peut exister.

mais démultipliés par des effets de mimétisme qui peuvent conduire à des ruptures de marché. La leptokurticité devient alors interne au marché.

La variable qui différencie les deux approches est l'information. L'information dont disposent, et qu'utilisent, les agents économiques peut être de deux natures : exogène ou endogène. Par information exogène, on entend une information relative à l'environnement économique ou financier dans lequel évolue le marché, décrite par exemple par les trois formes de Fama [1970], c'est-à-dire des éléments qui concernent l'économie dans laquelle les agents forment leurs choix de portefeuille. Selon cette définition, entrent dans la catégorie d'information exogène, des éléments bilantiels d'entreprises, des données économiques, ou toute autre information sur les facteurs non techniques agissant sur le marché (c'est-à-dire externes au marché). Au contraire, une information endogène caractérisera des éléments internes au marché, c'est-à-dire propres aux agents eux-mêmes : on entre alors dans les modèles d'apprentissage et d'imitation, où l'information est obtenue par observation des autres agents, conduisant aux analyses mimétiques des mouvements boursiers, dans le sens de Orléan [1989]. Nous présentons ces deux approches.

4.1 Leptokurticité, information et mimétisme

4.1.1 Leptokurticité d'origine externe au marché : l'économie des extrêmes

La justification économique classique de la distribution normale utilise l'argument d'un grand nombre de petits chocs d'information exogène, homogènes et non hiérarchisés. En faisant l'hypothèse qu'il n'existe pas d'aléa sur l'offre, que la seule information pertinente est exogène, et que l'efficacité du marché assure l'intégralité de la transmission de l'information exogène agrégée dans le prix, c'est donc cette hypothèse sur la nature des chocs d'information qui est à reconsidérer. Au lieu d'être percuté par un très grand nombre de petits chocs d'information homogènes et non hiérarchisés, le marché subirait plutôt l'effet de chocs hétérogènes et hiérarchisés, c'est-à-dire se répartissant entre un petit nombre de grands chocs, et un grand nombre de chocs insignifiants. Par exemple, l'annonce d'une OPA, ou d'une fusion entre sociétés, peut avoir plus d'impact que les résultats trimestriels publiés régulièrement. Cette conception revient à affaiblir la pertinence de la notion de "petit choc moyen", à l'instar de l'affaiblissement de celles de "perte moyenne" ou de "gain moyen" que nous avons évoqué précédemment.

En 1965, Fama avait suggéré que "si de grandes variations de cours surviennent fréquemment, cela peut vouloir indiquer que la structure de l'économie à l'origine de ces variations, est elle-même sujette à de grands et fréquents changements" (Fama [1965], p. 41). En cela, la non normalité est d'origine extérieure au marché et cela revient à chercher dans les variables de l'économie "réelle" des distributions non normales. Précisément, Samuelson [1972, pp 30,31] observait que "de telles distributions ultra-étirées se manifestent fréquemment en économie". De fait, on observe souvent des distributions non normales dans un grand nombre de situations économiques réelles : la taille des entreprises, les chiffres d'affaires annuels, la répartition des richesses, la population des pays etc. Zajdenweber [2000] a présenté un échantillon important de résultats, et a qualifié d'"économie des extrêmes" cette économie réelle où les quantités observées sont modélisables par des lois de puissance de type Pareto définie en (4).

On a vu comment ces distributions parétiennes se caractérisent par une forte hiérarchisation des contributions de chaque individu statistique au phénomène total et une concentration du résultat total sur quelques individus statistiques. Cette présence d'événements extrêmes en économie avait été qualifiée d'"effet Noé" par Mandelbrot [1973], par analogie avec l'épisode biblique du déluge, dans la mesure où une cascade informationnelle (un "déluge d'information") parvient soudainement sur le marché et se transmet dans les prix. La propriété d'efficience des marchés assure que la structure parétienne de la "nature" de l'économie se transmet intégralement dans les prix, en sorte que :

Hypothèse 1 *La structure leptokurtique des variations boursières reflète la structure parétienne de l'économie.*

Selon cette hypothèse, l'économie des extrêmes, décrite par la théorie des valeurs extrêmes, se transmet dans les variations extrêmes des marchés, qui peuvent donc être aussi décrites par la théorie des valeurs extrêmes.

4.1.2 Leptokurticité d'origine interne au marché : l'interaction extrême

La définition de l'efficience au sens de Fama [1976] fait intervenir la notion de "pertinence" de l'information, de "significativité" d'un choc d'information. L'amplitude que peut produire un choc sur la variation des prix, dépend de l'interprétation que les opérateurs de marché font de cette information. Un élément de subjectivité est introduit dans la hiérarchisation des chocs d'information exogène, et une formalisation du traitement de l'information devient nécessaire. L'hypothèse d'anticipations rationnelles avait pour objet la résolution de cette question, mais il est connu depuis Tirole [1982] que cette hypothèse ne suffit pas à permettre la réalisation d'un équilibre du marché : une polarisation des anticipations est possible, qui conduit le marché sur n'importe quel point fixe arbitraire. On observe bien ce phénomène de subjectivisation de la valeur de l'information en suivant l'évolution du choix des quantités économiques retenues comme pertinentes par les opérateurs. Des phénomènes de modes font passer sur le devant de la scène des indicateurs macro-économiques qui étaient considérés comme non significatifs la période précédente. Ainsi, selon les moments, c'est le nombre des mises en chantiers, le solde du commerce extérieur, le taux de chômage, le taux directeur de la banque centrale, ou d'autres, qui deviennent les indicateurs "à suivre", et qui provoquent donc les plus grands chocs d'information.

Orléan [1989] a montré comment l'introduction de la subjectivité des agents dans l'explication de la variation des cours instaure un régime d'opinion sur les marchés et comment le mimétisme pouvait affecter la réalisation d'un équilibre. A l'extrême, lorsque toute référence à l'information exogène disparaît, les modèles de type "tache solaire" illustrent les situations à information endogène pure. Selon cette conception, le système d'information représenté par les prix est un miroir déformant de l'économie : ce n'est pas une économie "objective" qui se transmet au système de prix par la propriété d'efficience, mais une économie filtrée par les croyances de ceux qui l'interprètent, c'est-à-dire une économie "intersubjective". Plus le mimétisme est généralisé, plus les opinions des opérateurs s'agglutinent pour interpréter les informations exogènes, plus les volumes de transactions augmentent, plus le marché s'emballe. Dans cette perspective :

Hypothèse 2 *La structure leptokurtique des variations boursières reflète la structure mimétique du comportements des agents.*

Selon cette hypothèse, la leptokurticité des marchés est d'origine endogène : les marchés sont la source de leur leptokurticité. Des modèles récents de formation des prix illustrent cette conception : par exemple Johansen, *et al.* [1998] ou Bouchaud et Cont [1998].

Cette intersubjectivité de l'information conduit à poser la question de la pertinence du temps "objectif", c'est-à-dire calendaire, pour mesurer les variations boursières : à l'instar du caractère intersubjectif de l'économie filtrée par les marchés, le temps boursier ne serait-il pas aussi un temps intersubjectif? Cette interrogation est développée ci-dessous.

4.2 Leptokurticité, durée du temps, et valeur du temps

En mesurant une "valeur" du temps par la quantité d'information déversée sur le marché, on voit que des intervalles de temps identiques n'ont pas la même valeur pour la mesure des variations des cours. Puisqu'une même durée du temps ne contient pas la même quantité d'information pertinente selon les moments du marché, on peut interpréter la défaillance de la distribution normale à représenter les répartitions de variations réelles des marchés mesurées à intervalles réguliers, comme un signe de ce que, pour les opérateurs des marchés, le temps écoulé n'a pas la même valeur selon les périodes d'observation. Ce que l'on peut formuler de manière suivante : sur les marchés, *la valeur du temps n'est pas égale à la durée du temps*. Pendant de longues périodes, aucune information disponible n'est considérée comme importante, et, comme disent les boursiers, "il ne se passe rien d'important en ce moment". Puis, tout-à-coup, l'environnement semble changer brutalement, et, pendant une très courte période, une quantité d'information considérée comme extrêmement importante déferle sur le marché, aboutissant à de forts décalages de cours. Sur les marchés, si l'on considère le critère de l'information importante, "mille ans sont comme un jour, et un jour est comme mille ans". Dans ce sens, on peut reformuler le problème de la non normalité en disant que :

Hypothèse 3 *La structure leptokurtique des variations boursières reflète la disjonction entre durée du temps physique et valeur du temps physique, la valeur étant appréciée par une quantité d'information.*

Ces considérations peuvent laisser supposer que le temps physique n'est pas nécessairement la bonne horloge pour mesurer la variabilité des cours, et donc le risque des marchés. Les marchés évolueraient plutôt en fonction d'une horloge interne, d'un "temps boursier", ou "temps interne", ou "temps intrinsèque", différent du temps calendaire, et qui correspondrait à la quantité d'information considérée comme pertinente par les opérateurs. Le "temps intrinsèque" serait davantage un temps intersubjectif, tandis que le temps calendaire est un temps objectif. Dès lors, l'étape intellectuelle suivante la plus naturelle est de se demander si la bonne manière de mesurer les fluctuations boursières ne serait pas d'introduire un changement de temps dans l'analyse des distributions, en "redressant" les distributions par une modification de l'échelle de temps.

Le nouveau temps pertinent serait, dans ce sens, un temps conditionné par la quantité d'information : un temps-information tel que, par un redressement adéquat du temps, la même quantité d'information parvienne sur le marché par périodes successives. La déformation serait d'autant plus importante que le flux d'information serait dense. Pour utiliser une image suggestive, on peut imaginer que tout se passe comme si le "poids" de l'information considérée comme pertinente "pliait" le temps physique, le contractait ou le dilatait selon les périodes de marché. Dans les marchés très calmes, le temps s'écoulerait plus lentement que le temps physique, tandis que, dans des marchés très agités, le temps s'écoulerait plus rapidement. On peut considérer par ailleurs que l'information peut ne pas être seulement exogène, mais relever aussi du poids des opinions identiques des acteurs. Autrement dit, dans ce courant de pensée, tant l'économie des extrêmes que le mimétisme des opérateurs sont des facteurs de déformation du temps des marchés.

Il faut ensuite mesurer cette "quantité d'information". En général (voir par exemple Harris [1986] et Chauveau [1997]), on fait l'hypothèse d'une causalité entre la taille du volume de transactions et celle de la variation du cours : le volume de titres échangés représenterait alors une mesure adéquate de la "quantité d'information" contenue dans le prix. Comme les professionnels des marchés disent souvent que la valeur d'un prix coté (et donc la pertinence de la mesure) n'est pas la même selon que ce prix correspond à un échange de cinq cent mille titres ou de cinq titres, faisant intervenir une notion de "profondeur" du marché, de "poids" de l'échange, la relation entre information et volume permet de donner une signification précise à un code langagier communément accepté.

Cette conceptualisation des marchés au moyen de l'introduction d'un temps déformé conduit à définir un cadre particulier de modélisation du temps, et se poser la question du choix du temps adéquat pour la mesure des variations boursières. Ceci a été rendu possible grâce aux données de haute fréquence. On dispose, par exemple entre deux dates 0 et T , de la totalité des entregistrements de cotations et volumes. A partir de cette série d'échanges ("transactions"), on peut donc construire deux séries chronologiques différentes pour l'analyse du marché, selon que l'on agrège les échanges par le nombre de minutes passées (temps d'horloge) ou par le nombre de titres échangés (temps-volume).

Temps de l'horloge Dans le premier cas, on relève les cours cotés entre 0 et T selon une fréquence d'observation fixe. Cela revient à subdiviser l'intervalle $[0, T]$ en n périodes égales de pas de temps de base $\tau = T/n$, cette durée τ définissant un "temps caractéristique" d'observation du marché. La variation du cours entre 0 et T est dans ce cas

$$S_T = S_0 + \sum_{k=1}^n \Delta S_k \quad n = T/\tau \quad (26)$$

où $\Delta S_k = S_k - S_{k-1}$ représente la variation du cours de l'actif S entre les dates $t - \tau$ et t , où $t = k\tau$ est exprimé en multiple du pas de temps de base τ .

Temps des échanges Dans le deuxième cas, on relève chaque cours coté à chaque échange ayant lieu entre 0 et T , ce qui revient à partitionner l'intervalle $[0, T]$ en $0 = t_0 < t_1 < \dots < t_n = T$, et à travailler en temps des échanges ("temps-transaction"), t_j étant le moment du j -ième échange. La variation du cours entre 0 et T est dans ce cas

$$S_T = S_0 + \sum_{j=1}^{N(T)} \Delta S_j \quad (27)$$

où $\Delta S_j = S_j - S_{j-1}$ représente la variation du cours de l'actif S entre les échanges $j - 1$ et j . Le processus $\{S_j\}$ des cours est alors indicé par un temps des échanges $\theta(t)$. Soit $V(T)$ le volume total de titres échangés entre 0 et T :

$$V_T = \sum_{j=1}^{N(T)} v_j \quad (28)$$

où v_j représente le volume de titres échangés au cours de la transaction j . Le processus $\{V_j\}$ des volumes est indicé par le temps des échanges. Dans ce cadre d'analyse, le cours coté en date T est donc la résultante de trois facteurs, ou trois processus, entre 0 et T : le processus des échanges $N(t)$, le processus des variations du cours entre deux échanges ΔS_j , le processus des volumes v_j .

L'étape suivante est donc :

- soit l'étude de la distribution des volumes d'échange (voir par exemple Maillet et Michel [1997])
- soit l'étude de la distribution des durées entre échanges (voir par exemple Ghysels *et al.* [1997])

Il est intéressant d'observer comment le passage en un temps opérationnel peut conduire à la disparition de la leptokurticité. Par exemple, Maillet et Michel [1997] proposent un redressement par les volumes, et utilisent pour la description des volumes un diagramme rang-fréquence selon la loi de Zipf définie en (18) introduite par Mandelbrot [1968] à partir des travaux du linguiste américain Zipf sur la distribution de la fréquence des mots dans le discours. Cette loi indique que les mots d'un texte se distribuent de telle sorte que, si l'on classe les mots par fréquence décroissante en une liste de ces fréquences, la fréquence d'un mot est inversement proportionnelle à son rang dans la liste : l'interprétation qui en est faite est que, plus un mot est rare dans un texte, et plus il contient de l'information. Adaptée aux mouvements boursiers, cette relation rang-fréquence revient à formaliser l'hypothèse selon laquelle les grandes variations des cours sont porteuses d'une information essentielle pour la compréhension du mécanisme gouvernant le marché. Les volumes les plus importants associés aux grandes variations correspondraient donc à la transmission dans les prix d'informations considérées comme extrêmement pertinentes pour l'avenir des marchés, par l'ensemble des opérateurs. Tandis que les petites variations et les volumes faibles traduiraient un faible contenu informationnel. Après renormalisation par le temps-volume, Maillet et Michel constatent une disparition du phénomène leptokurtique par un retour à une distribution normale des rentabilités⁴².

Plus généralement, les résultats obtenus avec l'agrégation des rentabilités et des volumes conduisent à retrouver un cadre gaussien pour les variations des cours : le temps opérationnel introduit par Mandelbrot et Taylor [1967], puis Clark [1973], mais récemment redécouvert, permettrait donc d'éviter le recours à des processus complexes, en conservant un mouvement brownien, mais avec un temps boursier : c'est la notion de processus subordonné (voir par exemple pour une étude récente, Ané et Geman [2000]).

En conclusion de cette partie, on peut avancer que ce cadre théorique nouveau permettra vraisemblablement un rapprochement entre les approches universitaire et techniques (au sens de

⁴²Le test porte sur les cotations en continu de l'action Elf Aquitaine entre mars et août 1996, soit 179 958 cotations.

l'analyse technique) des marchés, dans la mesure où les études de microstructure complétées par la prise en compte des effets de volume dans la variabilité des cours, rejoignent progressivement les préoccupations des analystes techniques. Dès l'instant où l'on admet que l'information endogène peut avoir une influence sur le niveau des cours, on réhabilite de fait simultanément l'attention portée par les analyses techniques aux comportements collectifs passés des opérateurs agissant sur les marchés. En un sens, on peut dire que le phénomène leptokurtique est un facteur de réintégration de l'analyse technique dans la théorie financière.

5 Conclusion

Le phénomène leptokurtique est devenu un fait d'expérience omniprésent dans la finance de marché. Nous avons retracé sa permanence au cours du temps dans les analyses des variations boursières, et montré comment il s'est transformé de "non observable" en "fait scientifique", au fur et à mesure que les préoccupations des professionnels se déplaçaient des valeurs moyennes vers les valeurs extrêmes, sous la pression des approches prudentielles internationales. Nous avons illustré ce déplacement de deux manières différentes : d'une part, quantitative, au moyen d'un remplacement de VaR gaussienne par une fonctionnelle de VaR en distributions α -stables symétriques, faisant apparaître les modifications résultant d'une prise en compte explicite des queues de distributions non gaussiennes ; en calculant également la concentration de la performance de l'indice CAC 40 sur 13 ans ; d'autre part, qualitative, en proposant une typologie de l'industrie de la gestion d'actifs qui incorpore l'hypothèse de normalité dans la construction des processus d'investissement, pour faire apparaître l'impact de la non normalité dans les manières de gérer l'épargne. Nous avons enfin utilisé un cadre interprétatif simple de lecture de la non normalité, en qualifiant l'information selon son origine : exogène ou endogène.

En résumé, on conjecture que le phénomène leptokurtique devrait être à l'origine d'un renouvellement important des formalisations en finance de marché, d'une reconceptualisation de certains éléments de cette finance, ouvrant la voie à de nouvelles pratiques dans l'industrie financière, et à une généralisation de la théorie de l'efficience incluant la prise en compte des événements rares, événements devenant alors assimilés à des événements "normaux" de la vie des marchés. Et, à partir du moment où les fluctuations extrêmes des marchés sont comprises comme faisant partie de leur fonctionnement normal, les crises financières, les bulles spéculatives et les krachs boursiers deviennent constitutifs de la structure même des bourses : toute mise en place d'un marché contiendrait alors en germe les conditions de ses propres dérèglements, de sa propre instabilité intrinsèque.

Les développements théoriques futurs pourront donc concerner, soit des redéfinitions du risque et de la rentabilité des actifs financiers, et leur relation ; soit des modélisations de comportements d'agents aboutissant aux fluctuations extrêmes ; soit des études sur l'analyse organisationnelle des bourses, la microstructure des marchés, pour déterminer les conditions d'efficacité maximale des marchés dans un environnement financier structurellement leptokurtique.

A Annexe : calcul numérique de la fonction de répartition d'une loi α -stable

Soit X une variable α -stable réduite, de paramètres α et β , de fonction de densité $f_{\alpha,\beta}$ et de fonction de répartition $F_{\alpha,\beta} : F_{\alpha,\beta}(x) = \Pr(X < x) = \int_{-\infty}^x f_{\alpha,\beta}(t)dt$. Dans la mesure où $f_{\alpha,\beta}$ n'a pas de forme analytique explicite, il est nécessaire de calculer numériquement $F_{\alpha,\beta}$ pour pouvoir utiliser simplement les lois α -stables. Ce calcul numérique utilise les représentations en série des lois α -stables introduites par Humbert [1945] et Bergström [1952], et repris par Fama et Roll [1968], Janicki et Weron [1994] et Rachev et Mittnik [2000] dans le cas symétrique. On présente ici ces développements en série.

On montre⁴³ que :

1. pour $|x| \rightarrow 0$, la fonction de densité $f_{\alpha,\beta}$ admet la représentation en série

$$f_{\alpha,\beta}(x) = \frac{1}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\Gamma\left(\frac{k+1}{\alpha}\right)}{\alpha} x^k \cos \left[k \left(\frac{\pi}{2} + \frac{\beta}{\alpha} \right) + \frac{\beta}{\alpha} \right] \quad (29)$$

2. pour $|x| \rightarrow \infty$, la fonction de densité $f_{\alpha,\beta}$ admet la représentation en série

$$f_{\alpha,\beta}(x) = -\frac{1}{\pi} \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \frac{\Gamma(\alpha k + 1)}{x|x|^{\alpha k}} \sin k \left(\frac{\alpha\pi}{2} + \beta - \alpha \arg x \right) \quad (30)$$

Dans le cas de variables $S\alpha S$ réduites, (29) et (30) se simplifient en posant $\beta = 0$. Pour de telles variables, dans le cas où $\alpha > 1$, Fama et Roll [1968, p. 819], Janicki et Weron [1994, p. 48] et Rachev et Mittnik [2000, p. 109] donnent les expressions suivantes de la fonction de répartition $F_{\alpha}(x) = \int_{-\infty}^x f_{\alpha}(t)dt$

1. pour $|x| \rightarrow 0$

$$F_{\alpha}(x) = \frac{1}{2} + \frac{1}{\pi\alpha} \sum_{k=1}^{\infty} (-1)^{k-1} \frac{\Gamma\left(\frac{2k-1}{\alpha}\right)}{(2k-1)!} x^{2k-1} \quad (31)$$

2. pour $|x| \rightarrow \infty$

$$F_{\alpha}(x) = 1 + \frac{1}{\pi} \sum_{k=1}^n \frac{(-1)^k}{k!} \frac{\Gamma(\alpha k)}{x^{\alpha k}} \sin \frac{\alpha k \pi}{2} - o \left(\frac{x^{-\alpha(n+1)}}{\alpha(n+1)} \right) \quad (32)$$

Les expressions (31) et (32) permettent de tabuler la fonction de répartition d'une variable $S\alpha S$. La question à résoudre est de savoir à quel moment x est "petit" ou "grand", c'est-à-dire quelle forme de F_{α} utiliser. Il s'agit de déterminer une valeur critique x_c pour laquelle on passe de (31) à (32). On aura

$$F_{\alpha}(x) = \begin{cases} F_{\alpha}(x) & \text{selon (31) si } x \leq x_c \\ F_{\alpha}(x) & \text{selon (32) si } x > x_c \end{cases}$$

Pour cela, il a été procédé à une investigation empirique des valeurs de $F_{\alpha}(x)$ pour α variant de 1,1 à 2 par pas de 0,1 et x variant de 0 à 40 par pas de 0,01. Il est apparu que la valeur

⁴³Voir Bergström [1952], p. 376.

α	x_c
1,1	1,6
1,2	2,3
1,3	3,2
1,4	4,0
1,5	5,1
1,6	6,3
1,7	7,6
1,8	9,1
1,9	10,7

TAB. 8 – Valeurs critiques x_c en fonction de α .

critique x_c qui déclenche le passage de (31) à (32) dépend de celle de α . On a alors déterminé l'ensemble des valeurs critiques x_c , fonction de α , pour obtenir une relation entre x_c et α . Le résultat est donné dans le tableau 8.

B Annexe : tabulations de la fonction de répartition d'une loi α -stable

α x	Cauchy 1	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	Gauss 2
2, 0	0,8524	0,8625	0,8718	0,8802	0,8879	0,8950	0,9013	0,9071	0,9123	0,9170	0,9214
2, 1	0,8585	0,8690	0,8787	0,8875	0,8956	0,9030	0,9098	0,9159	0,9215	0,9266	0,9312
2, 2	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 3	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 4	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 6	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 7	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 8	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
2, 9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 1	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 2	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 3	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 4	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 6	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 7	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 8	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4, 9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
5, 0	0,9372	0,9482	0,9677	0,9660	0,9731	0,9793	0,9847	0,9893	0,9934	0,9968	0,9998
6, 0	0,9474	0,9576	0,9663	0,9736	0,9797	0,9848	0,9891	0,9927	0,9956	0,9980	1,0000
7, 0	0,9548	0,9643	0,9721	0,9786	0,9839	0,9882	0,9918	0,9948	0,9969	0,9986	1,0000
8, 0	0,9604	0,9692	0,9764	0,9821	0,9868	0,9905	0,9935	0,9958	0,9976	0,9990	1,0000
9, 0	0,9648	0,9730	0,9798	0,9848	0,9889	0,9922	0,9947	0,9966	0,9981	0,9992	1,0000
10, 0	0,9683	0,9760	0,9820	0,9868	0,9905	0,9934	0,9956	0,9972	0,9985	0,9994	1,0000
11, 0	0,9711	0,9784	0,9840	0,9884	0,9917	0,9943	0,9962	0,9977	0,9987	0,9995	1,0000
12, 0	0,9735	0,9804	0,9856	0,9897	0,9927	0,9950	0,9967	0,9980	0,9989	0,9996	1,0000
13, 0	0,9756	0,9820	0,9870	0,9907	0,9935	0,9956	0,9971	0,9983	0,9991	0,9996	1,0000
14, 0	0,9773	0,9834	0,9881	0,9916	0,9942	0,9961	0,9975	0,9985	0,9992	0,9997	1,0000
15, 0	0,9788	0,9847	0,9891	0,9923	0,9947	0,9965	0,9977	0,9986	0,9993	0,9997	1,0000
16, 0	0,9801	0,9857	0,9899	0,9929	0,9952	0,9968	0,9980	0,9988	0,9994	0,9997	1,0000
17, 0	0,9813	0,9867	0,9908	0,9935	0,9956	0,9971	0,9982	0,9989	0,9994	0,9998	1,0000
18, 0	0,9823	0,9875	0,9912	0,9940	0,9959	0,9973	0,9983	0,9990	0,9995	0,9998	1,0000
19, 0	0,9833	0,9882	0,9918	0,9944	0,9962	0,9975	0,9985	0,9991	0,9995	0,9998	1,0000
20, 0	0,9841	0,9888	0,9923	0,9947	0,9965	0,9977	0,9986	0,9992	0,9996	0,9998	1,0000

TAB. 9 – Fonction de répartition F_α en fonction de α .

Le tableau ci-dessus, extrait de Walter [1994], donne les résultats des calculs des probabilités cumulées $F_\alpha(x) = \int_{-\infty}^x f_\alpha(t)dt$ des lois stables symétriques standards pour x variant de 2 à 20. Les lois stables d'exposant caractéristique α variant de 1 à 2 sont encadrées par les deux lois stables particulière à fonction de densité de forme analytique explicite : les lois de Cauchy ($\alpha = 1$) et de Gauss ($\alpha = 2$).

α x	1 Cauchy	1,2	1,4	1,6	1,8	2 Gauss	$N(0, 1)$
-1	24,92%	24,66%	24,45%	24,28%	24,13%	23,98%	-0,71
-2	14,76%	12,82%	11,21%	9,87%	8,77%	7,86%	-1,41
-3	10,24%	7,95%	6,02%	4,36%	2,93%	1,69%	-2,12
-4	7,80%	5,58%	3,81%	2,38%	<u>1,21%</u>	0,23%	-2,83
-5	6,28%	4,23%	2,69%	1,53%	0,66%	0,02%	-3,54
-6	5,26%	3,37%	2,03%	<u>1,09%</u>	0,44%	0,00%	-4,24
-7	4,52%	2,79%	1,61%	0,82%	0,31%	0,00%	-4,95
-8	3,96%	2,36%	1,32%	0,65%	0,24%		-5,66
-9	3,52%	2,04%	<u>1,11%</u>	0,53%	0,19%		-6,36
-10	3,17%	1,80%	0,95%	0,44%	0,15%		-7,07
-11	2,89%	1,60%	0,83%	0,38%	0,13%		-7,78
-12	2,65%	1,44%	0,73%	0,33%	0,11%		-8,49
-13	2,44%	1,30%	0,65%	0,29%	0,09%		-9,19
-14	2,27%	1,19%	0,58%	0,25%	0,08%		-9,90
-15	2,12%	1,09%	0,53%	0,23%	0,07%		-10,61
-16	1,99%	<u>1,01%</u>	0,48%	0,20%	0,06%		-11,31
-17	1,87%	0,94%	0,44%	0,18%	0,06%		-12,02
-18	1,77%	0,88%	0,41%	0,17%	0,05%		-12,73
-19	1,67%	0,82%	0,38%	0,15%	0,05%		-13,44
-20	1,59%	0,77%	0,35%	0,14%	0,04%		-14,14

TAB. 10 – Probabilité de perte : $\Pr(X > x) = 1 - F_\alpha(x)$ en fonction de x et α .

Le tableau ci-dessus donne la valeur de la probabilité de perte $\Pr(X > x) = 1 - F_\alpha(x)$ en fonction de x et α . Les lois α -stables d'exposant caractéristique α variant de 1 à 2 sont encadrées par les deux lois stables particulières à fonction de densité de forme analytique explicite : les lois de Cauchy ($\alpha = 1$) et de Gauss ($\alpha = 2$). On observe que la décroissance des probabilités de perte est rapide pour des valeurs élevées de α , et lente pour des valeurs faibles. Par exemple, pour $\Pr(\text{perte} > -5)$, on trouve 0,02% pour $\alpha = 2$, valeur qui augmente à 6,28% pour $\alpha = 1$. Avec la gaussienne, la conversion des quantiles s'effectue de la manière suivante : $\Pr(u_{S_{2,0}(1,0)} \leq -5) = \Pr(u_{N(0,\sqrt{2})} \leq -5) = \Pr(u_{N(0,1)} \leq -5/\sqrt{2}) = \Pr(u_{N(0,1)} \leq -3,54)$. Réciproquement, on peut se demander quelle est la valeur d'une perte x pour un niveau de probabilité fixé. Les valeurs soulignées font apparaître, pour un niveau de probabilité donné (dans l'exemple, voisin de 1%), l'ensemble des couples (α, x) possibles, et l'augmentation de la perte x en fonction de la décroissance de l'exposant caractéristique α . Par exemple, on trouve (approximativement) la même probabilité avec les couples $(1, 2; -16)$ et $(1, 8; -4)$. Ce qui veut dire qu'une perte de -16 est aussi probable avec une loi 1,2-stable que ne l'est une perte de -4 avec une loi 1,8-stable. Dans tous ces exemples, on suppose que $X \sim S_\alpha S$.

α x	1 Cauchy	1,2	1,4	1,6	1,8	2 Gauss	$N(0, 1)$
-1	4	4	4	4	4	4	-0,71
-2	7	8	9	10	11	13	-1,41
-3	10	13	17	23	34	59	-2,12
-4	13	18	26	42	<u>83</u>	435	-2,83
-5	16	24	37	65	152	4915	-3,54
-6	19	30	49	<u>92</u>	227	90531	-4,24
-7	22	36	62	122	323	$2,69 \cdot 10^6$	-4,95
-8	25	42	76	154	417	$1,30 \cdot 10^8$	-5,66
-9	28	49	<u>90</u>	189	526	$1,02 \cdot 10^{10}$	-6,36
-10	32	56	105	227	667	$1,30 \cdot 10^{12}$	-7,07
-11	35	62	120	263	769		-7,78
-12	38	69	137	303	909		-8,49
-13	41	77	154	345	1111		-9,19
-14	44	84	172	400	1250		-9,90
-15	47	92	189	435	1429		-10,61
-16	50	<u>99</u>	208	500	1563		-11,31
-17	53	106	227	556	1748		-12,02
-18	56	114	244	588	1942		-12,73
-19	60	122	263	667	2145		-13,44
-20	63	130	286	714	2500		-14,14

TAB. 11 – Durée de retour $d_\alpha(x) = 1/(1 - F_\alpha(x))$ en fonction de x et α .

Le tableau ci-dessus donne la valeur de la durée de retour $d_\alpha(x) = 1/(1 - F_\alpha(x))$ en fonction de x et α . Les lois α -stables d'exposant caractéristique α variant de 1 à 2 sont encadrées par les deux lois stables particulières à fonction de densité de forme analytique explicite : les lois de Cauchy ($\alpha = 1$) et de Gauss ($\alpha = 2$). Dans tous ces exemples, on suppose que $X \sim S_\alpha S$.

Le graphique ci-dessus représente les 1856 variations quotidiennes positives de l'indice CAC 40 entre le 1^{er} janvier 1988 et le 30 juin 2001 rangées par ordre croissant selon une représentation de type "amplitude-rang". En abscisse, l'amplitude Δr_k de la variation, mesurée en % quotidien, de la plus faible à la plus forte. En ordonnée, le rang de la variation, $k = \text{Nb}(\Delta r_i > \Delta r_{(k)})$, de $k = 1$ (la plus élevée) à $k = 1856$ (la plus faible). La plus forte variation a pour amplitude $\Delta r_{(1)} = 7,05\%$. Grâce au redressement bilogarithmique, les quelques points représentant les grandes variations (au-delà de 4%) apparaissent nettement isolés, ce qui illustre une queue de distribution importante pour les 20 plus grandes variations (sans valeurs extrêmes, on observerait un tassement de la courbe et une indistinction des points, comme c'est le cas pour les variations moins importantes). En divisant, sur l'axe des ordonnées, k par n , on obtient les valeurs k/n de la fonction de répartition empirique, et la courbe ci-dessus représente donc la courbe de la fonction de répartition empirique des variations positives du CAC40, ou courbe de fréquence cumulée $\text{Fr}(\Delta r_{(k)})$.

Le graphique ci-dessus représente les 1604 variations quotidiennes négatives de l'indice CAC 40 entre le 1^{er} janvier 1988 et le 30 juin 2001 rangées par ordre croissant en valeur absolue selon une représentation de type "amplitude-rang". En abscisse, la valeur absolue de l'amplitude Δr_k de la variation négative, mesurée en % quotidien, de la plus faible à la plus forte. En ordonnée, le rang de la variation, $k = \text{Nb}(\Delta r_i > \Delta r_{(k)})$, de $k = 1$ (la plus élevée) à $k = 1604$ (la plus faible). La plus forte variation a pour amplitude $\Delta r_{(1)} = -7,29\%$. Grâce au redressement bilogarithmique, les quelques points représentant les grandes variations (au-delà de 4%) apparaissent nettement isolés, ce qui illustre une queue de distribution importante pour les 20 plus grandes variations (sans valeurs extrêmes, on observerait un tassement de la courbe et une indistinction des points, comme c'est le cas pour les variations moins importantes). En divisant, sur l'axe des ordonnées, k par n , on obtient les valeurs k/n de la fonction de répartition empirique, et la courbe ci-dessus représente donc la courbe de la fonction de répartition empirique des variations négatives du CAC40, ou courbe de fréquence cumulée $\text{Fr}(\Delta r_{(k)})$.

Le graphique ci-dessus illustre la concentration de la performance positive de l'indice CAC 40 sur les plus fortes des 1856 variations quotidiennes positives, entre le 1^{er} janvier 1988 et le 30 juin 2001. Ce phénomène est représenté au moyen de la courbe de Gini-Lorenz. En abscisse, la répartition de l'effectif total rangé par ordre décroissant, soit k/n . 100% de l'effectif représentent les 1856 variations quotidiennes positives. En ordonnée, la répartition de la performance positive totale cumulée sur la durée, soit $r_{(k)}/r_{(n)}$. 100% de l'effectif représentent 1651,04% de performance positive cumulée. On observe que les 10% des variations positives les plus fortes (soit les 185 meilleures journées de bourse) représentent le tiers de la performance cumulée, et que les 30% des plus fortes variations positives contribuent à 60% de la performance totale. Si la performance était répartie de manière homogène, la courbe de Gini-Lorenz serait une droite : la bissectrice. L'écart à la bissectrice illustre que la répartition de la performance positive est inégalitaire, et se concentre sur les queues de distribution.

Le graphique ci-dessus illustre la concentration de la performance négative de l'indice CAC 40 sur les plus fortes des 1604 variations quotidiennes négatives entre le 1^{er} janvier 1988 et le 30 juin 2001. Ce phénomène est représenté au moyen de la courbe de Gini-Lorenz. En abscisse, la répartition de l'effectif total rangé par ordre décroissant, soit k/n . 100% de l'effectif représentent les 1604 variations quotidiennes négatives. En ordonnée, la répartition de la performance négative totale cumulée sur la durée, soit $r_{(k)}/r_{(n)}$. 100% de l'effectif représentent 1460,25% de performance négative cumulée. On observe que les 10% des variations positives les plus fortes (soit les 160

pires journées de bourse) représentent le tiers de la perte cumulée, et que les 30% des plus fortes variations négatives contribuent à 60% de la perte totale. Si la performance était répartie de manière homogène, la courbe de Gini-Lorenz serait une droite : la bissectrice. L'écart à la bissectrice illustre que la répartition de la performance négative est inégalitaire, et se concentre sur les queues de distribution.