

BERTRAND MAILLET

THIERRY MICHEL

Mesures de temps, information et distribution des rendements intra-journaliers

Journal de la société statistique de Paris, tome 138, n° 4 (1997), p. 89-120

http://www.numdam.org/item?id=JSFS_1997__138_4_89_0

© Société de statistique de Paris, 1997, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

MESURES DE TEMPS, INFORMATION ET DISTRIBUTION DES RENDEMENTS *INTRA*-JOURNALIERS

Bertrand MAILLET

Centre d'Economie Bancaire Internationale,
Université Paris I

Thierry MICHEL

Centre d'Economie Bancaire Internationale,
Université Paris I

Direction de la Prévision, Ministère de l'Economie et des Finances

Au cours de la préparation de cet article, nous avons bénéficié des commentaires de Thierry CHAUVEAU (CEBI, Université Paris I-CDC/FMR), Hélène RAYMOND (CEBI, Université Paris I), Richard TOPOL (OFCE), Catherine LUBOCHINSKY (Université Paris II), Gaëlle LE FOL (CEME, Université Paris I-CREST/ENSAE), Michel BOUTILLIER (Université d'Evry - CDC/FOE), Rahim LOUFIR (Université d'Evry), Philippe SPIESER (ESCP), François LONGIN (ESSEC), Hervé ALEXANDRE (Université de Besançon), ainsi que les deux rapporteurs du journal. Nous les remercions pour leurs commentaires, leurs suggestions et leurs conseils bibliographiques. Les erreurs, lacunes et imprécisions qui resteraient sont sous notre seule responsabilité. Une version préliminaire de cet article a été présenté au XIV^e colloque de l'AFFI, à Grenoble, en Juin 1997.

Correspondance :

Bertrand MAILLET, BMAILLET@asterix.univ-paris1.fr

Adresser tout courrier à
CEBI, Maison des Sciences Economiques,
106-112, boulevard de l'Hôpital, 75647 PARIS CEDEX 13.
Téléphone-Fax : 01.55.43.42.69.

Résumé

La conclusion principale de cette étude est que l'introduction du volume de transaction dans l'estimation du processus qui gouverne l'évolution du prix des actifs financiers permet d'améliorer celle-ci. Les caractéristiques des séries de rendements boursiers de haute fréquence sont incompatibles avec l'hypothèse standard de bruit gaussien. Nous montrons que les rendements mesurés à volume constant – *i.e.* en temps-volume – n'ont pas les mêmes propriétés que les rendements mesurés en temps-calendaire. Le volume de transaction, utilisé comme approximation de la quantité d'information arrivant sur le marché, permet, grâce à une renormalisation des cours, de reconstruire une série de rendements homoscedastiques et normaux.

Mots-clefs : volumes de transaction, effets ARCH, échelles de temps, distribution des rendements boursiers, données de haute fréquence.

Abstract

«Time measures, information and the distribution of the *intra*-day returns »

It is shown in this paper that the estimation of the price formation process can be improved by the use of transaction volume. The properties of high-frequency stock returns prove inconsistent with the usual gaussian noise hypothesis. The trading volume, as an approximation of the incoming information in the market, allows to define return series which are homoscedastic and normal. The distribution of returns, measured with constant volume, *i.e.* in the volume-time scale, are significantly different from those of returns measured in the calendar-time scale.

Keywords: Trading volume, ARCH effects, time-scale, stock-return distribution, high frequency data.

Classification J.E.L. : C.32 ; G.14.

La recherche du processus engendrant le cours des actifs a connu de nombreux développements. Si le processus gaussien a été longtemps le plus couramment utilisé, de nombreuses alternatives à une loi normale ont été proposées. Deux principales approches peuvent être distinguées d'après TUCKER (1992) : la première suppose que la loi de distribution des rendements est plus générale qu'une loi normale, la seconde suppose que les paramètres du processus des prix varient dans le temps. Certains auteurs, dans le cadre de la première approche, utilisent des lois stables de Pareto-Levy, des lois elliptiques, des mélanges de lois normales, la loi de Student, la loi Gamma, la loi de Bêta généralisée, des convolutions d'une loi normale et d'une loi de Poisson, des mélanges d'une loi normale et d'une loi d'Edgeworth-Sargent¹. Néanmoins, aucune de ces hypothèses à notre connaissance, ne fournit une description suffisamment bonne de la dynamique des rendements boursiers qui puisse faire l'objet d'un consensus. La seconde approche considère que les paramètres de la loi des rendements, et plus précisément son deuxième moment, varient dans le temps. Les exemples les plus célèbres de cette approche sont les processus de la famille des GARCH introduits par ENGLE (1982) et BOLLERSLEV (1986). Dans le cadre de cette approche, on peut supposer qu'il existe un « temps boursier » [Cf. CLARK (1973), DACOROGNA et *alii* (1992) et (1993), ZHOU (1996), MANDELBROT (1997)] qui soit différent du temps d'horloge. Si les paramètres de la loi varient dans le temps d'horloge, ils seraient, par définition, constants dans ce temps boursier. Il est, en effet, possible de définir trois échelles de temps : le temps-calendaire, le temps-transaction et le temps-volume. Dans le temps-calendaire, une période fixe de temps d'horloge sépare deux transactions ; le temps-transaction est caractérisé par un nombre fixe de transactions entre chaque observation ; l'unité de temps-volume est un nombre constant de titres échangés entre deux observations. Ainsi, une série exhaustive d'enregistrements de transactions est directement exprimée en temps-transaction. A partir de celle-ci, on peut construire deux séries, l'une en temps-calendaire et l'autre en temps-volume, selon que l'on agrège les ordres par le temps d'horloge ou par le volume de transaction. Le principal but de cet article est de montrer que la modélisation du processus qui gouverne les prix est améliorée lorsque ceux-ci sont exprimés dans l'échelle de temps « déformé » qui correspond à l'agrégation des volumes.

Le rôle des volumes de transaction, dans le processus de formation des cours, a en effet été mis en évidence dans la littérature à maintes reprises sur le marché des actions : les larges variations de prix sont associées à de forts volumes de transaction². Plusieurs études empiriques concluent, de plus, à l'existence d'une causalité entre les volumes de titres échangés et les variations de cours³. A la suite des travaux de BESSEMBINDER et SEGUIN (1993) utilisant des modèles à volatilité stochastique, AVOUYI-DOVI et *alii* (1996), ainsi

1. Cf., par exemple, BOUTILLIER (1992) et MAULEÓN et PEROTE (1997).

2. E.g. KARPOV (1987), GALLANT et *alii* (1992) et CHAUVEAU (1997) pour des revues de la littérature.

3. Cf. AVOUYI-DOVI et *alii* (1996) et CHAUVEAU (1997).

que CHAUVEAU (1997) montrent, en estimant simultanément les rendements boursiers et leur volatilité, que le volume a un effet significatif sur ces deux variables. LAMOUREUX et LASTRAPES (1990-a) montrent, dans le cadre d'un modèle de la famille des GARCH, sur des rendements journaliers de titres américains, que l'autocorrélation de la volatilité diminue lorsque le volume est pris en compte dans l'équation de la variance conditionnelle. Il est donc possible, de manière plus générale, que la persistance de la volatilité constatée sur les cours soit due à une erreur de spécification [LAMOUREUX et LASTRAPES (1990-a), GOURIÉROUX *et alii* (1996-a)], résultant de l'omission d'une variable d'information, par exemple le volume de transaction. Cette interaction entre les volumes de transaction et les cours pourrait être interprétée comme une conséquence de la corrélation entre le volume échangé et le flux d'information arrivant sur le marché⁴.

L'irrégularité de l'arrivée des informations serait à l'origine de certaines caractéristiques telles que l'hétéroscédasticité des cours boursiers et la leptokurticité des rendements. Le volume pourrait, comme le suggère par exemple HARRIS (1986), être considéré comme une variable instrumentale conditionnant le processus des rendements. Il permettrait ainsi d'expliquer, en partie, les brusques variations des cours.

Après avoir décrit notre base de données *intra-journalières*, nous montrons dans une première partie que la série des volumes associés à chaque transaction vérifie la relation dite de Zipf [Cf. MANDELBROT (1968)], qui lie la fréquence d'occurrence d'un volume de transaction et sa taille relative. Nous nous intéressons alors plus précisément à la distribution de ces volumes de transaction. Dans le cadre analytique des modèles ARCH, la seconde partie a pour objet d'étudier l'influence du volume sur la volatilité du rendement, mesuré cette fois en temps-calendaire, suivant en cela les travaux de LAMOUREUX et LASTRAPES (1990-a). La troisième partie concerne les conséquences de l'agrégation, par les volumes et par le temps, sur les caractéristiques d'hétéroscédasticité et de non-normalité des rendements boursiers. Nous illustrons enfin, dans la quatrième partie, ces différences à partir de deux mesures de la même série dans ces deux échelles de temps concurrentes.

1. Distribution des rendements, information et volume

Nous commençons cette partie par la description de notre échantillon de données de haute fréquence (section 1.1). Nous rappelons ensuite les faits stylisés relatifs aux rendements de haute fréquence, tels qu'ils sont rapportés dans la littérature empirique (section 1.2). Nous étudions, enfin, les caractéristiques de la série des volumes de transaction associés à chaque ordre renseigné dans la base (section 1.3).

4. Cf. CLARK (1973), HARRIS (1986).

1.1 Description des données

Notre étude porte sur l'action Elf-Aquitaine, choisie pour sa forte capitalisation, et sur laquelle les échanges, à Paris, de mars à août 1996 représentent plus de 150 000 transactions, soit plus de 64 millions de titres échangés, pour une valeur totale des échanges traités de plus de 21 milliards de francs. Les six mois d'observation choisis ne comprennent pas de période de crise majeure sur les marchés d'actions, ni d'activité particulière sur ce titre. Les données *intra-journalières* sont extraites de la base de la Société de Bourses Françaises, dans laquelle, pour chaque transaction ayant lieu sur le marché, sont enregistrés, entre autres, son horodatage, le prix auquel se fait l'échange et le volume traité. Afin d'éviter l'effet de «pré-ouverture» caractéristique du marché parisien, ainsi que l'effet de fin de journée, nous avons éliminé de la base les rendements des première et dernière demi-heures de cotation chaque jour⁵. Ainsi, nous avons en particulier éliminé les rendements «*overnight*» qui correspondent aux différences entre les cours de clôture et d'ouverture de la bourse. Nous avons de plus vérifié que l'élimination de ces rendements ne modifiait pas qualitativement les résultats généraux de cet article.

1.2 Les faits stylisés

Il est connu que le paradigme Markov-Gaussien, c'est-à-dire l'approximation des cours boursiers par une marche aléatoire à incréments normaux, n'est pas une bonne représentation des distributions empiriques des actifs financiers⁶. Les faits stylisés tels qu'ils sont décrits dans la littérature⁷ sont les suivantes :

- les distributions empiriques des rendements boursiers sont leptokurtiques⁸, «faiblement» asymétriques, avec un pic très prononcé pour la valeur 0 ;
- les autocorrélations des carrés de rendements boursiers (ou de leur valeur absolue) sont significativement positives, phénomène connu sous le nom de «mémoire longue» des séries⁹ ;
- les rendements boursiers ne sont pas homoscedastiques ; leurs volatilités sont autorégressives¹⁰.

Les premières constatations empiriques de cette inadéquation incitent MANDELROT (1963, 1973-a et 1973-b) à définir l'exposant de Hurst des cours boursiers, ou exposant H , défini comme le taux de croissance de l'écart-type de ses

5. Les cours correspondant à des «applications» ont été éliminés de la base ; soient 1238 observations sur les 179 958 observations de Mars à Août 1996 ; Cf. la documentation de la base SBF-bdm (Janvier 1996), p.26. L'élimination des rendements relatifs à la première et à la dernière demi-heure de cotation de chaque jour ouvré conduit à conserver, au final, 118 876 observations.

6. Cf. WALTER (1996) ; BOUCHAUD et POTTERS (1997).

7. Cf. PAGAN (1996) pour une revue de la littérature.

8. Une étude détaillée de cette caractéristique est fournie dans JANSEN et DE VRIES (1991) et LONGIN (1996).

9. Cf. par exemple DING *et alii* (1993).

10. E.g. ENGLE (1982), BOLLERSLEV (1986).

incréments par rapport au pas sur lequel sont calculés ces mêmes incréments. Dans le cas d'une marche aléatoire «pure», la valeur de cet exposant est $1/2$; elle est supérieure à $1/2$ si la série est persistante (à «mémoire longue») et inférieure à $1/2$ dans le cas où la série est «anti-persistante». Si la mesure proposée par Mandelbrot permet de prendre en compte les propriétés des rendements, aucune justification théorique de celles-ci n'a été apportée avant les travaux de CLARK (1973). Cet auteur explique la distribution particulière des rendements par l'hypothèse de «mélange de distributions»; la distribution totale étant ainsi la composée de distributions de variances différentes. Dans le modèle de Clark, le processus qui gouverne le rendement de l'actif est lui-même fonction d'un processus stochastique sous-jacent qui représente l'évolution aléatoire de l'arrivée d'information sur le marché. Cette arrivée d'information définit le «temps opérationnel», mesuré à chaque nouvelle information, par opposition au temps-calendaire mesuré à des intervalles de temps réguliers. Dans cette optique, le volume de transaction a souvent été utilisé comme facteur explicatif du rendement ou de sa variance, dans la mesure où il est lui-même fonction de l'arrivée d'information. L'hypothèse de «mélange de distributions» est ainsi résumée par HARRIS (1986), en introduisant le volume :

$$\begin{cases} (1) : & \Delta P_t \sim N(an_t, bn_t | n_t) \\ (2) : & Vol_t \sim N(cn_t, dn_t | n_t) \\ (3) : & Cov(\Delta P_t, Vol_t | n_t) = 0 \end{cases}$$

où ΔP_t est la variation du prix entre $t - 1$ et t , $N(\cdot)$ est la densité de la loi normale, Vol_t est le volume de transaction, et, a , b , c et d sont des paramètres positifs. Le processus n_t gouvernant l'évolution des volumes et des prix, peut être aléatoire, déterministe ou saisonnier. Cette spécification permet à Harris de déduire des propriétés théoriques sur les volumes et les rendements, qui s'avèrent correspondre aux faits stylisés rappelés précédemment.

L'hypothèse commune à de tels modèles est donc que les volumes de transaction, ou à défaut, le nombre de transactions, sont une bonne approximation de la «quantité d'information» contenue dans les prix.

En accord avec cette hypothèse, nous supposons ici que la non-normalité de la on des rendements en temps-transaction est liée à la manière dont l'information est véhiculée par les échanges. Nous étudions donc tout d'abord les volumes en nous basant sur une méthodologie empruntée à Zipf et décrite par GIRAUD (1968) et MANDELBROT (1997).

1.3 Description des volumes de transaction associés à chaque ordre

L'équation de Zipf spécifie une relation entre la fréquence de la modalité d'un événement et son rang dans le classement cardinal des modalités, telle que :

$$i(v) = \frac{k}{N(v)}$$

où k est une constante, $N(v)$ est le nombre d'ordres de volume v et $i(v)$ le rang correspondant au volume v dans la série de volumes ordonnée dans un ordre croissant.

Nous avons voulu vérifier ici que cette relation s'appliquait aux volumes de transaction. Nous présentons ci-après l'estimation de l'équation de Zipf pour la série des volumes associés à chaque transaction enregistrée pour le titre Elf-Aquitaine :

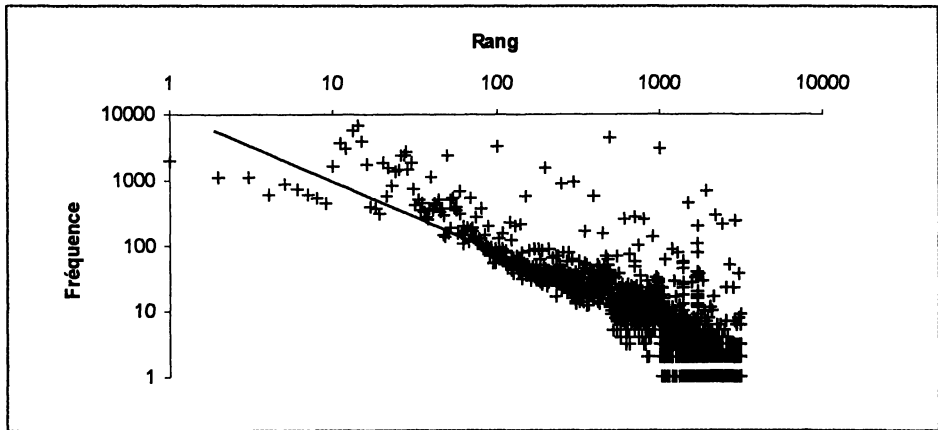
$$\ln(N(v)) = 11.21 - 1.38 \ln(i(v))$$

Le graphique 1, en coordonnées logarithmiques, illustre la loi de Zipf sur notre échantillon de volumes à partir d'une estimation par les moindres carrés ordinaires.

Cette relation rang-fréquence est vérifiée pour de nombreuses séries, plus exactement «partout où l'on définit les catégories observées comme la somme d'un certain nombre d'unités de base»¹¹. Dans cette interprétation, les volumes les plus importants résultent en fait de l'agrégation de transactions élémentaires.

Nous présentons également les principales statistiques descriptives de cette série de volume dans le tableau 1, ainsi que sa densité de probabilité estimée dans le graphique 2.

Graphique 1 : Loi de Zipf appliquée aux volumes de transaction



où la fréquence représente le nombre d'ordres de volume de i -ème rang.

11. GIRAUD (1968), p.157.

Tableau 1 : Statistiques descriptives des volumes de transaction

Nombre d'observations	118 876
Moyenne*	306.50
Ecart-type*	785.10
Coefficient d'asymétrie	13.58
Indice de <i>Kurtosis</i>	483.89
Minimum*	1
Maximum*	51 000
Médiane*	41
Ordre d'Intégration	I(0)**
Tests de normalité	rejet**

* l'unité des statistiques descriptives est le titre échangé.

** Les tests de PHILIPS et PERRON (1988) et de DICKEY et FULLER augmenté (1981) rejettent l'hypothèse de racine unitaire au seuil de 1 %, bien que le test de KWIATOWSKI, PHILIPS, SHIN et SCHMIDT (1992) n'accepte la stationnarité qu'au seuil de 1 %. L'autocorrélation, étudiée *infra*, est pourtant relativement faible.

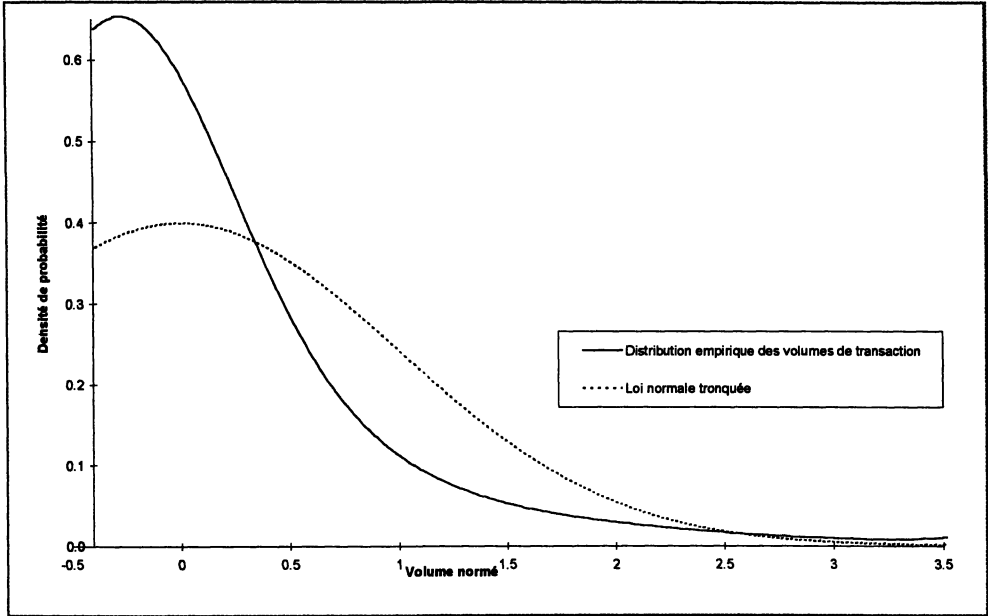
*** Les tests d'ajustement du Chi-deux, de Jarque-Bera, de Kuiper et de Kolmogorov-Smirnov conduisent au rejet de l'hypothèse de normalité au seuil de significativité de 5 %. Nous utilisons, ici, plusieurs tests de normalité étant donné l'importance de cette propriété dans notre article. Cf. TASSI (1992) pour une description des tests ; voir aussi RONCALLI (1995).

Cette série de volumes de transaction du titre Elf-Aquitaine, l'un des titres les plus liquides du marché parisien, est caractérisée par une très grande disparité entre les transactions les plus faibles et les quelques transactions de fort volume. On constate, en effet, une forte asymétrie de la distribution : les transactions de moins de 41 titres représentent environ 50 % des transactions totales, et la transaction la plus forte est supérieure à 166 fois la transaction moyenne. Le graphique 2 représente la densité de probabilité des volumes de transaction, comparée à une loi normale tronquée.¹² Bien que les événements les plus extrêmes ne soient pas représentés ici, il apparaît nettement sur ce graphique que la loi normale n'est pas une bonne approximation de la distribution empirique.

12. Pour représenter cette densité, nous avons utilisé un estimateur non-paramétrique à noyau gaussien.

Cf. DEHEUVELS (1977), SILVERMAN (1986), et GOURIÉROUX (1996-a).

Graphique 2 : Distribution des volumes de transaction



La distribution empirique correspond, de plus, aux propriétés énoncées précédemment : il existe un très grand nombre de transactions d'un faible montant et un nombre très faible de très grands volumes. Si l'on approxime l'information par le volume, la loi de Zipf permettrait alors de décrire la distribution de l'information arrivant sur le marché. Cela revient à faire l'hypothèse qu'aux transactions élémentaires correspondent des informations unitaires. Les variations de cours induites par ces transactions n'auraient alors qu'un faible contenu informationnel. En transformant la série brute des rendements afin de rendre le volume uniforme, la même valeur informationnelle serait accordée à toutes les variations des cours ainsi transformés.

Avant d'effectuer cette transformation, nous avons voulu vérifier le lien, clairement établi dans la littérature [Cf. KARPOV (1987), CHAUVEAU (1997)], qui existe entre volume de transaction et prix. Plus précisément, nous nous proposons de tester la relation entre le risque, mesuré par la volatilité du rendement, et le flot d'information arrivant sur le marché, mesuré par le volume des échanges, en reprenant les travaux de LAMOUREUX et LASTRAPES (1990-a) et en appliquant leur méthode à notre série de haute fréquence.

2. Hétéroscédasticité des rendements de haute fréquence mesurés en temps-calendaire et influence des volumes de transaction

Une des modélisations possibles de l'hétéroscédasticité des rendements boursiers est fournie par les processus ARCH (*AutoRegressive Conditional Heteroskedasticity*) tels qu'ils ont été définis par ENGLE (1982) et BOLLERSLEV (1986). Ces modèles ont la particularité principale de spécifier une variance conditionnelle autorégressive. Depuis ces articles fondateurs, la modélisation a été progressivement enrichie pour tenir compte de certaines caractéristiques des séries financières comme l'asymétrie des effets, le degré d'intégration des séries, l'existence de seuils, ou encore la non-normalité des résidus normés par la variance conditionnelle [Cf. BERA et HIGGINS (1993), HAMILTON (1994), BOLLERSLEV *et alii* (1994), SHEPHARD (1996)]. L'hétéroscédasticité des rendements boursiers peut aussi être décrite par des modèles dits de volatilité stochastique, dans lesquels la variance conditionnelle peut dépendre, non seulement de ses valeurs passées, mais aussi de composantes non-observées ou de variables exogènes explicites [Cf. HAMILTON (1994), TAYLOR (1994), PAGAN (1996), SHEPHARD (1996)].

TAYLOR (1986), propose, par exemple, le modèle suivant :

$$\begin{cases} \tilde{r} &= \tilde{\varepsilon}_t \beta \exp(\tilde{h}_t/2) \\ \tilde{h} &= \delta_0 + \delta_1 \tilde{h}_{t-1} \end{cases}$$

où $\tilde{\varepsilon} | \Phi_{t-1} \sim N(0, \tilde{h}_t)$, avec $\Phi_{t-1} = \{\tilde{\varepsilon}_{t-1}, \tilde{\varepsilon}_{t-2}, \dots\}$ et $t = [1, 2, \dots, T]$.

Une interprétation courante de la variable latente \tilde{h}_t est qu'elle représente le flot d'information arrivant sur le marché. L'apport de LAMOUREUX et LASTRAPES (1990-a) est d'introduire explicitement les volumes de transaction dans l'information conditionnant la variance, dans le cadre d'un modèle GARCH(P,Q). Le modèle proposé peut alors s'écrire :

$$\begin{cases} \tilde{r}_t &= \mu + \tilde{\varepsilon}_t \\ \tilde{h}_t &= \alpha_0 + \sum_{p=1}^P \beta_p \tilde{h}_{t-p} + \sum_{q=1}^Q \alpha_q \tilde{\varepsilon}_{t-q} + \sum_{v=0}^V \delta_v \widetilde{Vol}_{t-v} \end{cases}$$

où $\tilde{\varepsilon} | \Phi_{t-1}' \sim N(0, \tilde{h}_t)$, avec $\Phi_{t-1}' = \{\widetilde{Vol}_t, \tilde{\varepsilon}_{t-1}, \tilde{\varepsilon}_{t-2}, \dots\}$, $t = [1, 2, \dots, T]$, et \widetilde{Vol}_t représente le volume échangé en t .

Lamoureux et Lastrapes comparent un modèle GARCH(1,1), estimé par la méthode du maximum de vraisemblance, dont les paramètres δ_v sont contraints à 0, avec un modèle où le coefficient associé au volume instantané n'est pas contraint. Leurs estimations sur données quotidiennes les conduisent à constater une diminution significative de la somme des coefficients ($\alpha_1 + \beta_1$),

lorsque le volume instantané intervient dans l'équation de la variance conditionnelle. Leur conclusion est alors que l'hétéroscédasticité des rendements peut s'expliquer en partie par les fluctuations des volumes échangés.

Nous avons appliqué, sur nos données de haute fréquence, le modèle décrit précédemment. Les rendements et volumes ont été mesurés, en temps-calendaire, avec une périodicité de 5 minutes, soit au total 7550 observations. A l'instar de WOOD *et alii* (1985) et FEINSTONE (1987) sur données *intra*-journalières américaines et GOURIÉROUX *et alii* (1996-a, 1996-b) sur l'action Alcatel sur le marché français, nous prenons en compte les saisonnalités *intra*-journalières. Les volumes de transaction ainsi que les rendements sont donc corrigés selon l'heure à laquelle la transaction a eu lieu. Nous avons, à l'instar de ANDERSEN et BOLLERSLEV (1997), désaisonnalisé les séries en normalisant les rendements et les volumes de transaction par leurs valeurs absolues moyennes pour chaque date de transaction à l'intérieur de la journée. Les autocorrélogrammes des volumes et des rendements avant désaisonnalisation sont donnés en annexe. Le tableau 2 contient les valeurs des coefficients ($\mu; \alpha_0; \alpha_1; \beta_1; \delta_0$) estimés par la méthode standard du *pseudo-maximum* de vraisemblance.

Tableau 2 :
Estimation de l'équation de volatilité d'un modèle GARCH(1,1)

Coefficients (Statistiques de Student)	Modèle sans volume	Modèle avec volume
μ (t-statistique)	0.3997 (698.43)	0.4005 (834.40)
α_0 (t-statistique)	0.0004 (17.46)	0.0004 (15.99)
β_1 (t-statistique)	0.7509 (69.59)	0.2589 (15.83)
α_1 (t-statistique)	0.1492 (21.71)	0.1996 (18.53)
δ_0 (t-statistique)	0 (0)	0.0057 (34.51)
Log-Vraisemblance (P-statistique du ratio de Quandt : 0.0000)	$\ln L(\hat{\theta}) = 11286.28$	$\ln L(\hat{\theta}) = 11860.92$
Autocorrélation d'ordre 1 des résidus normés : test de Durbin-Watson	$DW(\tilde{\varepsilon}_t/\tilde{h}_t) = 2.23$	$DW(\tilde{\varepsilon}_t/\tilde{h}_t) = 2.32$
Autocorrélation des résidus normés : test de Ljung-Box pour 4 retards	$Q_4(\tilde{\varepsilon}_t/\tilde{h}_t) = 104.37^*$	$Q_4(\tilde{\varepsilon}_t/\tilde{h}_t) = 207.41^*$
Indice de Kurtosis $k(f(\tilde{\varepsilon}_t)) = 6.80$	$k(f(\tilde{\varepsilon}_t/\tilde{h}_t)) = 6.98$	$k(f(\tilde{\varepsilon}_t/\tilde{h}_t)) = 2.91$
Normalité des résidus : test de Jarque-Béra	$JB(\tilde{\varepsilon}_t/\tilde{h}_t) = 15512.48^*$	$JB(\tilde{\varepsilon}_t/\tilde{h}_t) = 2729.44^*$

* Statistiques significatives au seuil de 5 %.

Les statistiques de Student en gras sont significatives au seuil de 5 %.

La somme des coefficients $\alpha_1 + \beta_1$ du modèle GARCH(1,1) sans volume est à peu près égale à 0.90, ce qui est conforme par exemple aux résultats de LAMOUREUX et LASTRAPES (1990-a). Dans le cadre de ce modèle, cela impliquerait donc une forte inertie de la variance¹³, ce qui est jugé peu crédible par plusieurs auteurs tels que LAMOUREUX et LASTRAPES (1990-b) ou HAMILTON et SUSMEL (1994). Ces résultats sont, en effet, incompatibles avec le faible pouvoir prédictif des modèles GARCH [Cf. HAMILTON et SUSMEL, (1994), p.314]. Plusieurs hypothèses ont été retenues pour expliquer cette contradiction : l'existence de changements de régime [Cf. LAMOUREUX et LASTRAPES (1990-b) ; HAMILTON et SUSMEL (1994)], la présence de composantes déterministes *intra*-journalières [GOURIÉROUX (1996-a) ; ANDERSEN et BOLLERSLEV (1997)], ou, plus simplement, l'omission d'un facteur explicatif comme le supposent LAMOUREUX et LASTRAPES (1990-a).

C'est ce dernier résultat qui a conduit ces auteurs à introduire le volume de transaction dans l'équation de la variance conditionnelle comme précisé *supra*. L'estimation de cette spécification sur nos données montre que le coefficient affectant le volume de transaction instantané est significativement positif ; le test du rapport des vraisemblances permet de rejeter la spécification de la variance conditionnelle sans volume.¹⁴ L'hétéroscédasticité de la série est atténuée par la prise en compte de l'influence du volume sur la volatilité. Cette atténuation constatée sur la série de données *intra*-journalières est similaire à la réduction de l'hétéroscédasticité mise en évidence par LAMOUREUX et LASTRAPES (1990-a) sur données quotidiennes. Lamoureux et Lastrapes supposent que cette réduction de l'autocorrélation de la variance conditionnelle, lorsque les volumes sont pris en compte, est due à une forte autocorrélation des volumes. Nous avons procédé à l'estimation d'un AR(1) sur les volumes de transaction ; les résultats sont les suivants :

$$\widetilde{Vol}_t = 270,95 + 0,116 \widetilde{Vol}_{t-1} + \tilde{\varepsilon}_t$$

(78,91*) (12,00*)

$$DW(\tilde{\varepsilon}_t) = 2.02 \quad \bar{R}^2 = 0.013$$

* Statistiques de Student significatives au seuil de 5 %.

L'autocorrélation des volumes *intra*-journaliers est significative au seuil de 5 % ; il est donc probable que la dynamique du volume de transaction explique, au moins en partie, le caractère autorégressif de la volatilité.¹⁵

13. Le test du Chi-deux sur la restriction ($H_0 : \alpha_1 + \beta_1 = 1$) permet de rejeter l'égalité à 1 de la somme des coefficients du GARCH au seuil de significativité de 5%. Une spécification IGARCH(1, 1, 1) est donc ici à rejeter.

14. Le test de rapport des vraisemblances conduit à rejeter la restriction de nullité du coefficient du volume dans l'équation de variance conditionnelle.

15. Les statistiques de Student sont corrigées de l'hétéroscédasticité des séries (NEWBY et WEST, 1987) ; Cf. le descriptif de l'option *Robusterror* dans la documentation de RATS. L'autocorrélogramme des volumes est représenté en annexe.

Nos conclusions sont, néanmoins, différentes de celles de LAMOUREUX et LAS-TRAPES (1994) et de ACOSTA GONZALEZ et PEREZ RODRIGUEZ (1997) sur données journalières ; pour ces derniers auteurs, si le volume de transaction est une variable explicative significative dans l'équation de la volatilité conditionnelle des rendements journaliers, l'estimation de celle-ci n'est pas significativement améliorée par la prise en compte du volume de transaction instantané dans le cadre d'un modèle GARCH (1,1).

Ces résultats montrant l'importance de conditionner la mesure du rendement par le volume de transaction, nous avons cherché à les compléter par l'étude de la distribution des rendements conditionnels.

3. Distribution des rendements bruts et conséquences de l'agrégation par le temps et par les volumes

Dans la section 3.1, nous nous intéressons aux rendements correspondant à chaque variation de prix entre deux transactions renseignées dans la base (*i.e.* mesurés en temps-transaction). Nous étudions ensuite les conséquences de l'agrégation des cours sur les caractéristiques de persistance de la volatilité conditionnelle et de normalité de la distribution des rendements. Dans la section 3.2, nous résumons les résultats théoriques concernant les effets de l'agrégation temporelle. Dans la section 3.3, nous effectuons des tests d'effet ARCH et de normalité sur des séries de période croissante mesurées en temps-calendaire. Nous complétons les tests de normalité par une étude de l'indice de *Kurtosis* de ces séries. Dans la section 3.4, enfin, nous reproduisons la démarche de la section 3.3 (effets de l'agrégation sur l'hétéroscédasticité et la normalité des rendements), cette fois sur des séries mesurées en temps-volume. La comparaison des deux échelles de temps révèle qu'à degré d'agrégation comparable, la normalité des séries est plus souvent vérifiée en temps-volume, et que l'approximation gaussienne classique se vérifie en temps-volume pour un degré d'agrégation moindre qu'en temps-calendaire.

3.1 Distribution des rendements entre deux transactions de la série d'ordres originale

Nous présentons ci-après (tableaux 3 à 5 et graphique 3) les principales caractéristiques de la distribution des rendements mesurés entre deux transactions renseignées dans la base originale (distribution en temps-transaction)¹⁶.

16. Pour représenter cette densité, nous avons utilisé un estimateur non-paramétrique à noyau gaussien. Cf. DEHEUVELS (1977), SILVERMAN (1986) et GOURIÉROUX (1996-a).

Tableau 3 :

Statistiques descriptives des rendements entre deux transactions

Nombre d'observations	118 876
Intervalle de temps moyen entre deux transactions	21.96 secondes
Moyenne (en %)	0.00000
Ecart-type (en %)	0.00079
Coefficient d'asymétrie	0.1139
Indice de <i>Kurtosis</i>	50.0659
Minimum en %	-0.02026
Maximum en %	0.02197
Médiane en %	0
Ordre d'intégration*	$I(0)$

* Les tests de PHILIPS et PERRON (1989), DICKEY-FULLER augmenté (1986) et KWIATOWSKI, PHILIPS, SHIN et SCHMIDT (1992) ont ici une conclusion identique au seuil de 5 %.

Graphique 3 : Distribution des rendements entre deux transactions

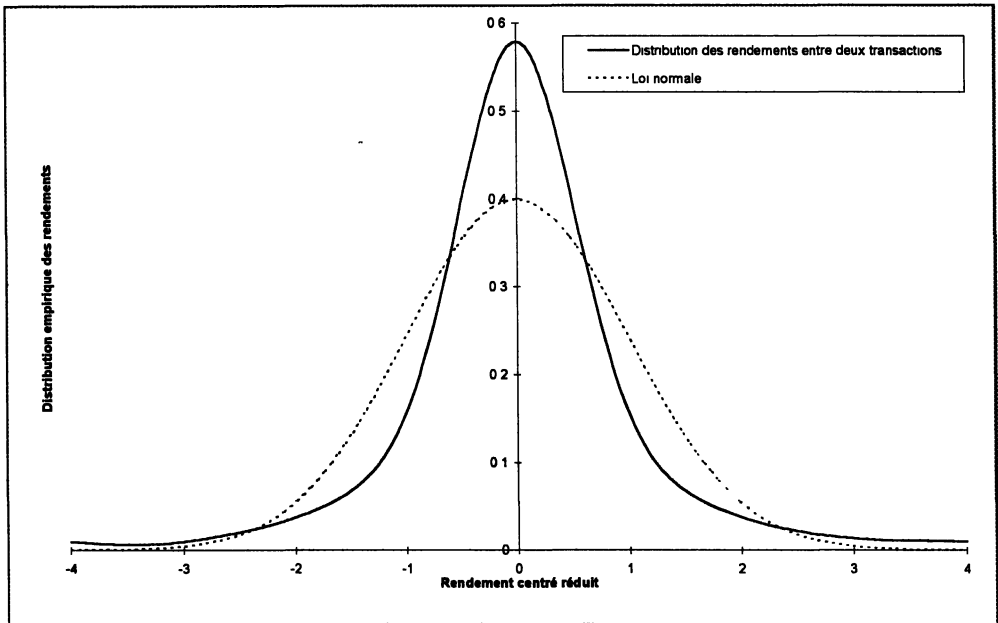


Tableau 4 : Tests paramétriques de la normalité de la distribution des rendements mesurés entre deux transactions¹⁷

$$H_0 : \tilde{r}_{TT_i} \sim N(0, 1)$$

Test	Jarque-Béra	Skewness	Kurtosis
Statistique	10 972 498.4	0.11394	50.06588
P-statistique	0.0000	0.0000	0.0000
Conclusion du test (au seuil de significativité de 5 %)	Rejet de H_0	Rejet de H_0	Rejet de H_0

Tableau 5 : Tests paramétriques de la normalité de la distribution des rendements mesurés entre deux transactions¹⁸

$$H_0 : \begin{cases} \tilde{r}_{TT_i} \sim N(0, 1) \\ \tilde{r}_i \sim N(0, 1) \end{cases}$$

Test	Kolmogorov-Smirnov	Kuiper
Statistique	0.29963	0.59608
Seuil à 5 %	0.00473	0.00473
Conclusion du test (au seuil de significativité de 5 %)	Rejet de H_0	Rejet de H_0

Les résultats des tests de normalité¹⁹ recensés dans les tableaux 4 et 5 montrent le caractère asymétrique à droite et leptokurtique de la distribution des rendements. Les événements «extrêmes» et les variations nulles de prix sont beaucoup plus fréquents que dans le cas d'une loi normale.

17. Où \tilde{r}_{TT_i} est le rendement centré et réduit, calculé entre deux cotations successives (temps-transaction).

18. Où \tilde{r}_{TT_i} est le rendement centré et réduit, calculé entre deux cotations successives (temps-transaction) et \tilde{r}_i est une série de rendements simulés, normalement distribués et de même longueur.

19. Nous utilisons plusieurs tests de normalité, étant donnée l'importance de cette propriété dans notre article.

3.2 Conséquences de l'agrégation temporelle sur la persistance de la variance conditionnelle et la normalité des rendements

La plupart des études réalisées dans la littérature utilisent l'échelle de temps-calendaire sur des périodicités d'observation quotidienne, hebdomadaire et mensuelle. Deux résultats traditionnels apparaissent alors : la disparition de l'hétéroscédasticité et la normalité des séries lorsque l'intervalle de temps séparant deux observations augmente suffisamment [Cf. FELLER (1968), DIEBOLD (1988), DROST et NIJMAN (1993), BERA et HIGGINS (1993)].

Si on considère en effet le modèle original ARCH(1) de ENGLE (1982) où :

$$\begin{aligned} & \tilde{\varepsilon}_t | \Phi_{t-1} \sim N(0, h_t) \\ \text{avec } & h_t = E(\tilde{\varepsilon}_t^2 | \Phi_{t-1}) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \quad t = [1, 2, \dots, T]. \end{aligned}$$

Quand $t = [2, 4, \dots, T]$, c'est-à-dire quand l'ensemble d'information ne contient qu'une donnée sur deux, c'est-à-dire :

$$\Phi_{t-(2)} = \{\tilde{\varepsilon}_2, \tilde{\varepsilon}_4, \dots, \tilde{\varepsilon}_T\},$$

le modèle correspondant pour $\tilde{\varepsilon}$ s'écrit :

$$\begin{cases} E(\tilde{\varepsilon}_t | \Phi_{t-(2)}) = 0 \\ h_{t-(2)} = E(\tilde{\varepsilon}_t^2 | \Phi_{t-(2)}) = \alpha_0(1 + \alpha_1) + \alpha_1^2 \varepsilon_{t-2}^2, \quad t = [2, 4, \dots, T] \end{cases}$$

De façon plus générale, si $t = [m, 2m, \dots, T]$, alors :

$$\begin{cases} E(\tilde{\varepsilon}_T | \Phi_{t-(m)}) = 0 \\ h_{t-(m)} = E(\tilde{\varepsilon}_T^2 | \Phi_{t-(m)}) = \alpha_0 \frac{1 + \alpha_1^m}{1 - \alpha_1} + \alpha_1^m \varepsilon_{t-m}^2, \quad t = [m, 2m, \dots, T] \end{cases}$$

Cette dernière relation exprime le lien entre les paramètres d'un modèle de haute fréquence et ceux d'un modèle de basse fréquence. De plus, quand m devient grand, c'est-à-dire quand l'agrégation devient importante, alors le modèle ARCH(1) le plus approprié pour une haute fréquence d'observation devient un modèle homoscédastique standard à basse fréquence. On a ainsi :

$$\lim_{m \rightarrow +\infty} \left[E(\tilde{\varepsilon}_t^2 | \Phi_{t-(m)}) \right] = \alpha_0 / (1 - \alpha_1).$$

Nous faisons ici l'hypothèse que les modèles GARCH décrivent l'effet de l'arrivée d'information sur le marché et que la disparition du caractère autorégressif de la variance conditionnelle est due à l'agrégation progressive de l'information contenue dans chaque transaction. Si cette hypothèse s'avérait vérifiée, il serait alors plus pertinent de chercher à rendre constant non pas l'intervalle de temps séparant deux observations, mais le montant des volumes des transactions effectuées entre deux observations. Le temps-volume est donc un cas particulier du temps opérationnel défini par CLARK (1973).

3.3 Etude empirique des effets de l'agrégation temporelle sur la persistance de la variance conditionnelle et la normalité normalité des rendements mesurés en temps-calendaire

Nous étudions tout d'abord la disparition progressive de l'hétéroscédasticité, obtenue en faisant croître l'unité de durée séparant deux observations. En renormalisant ainsi la série originale avec un intervalle de temps croissant, nous obtenons 5000 séries de tailles décroissantes. Sur chacune de ces 5000 séries, nous avons mené des tests d'effet ARCH d'ENGLE (1982), en estimant le modèle suivant :

$$\begin{cases} \tilde{r}_t = \mu + \tilde{\varepsilon}_t \\ \hat{\varepsilon}_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i \hat{\varepsilon}_{t-i}^2 + \tilde{\xi}_t \end{cases}$$

où $TR_{\hat{\varepsilon}}^2 \sim \chi^2(m)$ si $H_0 : \tilde{\varepsilon} \sim N(0, \sigma^2)$.

Nous avons considéré ensuite la propriété de normalité des séries exprimées en temps-calendaire. En utilisant les mêmes 5000 séries que précédemment, nous avons mené pour chaque série de rendement normée, les tests de Jarque-Bera, d'ajustement du Chi-deux, de Kolmogorov-Smirnov et de Kuiper. Nous n'avons conclu à la normalité des séries testées que lorsque les quatre tests ne rejettent pas l'hypothèse nulle de normalité au seuil de significativité de 5 concluent à l'hypothèse de normalité et elle vaut 0 dans le cas contraire.

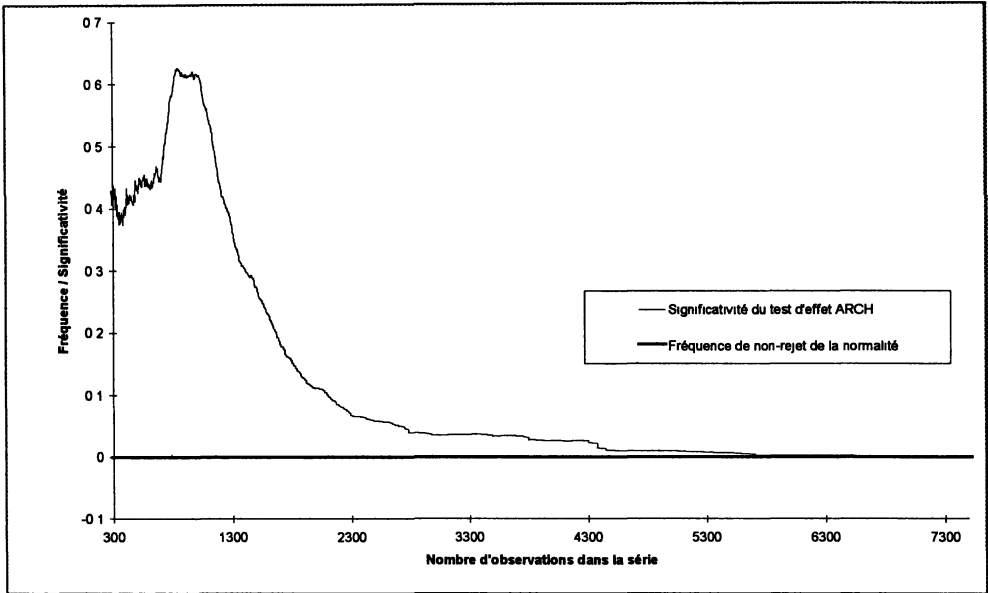
Pour toutes les séries contenant plus de 20730 observations, la probabilité moyenne de rejeter les effets ARCH, telle que nous l'avons définie précédemment, est inférieure à 5 %. De plus, aucune des séries reconstituées n'est normale.

Le graphique 4 résume simultanément les résultats concernant la normalité et l'hétéroscédasticité des séries. Sur l'axe des abscisses est représenté le nombre d'observations de chaque série renormée, ce nombre d'observations étant compris entre 300 pour la plus grande périodicité et 7400 pour la plus petite. Sur l'axe des ordonnées est représentée, en premier lieu, la p-statistique moyenne²⁰ associée au coefficient α_1 dans le test d'effet ARCH(1) pour les séries dont le nombre d'observations est porté sur l'axe des abscisses. Nous avons représenté, en second lieu, le pourcentage des séries (dont le nombre d'observations est inférieur à la valeur lue en abscisse), pour lesquelles les tests concluent à la normalité (*i.e.* dont l'indicatrice de normalité est égale à 1).

20. Obtenue en appliquant un noyau uniforme pour une fenêtre de longueur égale à 200. A la différence de DIEBOLD (1988), nous ne représentons pas la statistique de Chi-deux associée au test, mais le niveau de significativité de celle-ci.

Graphique 4 :

Tests de normalité et d'effets ARCH(1) en temps-calendaire



Ce graphique illustre bien le résultat de la disparition progressive de l'effet ARCH lorsque l'échelle de résolution temporelle augmente, ainsi que la non-normalité des rendements mesurés en temps-calendaire.

Sachant que ce sont les tests relatifs au quatrième moment centré qui sont les plus défavorables à l'hypothèse de normalité des rendements boursiers [Cf. BOUTILLIER (1992)], nous avons calculé les indices de *Kurtosis*²¹ des séries mesurées en temps-calendaire.

Nous illustrons l'effet de l'agrégation sur le graphique 5 suivant où sont reportés, sur l'axe des abscisses, le nombre d'observations dans la série considérée, et sur l'axe des ordonnées, l'indice moyen de *Kurtosis*²² correspondant. On vérifie alors que l'augmentation de la résolution temporelle conduit à élever l'indice de *Kurtosis*²³.

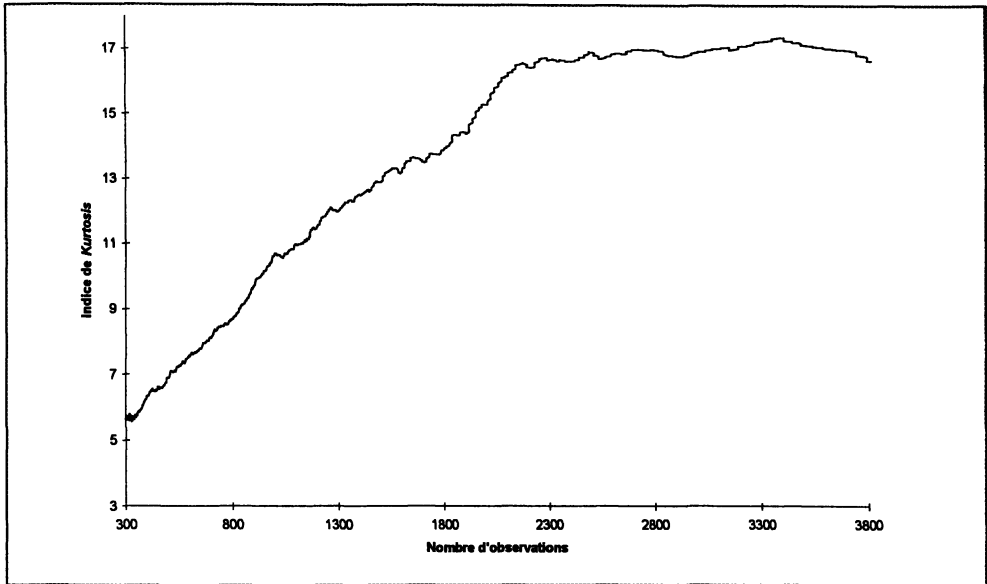
21. Cf. RUPPERT (1987) pour la définition standard de cet indice

22. Obtenu en appliquant un noyau uniforme pour une fenêtre de longueur égale à 200.

23. Nous retrouvons ici le résultat théorique de DIEBOLD (1988), illustré sur la figure 2.1. p.31, dans lequel les abscisses sont croissantes avec le degré d'agrégation.

Graphique 5 :

Indice de Kurtosis et effet de l'agrégation en temps-calendaire



3.4 Etude empirique des effets de l'agrégation temporelle sur la persistance de la variance conditionnelle et la normalité des rendements mesurés en temps-volume

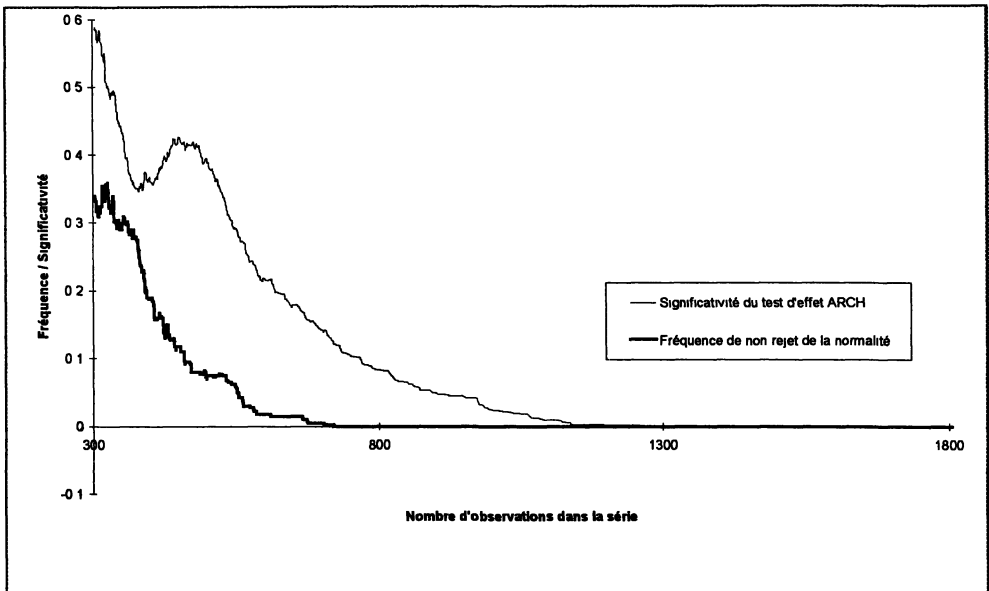
Nous reproduisons ici les mêmes tests que dans la section précédente, mais cette fois sur des séries mesurées en temps-volume. Nous avons ainsi renormalisé la série originale en augmentant le volume de titres échangés entre deux observations, jusqu'à obtenir 5 000 séries de tailles décroissantes. Nous avons ensuite mené sur ces séries le même test d'effet ARCH, ainsi que les tests non-paramétriques de normalité (Chi-deux, Kolmogorov-Smirnov et Kuiper).

Le graphique 6 illustre les résultats concernant la normalité et l'hétéroscédasticité des séries. Sur l'axe des abscisses est représenté le nombre d'observations de chaque série renormée selon le critère de volume de transaction ; ce nombre d'observations est, cette fois, compris entre 300 pour le plus grand seuil et 1 800 pour le plus petit seuil. Comme précédemment, sont représentés sur l'axe des

ordonnées, premièrement, la p-statistique moyenne²⁴ associée au coefficient α_1 dans le test d'effet ARCH(1) pour les séries dont le nombre d'observations est porté sur l'axe des abscisses, et deuxièmement, le pourcentage des séries (dont le nombre d'observations est inférieur à la valeur lue en abscisse), pour lesquelles les tests concluent à la normalité.

Graphique 6 :

Tests de normalité et d'effets ARCH(1) en temps-volume



Dans l'échelle de temps-volume, la probabilité moyenne de rejet de l'effet ARCH est inférieure à 5 % pour les séries comportant plus de 890 observations.

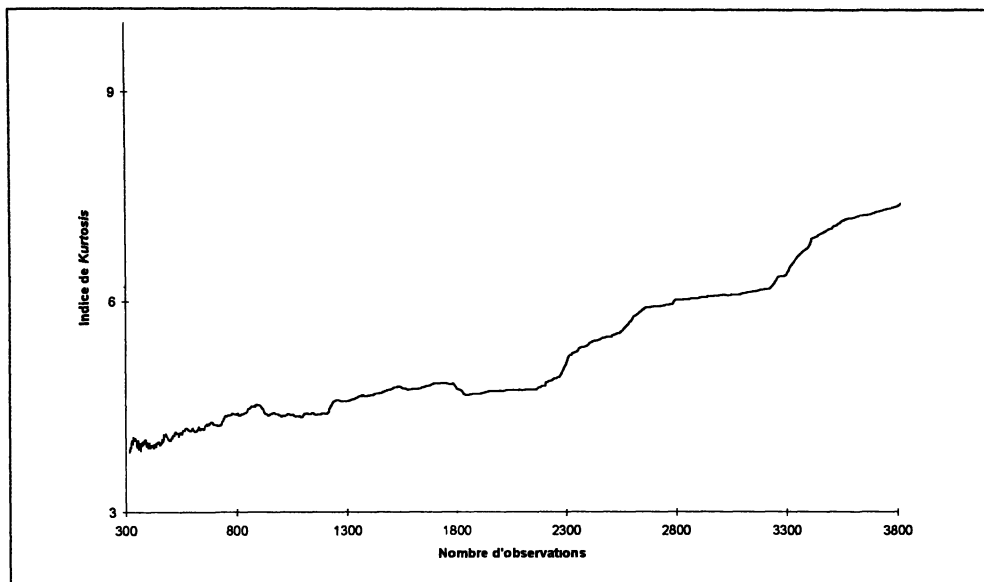
Le graphique 6 illustre donc la disparition progressive de l'effet ARCH lorsque l'échelle de résolution augmente en temps-volume. De plus, 11,03 % des 2 506 séries (contenant entre 300 et 1 881 observations) sont ici considérées comme normales alors qu'aucune des séries comparables ne l'est en temps-calendaire. Plus des deux tiers des séries de moins de 328 observations sont, par ailleurs, normales.

Le graphique 7 suivant représente, comme dans la section précédente, les effets de l'agrégation temporelle sur l'indice de *Kurtosis*.

24. Obtenue en appliquant un noyau uniforme pour une fenêtre de longueur égale à 200.

Graphique 7 :

Indice de *Kurtosis* et effet de l'agrégation en temps-volume



Là encore, l'indice de *Kurtosis* augmente avec la fréquence d'observation.

La comparaison des renormalisations dans les deux échelles de temps, illustrée par les graphiques 4 à 7, amène à conclure à des différences notables entre ces deux échelles, à degré d'agrégation comparable. Pour un même degré d'agrégation, l'indice moyen de *Kurtosis* est toujours inférieur dans le cas de l'agrégation par les volumes à ce qu'il est dans le cas de l'agrégation par le temps. Nous constatons que la renormalisation par les volumes de la série permet d'obtenir la propriété de normalité à un niveau d'agrégation pour lequel la renormalisation par le temps ne le permet pas. L'approximation gaussienne des variations des prix est alors justifiée pour une fréquence d'observation encore élevée. Ce phénomène est illustré plus en détail dans la partie suivante.

4. Comparaison des propriétés des rendements mesurés en temps-volume et en temps-calendaire : une illustration

A titre d'exemple, nous avons choisi d'étudier les propriétés statistiques de deux séries de tailles comparables, chacune étant mesurée dans l'une des deux échelles de temps concurrentes. La première série considérée, exprimée en temps-volume, est obtenue en mesurant les rendements toutes les 71 084 actions échangées, soit 503 observations au total sur six mois, avec un intervalle moyen d'une heure et trente minutes environ séparant deux observations successives. Cette série est comparée, dans la suite de ce texte, avec la série mesurée en temps-calendaire pour laquelle la période moyenne séparant deux observations est la même, c'est-à-dire dont le nombre d'observations au total sur les six mois est identique. Le tableau 6 permet une comparaison directe des périodes séparant deux observations²⁵.

Tableau 6 :

Comparaison de la fréquence d'observation dans les deux échelles

Echelle de temps	Temps calendaire	Temps-volume
Temps moyen* entre deux observations (en secondes)	5187.23658	5266.4072
Ecart-type (en secondes)	77.7197	3481.0178
Minimum (en secondes)	5140	380
Maximum (en secondes)	5850	17960

* L'intervalle de temps séparant deux observations en temps-calendaire est variable, car plutôt que d'effectuer une interpolation, nous avons choisi de prendre en considération la première cotation après que se soit écoulée la période de renormalisation considérée [Cf. FUNDING LA COUR (1997), p.4, note 3]. C'est aussi pourquoi les temps moyens séparant deux observations ne sont pas strictement égaux entre les deux échelles de temps.

La renormalisation en temps-volume implique que l'intervalle de temps d'horloge séparant deux observations est très variable (l'écart-type étant 45,2 fois plus important que celui des rendements du temps-calendaire). Néanmoins, dans cette échelle, l'intervalle le plus grand séparant deux observations n'est que d'environ 5 heures.

Nous détaillons tout d'abord les caractéristiques des deux séries reconstituées dans les échelles de temps concurrentes. Le tableau 7 présente les principales statistiques descriptives de ces séries.

25. Pour une étude détaillée des intervalles de temps séparant deux cours, et donc des « durations » de chaque rendement, voir GOURIÉROUX (1996-a).

Tableau 7 :

Statistiques descriptives de la série de rendements normalisés

Echelle de temps	Temps calendaire	Temps-volume
Moyenne (en %)	0.00001	0.00001
Ecart-type (en %)	0.00473	0.00437
Coefficient d'asymétrie	0.04418	-0.03997
Indice de <i>Kurtosis</i>	6.78618	3.47023
Minimum (en %)	-0.02348	-0.01748
Maximum (en %)	0.02443	0.01346
Médiane (en %)	0	0

Alors que les moyenne et écart-type des rendements ne diffèrent pas significativement, les moments centrés d'ordre supérieur, ainsi que les *extrema*, semblent différents. En particulier, l'indice d'aplatissement de la série mesurée en temps-calendaire est ici encore nettement supérieur à celui d'une loi normale. Pour cette série, les tests effectués ne permettent pas d'ailleurs de conclure à la normalité de la distribution des rendements centrés réduits, et ceci alors que la série mesurée en temps-volume est normale d'après les tests effectués. Le tableau suivant présente les valeurs des tests standards de normalité que nous avons utilisés.

Tableau 8 : Tests de normalité de la série de rendements en temps-volume et en temps-calendaire

$$H_0 : \tilde{r}_t \sim N(\mu, \sigma^2)$$

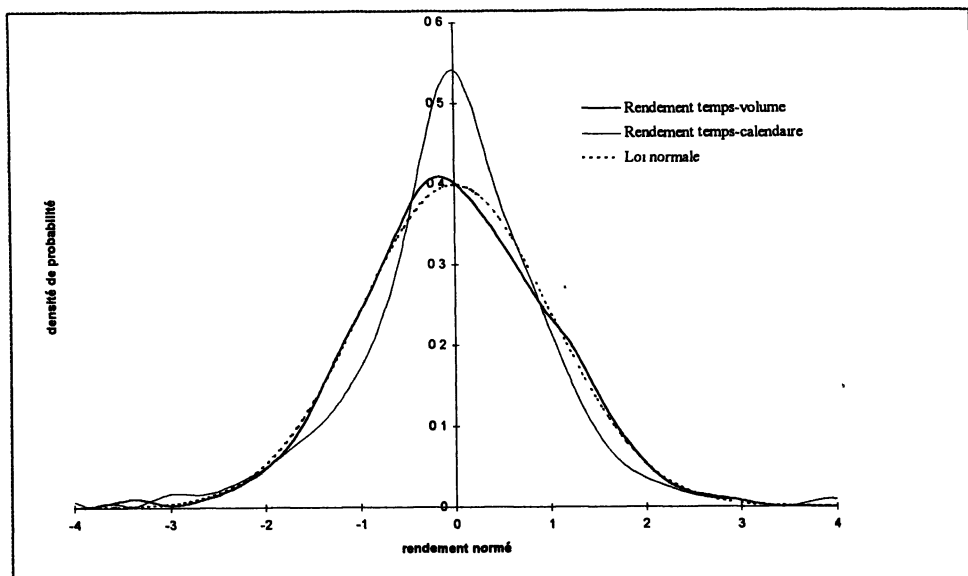
Echelle de temps	Temps calendaire	Temps-volume
Test :	Valeur (p-stat ou seuil) Conclusion au seuil de 5 %	Valeur (p-stat ou seuil) Conclusion au seuil de 5 %
Chi-deux	47.8388 (p-stat = 0.00000) Rejet de H_0	4.9439 (p-stat = 0.8391) H_0
Kolmogorov-Smirnov	0.07607 (seuil = 0.07268) Rejet de H_0	0.02909 (seuil = 0.07268) H_0
Kuiper	0.12862 (seuil = 0.07268) H_0	0.04447 (seuil = 0.07268) H_0

La distribution des rendements en temps-volume est en effet plus proche d'une distribution normale que la distribution des rendements en temps-calendaire. Le rendement moyen en temps-calendaire a, notamment, une probabilité d'occurrence nettement plus forte que dans la loi normale. Les tests de Kolmogorov-Smirnov et Kuiper montrent que les écarts de probabilité entre la distribution normale théorique et la distribution empirique des rendements en temps-calendaire sont significatifs.

Pour illustrer ces tests, nous avons représenté les fonctions de densité de ces deux séries, obtenues en utilisant comme précédemment une estimation non-paramétrique par noyau gaussien. Le graphique 8 présente ces estimations de densité, comparées à une distribution normale.

La normalisation par les volumes tend à éliminer les cotations correspondant à des transactions non significatives en terme de contenu informationnel ; elle permet donc de réduire significativement la probabilité d'occurrence correspondant à un rendement nul. Les grandes variations de prix associées à de faibles volumes de transaction disparaissent également du fait de l'agrégation. D'autre part, les prix étant mesurés à volume constant, donc à information constante, les variations extrêmes, liées à l'arrivée de «*news*», sont moins fréquentes et de moindre ampleur (comme les statistiques reportées dans le tableau 7 le confirment). Néanmoins, les deux premiers moments des distributions des rendements en temps-volume et en temps-calendaire ne sont pas significativement différents.

Graphique 8 : Comparaison des densités de probabilité des rendements en temps-volume et en temps-calendaire



Conclusion

Nous avons tenté dans cet article d'exploiter la notion de temps opérationnel définie par CLARK (1973) pour expliquer les propriétés des rendements des actifs boursiers. Ce temps opérationnel étant normé par l'arrivée d'information sur le marché, nous avons cherché par quelle variable approximer ce flux d'information. En accord avec la littérature, la variable retenue ici est le volume de transaction. Nous avons décrit les propriétés des volumes de transaction, et en particulier à la relation rang-fréquence, dite loi de Zipf, décrite par MANDELBROT (1968). L'étude des volumes suggère alors que ceux-ci permettent non seulement de détecter les arrivées d'information, mais aussi de mesurer «l'intensité d'information» contenue dans chaque transaction. Après avoir rappelé les résultats de LAMOUREUX et LASTRAPES (1990-a) sur le lien entre volume de transaction et volatilité du rendement, nous avons utilisé alors la série de volumes pour reconstituer les cours en temps opérationnel, baptisé ici temps-volume. Nous vérifions que l'agrégation des rendements, par le temps et par les volumes permet de retrouver l'approximation classique qui consiste à considérer que le processus gouvernant l'évolution des prix est gaussien. Néanmoins, les résultats mentionnés précédemment tendent à montrer que c'est l'agrégation par les volumes de transaction qui permet de vérifier l'hypothèse de normalité des rendements. Il existe bien évidemment une relation positive entre nombre de transactions, volume agrégé de transaction et période séparant deux observations consécutives. L'hypothèse que nous formulons est que cette renormalisation nous permet de prendre en considération des prix ayant un contenu informationnel minimum, voire de mesurer les prix à flux d'information constant. Toute l'information pertinente est alors contenue dans les variations de prix en temps-volume, si le niveau d'agrégation est suffisant. L'intérêt d'une telle transformation est alors, comme nous l'avons montré en dernière partie, de reconstituer des séries de rendements de volatilité constante, et pour lesquelles l'approximation gaussienne est justifiée. Pour généraliser ces résultats, il conviendrait, néanmoins, d'étendre l'étude à d'autres actifs financiers, ainsi qu'à des périodes de turbulence sur les marchés financiers. Des études de DACOROGNA *et alii* (1993) et ZHOU (1996) cherchent également à transformer des séries de taux de change de haute fréquence, mesurées en temps-calendaire, en de nouvelles séries exprimées en un temps boursier, tel que la volatilité soit constante. ZHOU (1996) en particulier montre l'intérêt pour la prévision à court terme d'une telle transformation.

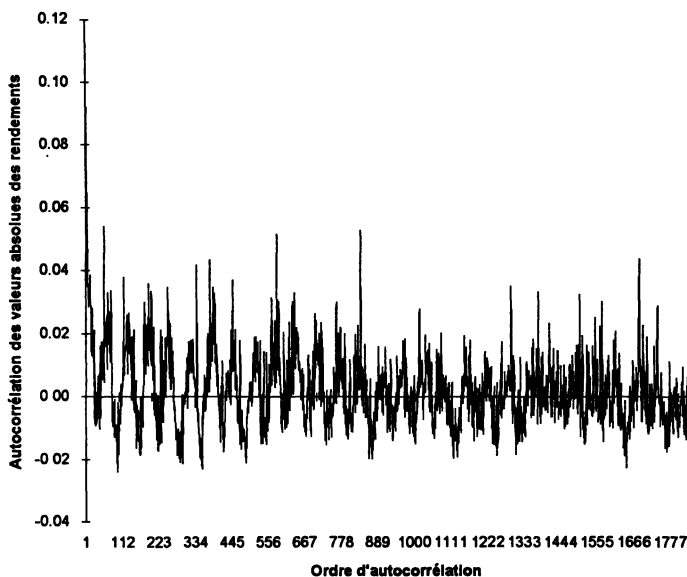
Si ces premiers résultats se confirmaient, la déformation des cours en temps-volume pourrait alors présenter aussi plusieurs applications : de nombreuses procédures économétriques, ainsi que le souligne PAGAN (1996, p.34), reposent sur l'hypothèse de normalité; de plus, la plupart des modèles financiers présupposent l'existence et la stabilité des moments d'ordre deux. Dans le cadre gaussien du temps-volume, on pourrait alors penser à des applications concernant des modèles de prévision de la volatilité, des modèles de *VaR* [Cf. LONGIN (1997)] ou encore des modèles d'options comme alternatives au modèle GARCH d'évaluation des options [Cf. DUAN (1995); VENUTELO et

PÉLLIEUX (1997)]. L'hypothèse de normalité des rendements en temps-volume pourrait aussi être utilisée dans le cadre d'une modélisation *intra-journalière* de la relation du MEDAF [Cf. KIM *et alii* (1996)], pour déterminer la structure optimale de portefeuilles ou comme alternative à l'utilisation de distributions non-gaussiennes [Cf. FRANKFURTER et LAMOUREUX (1987), HARVEY et ZHOU (1993), BELKACEM *et alii* (1995)].

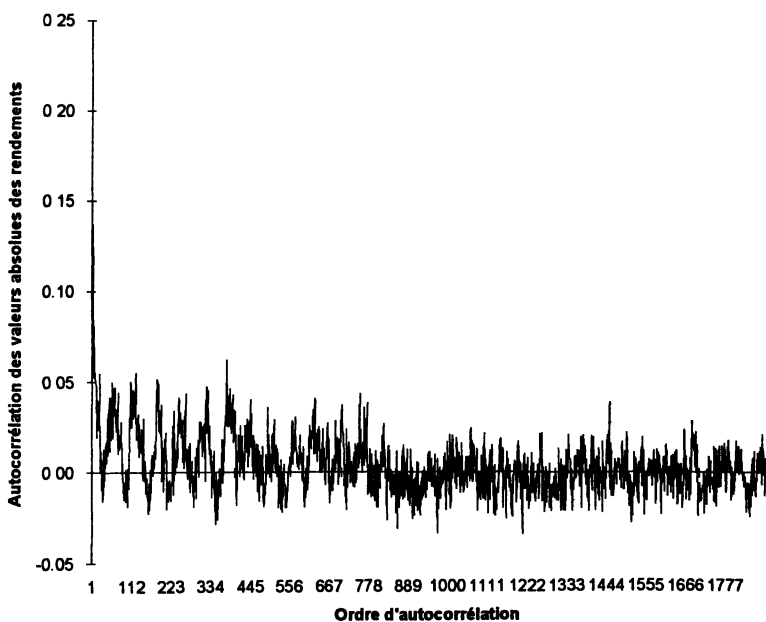
Nous envisageons, par la suite, de prolonger cette étude en nous intéressant plus particulièrement, à l'instar de JANSEN et de VRIES (1991) et de LONGIN (1996), aux rendements extrêmes, afin de mieux caractériser les distributions des rendements dans les deux échelles de temps.

Annexes

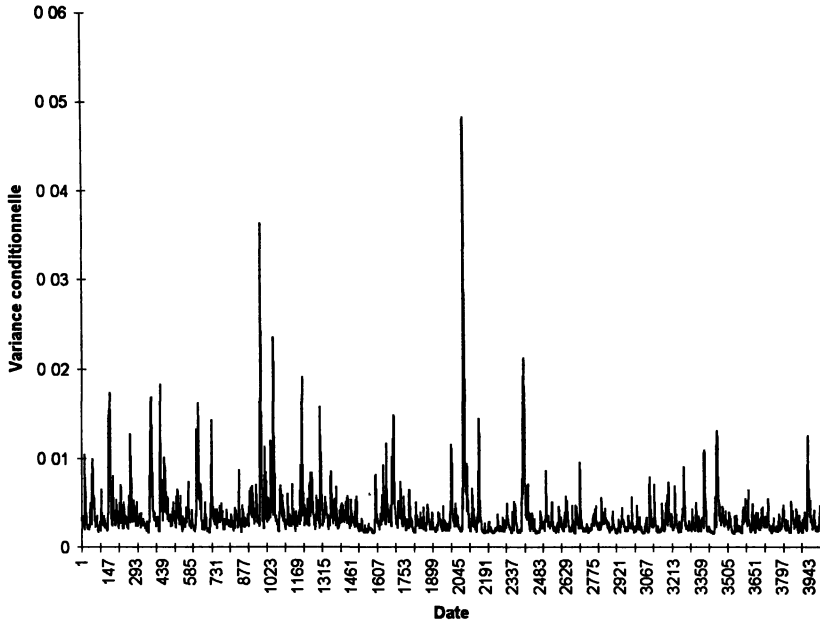
Graphique A.1
Autocorrélation des volumes de transaction bruts



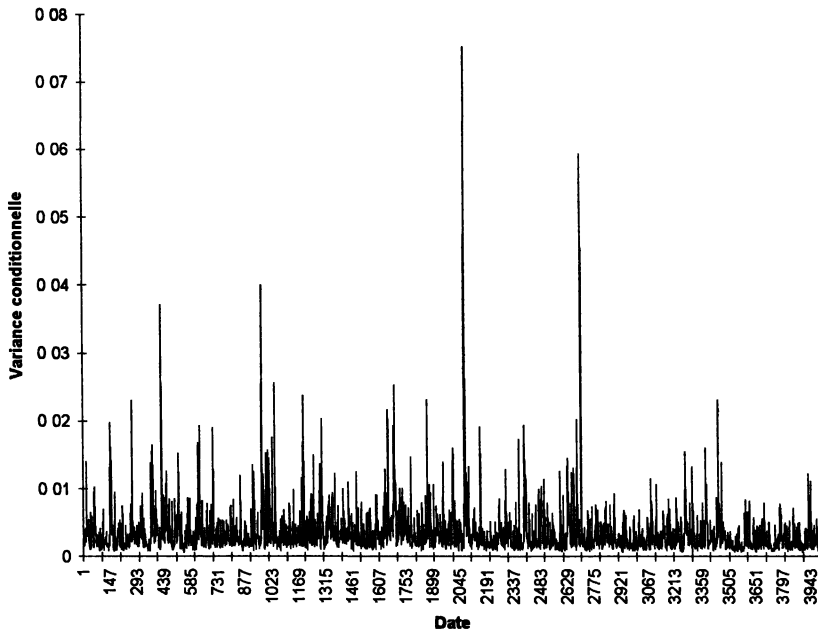
Graphique A.2
Autocorrélation des valeurs absolues des rendements bruts



Graphique A.3
Variance conditionnelle, modèle «sans volume de transaction»



Graphique A.4
Variance conditionnelle, modèle «avec volumes de transaction»



Références

- ACOSTA GONZALEZ E. et PEREZ RODRIGUEZ J. (1997) *Trading Volume and Price Changes of Assets : Causality, Mixture-Distribution and Multivariate GARCH Approach*, Actes du colloque A.E.A.-Imperial College-BNP, Londres, Mai 1997, 21 pages.
- ANDERSEN T. et BOLLERSLEV T. ((1997)) "Heterogeneous Information Arrivals and Return Volatility Dynamics : Uncovering the Long-Run in High Frequency Returns", *Journal of Finance*, vol. 52, n° 3, July 1997, pp. 975-1005.
- AVOUYI-DOVI S., JONDEAU E. et LAI TONG Ch. (1996) *Effet volume, volatilité et transmissions internationales sur les marchés boursiers dans le G5*, Communication au Colloque du GDR *Monnaie et Financement* du CNRS, Aix-en-Provence, Juin 1996, 22 pages.
- BELKACEM L., LEVY VÉHEL J. et WALTER C. (1995) *Generalized Market Equilibrium : 'Stable' CAPM*, Actes du colloque de l'AFFI, Bordeaux, Juin 1995, 21 pages.
- BERA A. et HIGGINS L. (1993) "ARCH Models : Properties, Estimation and Testing", *Journal of Economic Survey*, vol. 7, n° 4, pp. 305-365.
- BESSEMBINDER H. et SEGUIN P. (1993) "Price Volatility, Trading Volume and Market Depth : Evidence from Futures Markets", *Journal of Financial and Quantitative Analysis*, vol. 28, n°1, pp. 21-39.
- BOLLERSLEV T. (1986) "Generalized AutoRegressive Conditional Heteroskedasticity", *Journal of Econometrics*, vol. 69, 1993, pp. 307-327.
- BOLLERSLEV T., ENGLE R. et NELSON D. (1994) "ARCH Models", in *Handbook of Econometrics, volume IV, chapter 49*, pp. 2960-3038, Engle and McFadden Editors, Elsevier Science, 1994.
- BOUCHAUD J.-P. et POTTERS M. (1997) *Théorie des risques financiers*, Collection Aléa-Saclay, Diffusion Eyrolles, 1997, 192 pages.
- BOUTILLIER M. (1992) *Critique des processus de diffusion en finance : le cas des taux de change*, Journal de la Société de Statistique de Paris, tome 133, n° 4, pp. 113-122.
- CAMPBELL J., GROSSMAN S. et WANG J. (1993) "Trading Volume and Serial Correlation on Stock Returns", *The Quarterly Journal of Economics*, n°435, pp. 905-939.
- CHAUVEAU TH. (1997) "Peut-on exploiter le lien statistique entre cours et volumes ? Le cas de quatre bourses de valeurs", *Document de travail de la Caisse des Dépôts*, n°97-04/FI, Juin 1997, 72 pages.
- CLARK P. (1973) "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices", *Econometrica*, vol.41, n°1, pp. 135-155.
- DACOROGNA M., GAUVREAU C., MÜLLER U., OLSEN R. et PICTET O. (1992) *Short Term Forecasting Models of Foreign Exchange Rates*, Actes du colloque de la Caisse des Dépôts et Consignations, Paris, Septembre 1993, 22 pages.
- DACOROGNA M., MÜLLER U., NAGLER R., OLSEN R. et PICTET O. (1993) "A Geographical Model for the Daily and Weekly Seasonal Volatility in the FX Market", *Journal of International Money and Finance*, vol.12, n° 4, 1993, pp. 413-438.
- DEHEUVELS P. (1977) "Estimation non-paramétrique de la densité par histogrammes généralisés", *Revue de Statistique Appliquée*, vol.25, n°3, pp. 5-42.

- DIEBOLD F. (1988) *Empirical Modeling of Exchange Rate Dynamics*, Lecture Notes in Economics and Mathematical System, Springer-Verlag, 143 pages.
- DROST F. et NIJMAN T. (1993) "Temporal Aggregation of GARCH processes", *Econometrica*, vol.61, n°4, 1993, pp. 909-927.
- DUAN J. (1995) "The GARCH Option Model Pricing", *Journal of Mathematical Finance*, vol.5, pp. 13-32.
- ENGLE R. (1982) "AutoRegressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica*, 50, pp. 987-1008.
- FEINSTONE L. (1987) "Minute by Minute : Efficiency, Normality and Randomness in the Intra-daily Asset Prices", *Journal of Applied Econometrics*, vol.2, pp. 193-214.
- FELLER W. (1968) *An Introduction to Probability Theory and Its Applications*, John Wiley Editor, 1968, vol.2, 619 pages.
- FRANKFURTER G. et LAMOUREUX C. (1987) "The Relevance of the Distributional Form of Common Stock Returns to the Construction of Optimal Portfolios", *Journal of the Financial and Quantitative Analysis*, vol.22, n°4, December 1987, pp.505-511.
- FUNDING LA COUR L. (1997) *Regime Shift Models as an Alternative to the Use of a 'Business time Scale' for High-Frequency Exchange Rate Data*, Actes du colloque A.E.A.-Imperial College-BNP, Londres, Mai 1997, 24 pages.
- GALLANT A., ROSSI P. et TAUCHEN G. (1992) "Stock Prices and Volume", *Review of Financial Studies*, vol.5, n°2, 1992, pp.199-242.
- GALLO G. et PACINI B. (1997) *The Effects of Trading Activity on Market Volatility*, Actes du colloque A.E.A.-Imperial College-BNP, Londres, Mai 1997, 25 pages.
- GIRAUD P. (1968) "Language et théorie de la communication", *Encyclopédie de la Pleïade : linguistique*, Gallimard, vol. XXV-6, pp. 145-167.
- GOURIÉROUX CH., JASIAK J. et LE FOL G. (1996-a) "Intra-Day Market Activity", *Documents de travail du CREST, INSEE*, n°9633, Juin 1996, 47 pages.
- GOURIÉROUX CH., LE FOL G. et MEYER B. (1996-b) "Analysis of Order Queues", *Documents de travail du CREST, INSEE*, n°9653, Octobre 1996, 41 pages.
- HAMILTON J. (1994) *Times Series Analysis*, Princeton University Press, 799 pages.
- HAMILTON J. et SUSMEL R. (1994) "AutoRegressive Conditional Heteroskedasticity and Changes in Regime", *Journal of Econometrics*, vol.64, pp. 307-333.
- HARRIS L. (1986) "Cross-security Tests of the Mixture of Distributions Hypothesis", *Journal of Financial and Quantitative Analysis*, vol.21, pp. 39-46.
- HARVEY C. et ZHOU G. (1993) "International Asset Pricing with Alternative Distributional Specifications", *Journal of Empirical Finance*, n°1, 1993, pp. 107-131.
- HSIEH D. (1986) "The Statistical Properties of Daily Foreign Exchange Rates : 1974-1983", *Journal of International Economics*, vol.24, 1986, pp. 129-145.
- JANSEN D. et DE VRIES C. (1991) "On the Frequency of Large Stock Returns : Putting Booms and Busts into Perspective", *Review of Economic and Statistics*, vol.73, 1991, pp.18-24.
- KARPOFF J. (1987) "The Relation between Price Changes and Trading Volume : A Survey", *Journal of Financial and Quantitative Analysis*, n°22, pp. 109-126.

- KIM S., LOCKWOOD L. et MCINISH T. (1996) "Intraday Beta Stability", in *Advances in Quantitative Analysis of Finance and Accounting*, Jay Press Inc. Eds, vol.4, 1996, pp. 35-46.
- LAMOUREUX C. et LASTRAPES W. (1990-a) "Heteroskedasticity in Stock Returns Data : Volume versus GARCH Effects", *Journal of Finance*, vol.45, pp. 221-229.
- LAMOUREUX C. et LASTRAPES W. (1990-b) "Persistence in Variance, Structural Change and the GARCH Model", *Journal of Business and Economic Statistics*, vol.8, pp. 225-234.
- LAMOUREUX C. et LASTRAPES W. (1994) "Endogenous Trading Volume and Momentum in Stock-Return Volatility", *Journal of Business and Economic Statistics*, vol.12, n°2, pp. 225-234.
- LONGIN F. (1996) "The Asymptotic Distribution of Extreme Stock Markets Returns", *Journal of Business*, vol.69, n°3, pp. 383-407.
- LONGIN F. (1997) *Value at Risk : une nouvelle méthode fondée sur la théorie des valeurs extrêmes*, Actes du colloque de l'AFFI, Grenoble, Juin 1997, 32 pages.
- MANDELBROT B. (1963) "The Variation of Certain Speculative Prices", *Journal of Business*, n°36, pp. 394-419.
- MANDELBROT B. (1968) "Les constantes chiffrées du discours", *Encyclopédie de la Pleiade : linguistique*, Gallimard, vol.XXV-3, pp. 46-56.
- MANDELBROT B. (1973) "Formes nouvelles du hasard dans les sciences et Le syndrome de la variance infinie", *Economie Appliquée*, vol.26, pp. 307-348.
- MANDELBROT B. (1997) *Fractales, hasard et finance*, Flammarion, collection Champs, 246 pages.
- MAULEÓN I. et PEROTE J. (1997) *Testing Densities with Financial Data*, Actes du colloque A.E.A.-Imperial College-BNP, Londres, Mai 1997, 24 pages.
- PAGAN T. (1996) "The Econometrics of Financial Markets", *Journal of Empirical Finance*, vol.3, pp. 15-101.
- RONCALLI T. (1995) *Introduction à la programmation sous GAUSS, volume I et II*, Global Design, 299 et 356 pages.
- RUPPERT D. (1987) "What is Kurtosis ? An Influence Approach", *The American Statistician*, vol.41, n°1, February 1987, pp.1-5.
- SBF-BOURSE DE PARIS (1996) *BDM-Bourse de Paris*, Base de données du système CAC sur CD-ROM, Mars-Août 1996.
- SHEPHARD N. (1996) "Statistical Aspects of ARCH and Stochastic Volatility", in *Time Series Models in Econometrics, Finance and Other Fields*, Cox, Hinkley and Barndorff-Nielsen Eds, Chapman and Hall, pp. 1-67.
- SILVERMAN B. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, Monographs on Statistics and Applied Probability 26, 175 pages.
- TASSI P. (1992) *Méthodes statistiques*, Economica.
- TAYLOR S. (1986) *Modeling Financial Time Series*, John Wiley Editor, 268 pages.
- TAYLOR S. (1994) "Modeling Stochastic Volatility : A Review and Comparative Study", *Mathematical Finance*, vol.4, n°2, April 1994, pp. 183-204.
- TUCKER A. (1992) "A Reexamination of Finite -and Infinite- Variance Distributions as Models of Daily Stock Returns", *Journal of Business and Economic Statistics*, vol.10, n°1, Janvier 1992, pp. 73-81.

- VENUTOLO T. et PÉLLIEUX G. (1997) *Empirical Results from the GARCH Option Pricing Model*, Actes du colloque A.E.A.-Imperial College-BNP, Londres, Mai 1997, 30 pages.
- WALTER CH. (1996) "Une histoire du concept d'efficience sur les marchés financiers", *Annales HSS*, n°4, juillet-août 1996, pp.873-905.
- WOOD R., McINISH T. et ORD J. (1985) "An Investigation of Transactions Data for NYSE Stocks", *Journal of Finance*, vol.40, n°3, 1985, pp. 723-741.
- ZHOU B. (1996) "Forecasting Foreign Exchange Rates Subject to De-Volatilization", in *Forecasting Financial Markets : Exchange Rates, Interest and Asset Management*, Dunis Editor, John Wiley and Sons, 287 pages.