

# Volatilité du CAC40 et Changements Structurels

Abdessamad Karem\*, Ahmed Louichi\*

Juin 2001

\*GEMMA-CREME

Faculté de sciences économiques et de gestion

Université de Caen, 14032 Caen Cedex FRANCE

akarem@hotmail.com

adli@econ.unicaen.fr

## Résumé

La dynamique du CAC40 est explorée sur une longue période. La volatilité de la série est modélisée par des processus ARCH, GARCH, GARCH-L et t-GARCH. Le problème de la persistance est appréhendé à travers les changements structurels de la série par une spécification SWARCH (ARCH avec changements de régime) dûe à J.Hamilton et R.Susmel(1994).

**Mots clés :** ARCH, GARCH, SWARCH, persistance, processus de Markov.

# Introduction

Dans ce papier, nous nous interrogeons sur la dynamique du CAC40 sur une longue période. La série de 3376 observations journalières (jours ouvrables, du 09/07/1987 au 29/12/2000) exhibe un aspect hétéroscédastique ainsi que quelques variations brutales et de grandes amplitudes. Ces variations notables peuvent correspondre à des changements structurels dans le comportement de la série. Les changements structurels du processus générant les données peuvent être imputés à des chocs de certaines variables de l'environnement économique, tels le crash d'Octobre 1987 ou à des changements de politique monétaire ou fiscale. Ces différentes causes ont des conséquences diverses sur la volatilité en question.

Dans une première section, nous présenterons une synthèse sur la classe des modèles ARCH (voir R.F. Engel 1982) et ARCH généralisé (GARCH, T.Bollerslev 1986). Dans la seconde section, nous établirons une série d'ajustements par ces modèles dont nous retiendrons les meilleures spécifications. Les formes GARCH(2,1) et GARCH(1,1) sont retenues comme ajustements adéquats du CAC40 sur la période considérée. Or, comme les modèles ARCH imputent, souvent, beaucoup de persistance dans l'estimation de la volatilité (voir J.D. Hamilton et R. Susmel 1994), nous présenterons dans la troisième section, un cadre général des modèles ARCH avec changements de régimes (Switching ARCH ou SWARCH). Cette classe de modèles permet de décrire de manière plus appropriée le comportement de beaucoup de séries financières. En effet, ces modèles autorisent les paramètres d'un processus ARCH à provenir d'un régime parmi plusieurs régimes différents, tout en tenant compte des transitions lors du passage d'un régime à l'autre. Ces transitions sont gouvernées par un processus de type Markovien. cela suppose l'existence de plusieurs états de la nature caractérisés par un processus stochastique à valeurs discrètes (chaînes de Markov). Ainsi, la série change dans son comportement en fonction de l'état prévalant. D'un point de vue théorique, la justification des modèles à changements discrets repose sur des anticipations rationnelles avec plusieurs régimes possibles.

Dans une quatrième section, nous exposerons les méthodes ou procédures de l'estimation de ces modèles et nous présenterons les résultats obtenus sur la base de 2 à 4 régimes possibles dans lesquels les innovations latentes proviennent de distribu-

tions Gaussiennes ou de Student. Ces résultats montrent que la spécification SWARCH constitue une réponse adéquate au problème de la persistance dans l'estimation de la volatilité dans les modèles GARCH obtenus dans la troisième section. Ils montrent également que ce type d'ajustement a une capacité plus grande de prévoir une variance variant dans le temps en dehors de l'échantillon, et par conséquent, obtenir de meilleures prévisions sur la variable dépendante.

Dans la cinquième section, nous concluons en mettant en évidence le degré de flexibilité que le modèle introduit dans l'estimation des modèles linéaires ou non linéaires, univariés ou multivariés, permettant ainsi de résoudre certains problèmes de biais et d'efficacité et corrigeant certaines formes de mauvaise spécification. Nous soulignerons également l'aptitude du modèle à capter les événements exceptionnels et à retourner à un comportement typique dans des situations normales.

Les limitations théoriques de ces modèles sont esquissées brièvement.

# 1 Les modèles ARCH

Les modèles de séries temporelles ont été historiquement développés, soit dans un but descriptif (désaisonnalisation, prévision), soit dans des buts de contrôle de systèmes. Ces divers problèmes ont, vers le milieu des années 70 conduit au développement d'une classe de modèles particuliers : les modèles autorégressifs moyenne mobile (ARMA). Ces modèles consistent à exprimer la valeur présente de la série  $\{Y_t\}$  comme fonction linéaire de ses valeurs passées  $\{Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}\}$  et des erreurs passées  $\{\epsilon_t, \epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-q}\}$  à laquelle s'additionne un bruit blanc :

$$Y_t = a_0 + \sum_{i=1}^{i=p} a_i Y_{t-i} + \sum_{j=1}^{j=q} \epsilon_{t-j} + u_t \quad (1)$$

avec  $E(u_t) = 0$ ,  $V(u_t) = \sigma^2$  et  $cov(\epsilon_i, \epsilon_j) = 0$  pour  $i \neq j$

Ces formulations présentent le double avantage d'être simples à mettre en oeuvre (à cause de l'aspect linéaire, qui facilite l'estimation et les calculs de prévisions) et de posséder d'importantes propriétés de robustesse. En effet même si la véritable dynamique sous-jacente n'est pas linéaire, les prévisions des valeurs et plus généralement de toutes les fonctions linéaires de ces valeurs, sont non biaisées. Ainsi une erreur de spécification éventuelle sur la dynamique n'a pas d'effet trop important au niveau de ces prévisions.

La question fondamentale est évidemment de savoir si en pratique nous ne sommes intéressés qu'à des prévisions de telles fonctions linéaires ou si nous le sommes aussi dans des prévisions de fonctions non linéaires des observations, qui, elles, pourraient être fortement biaisées si la dynamique représentée ci-dessus n'est pas exacte. La réponse est malheureusement affirmative. Les modèles ARCH constituent une classe de modèles non linéaires qui peut répondre à ce type de problèmes.

## 1.1 Le modèle ARCH(p) linéaire :

Les modèles ARCH (AutoRegressive Conditional Heteroskedasticity) : AutoRégressifs Conditionnellement Hétéroscédastiques, sont introduits récemment dans le but de mieux analyser les séries financières et monétaires. L'article fondateur de R. ENGLE est ainsi paru en 1982, où une paramétrisation possible pour la variance a été suggérée.

On peut exprimer la variance  $\sigma_t^2$  linéairement en fonction des carrés des valeurs passées du processus :

$$Y_t = \mu + \varepsilon_t + b\varepsilon_{t-1} \quad (2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \quad (3)$$

l'espérance conditionnelle de  $Y_t$  (équation (2)) est représentée sous forme d'un processus moyenne mobile d'ordre un. La présence de ce terme moyenne mobile ( $b\varepsilon_{t-1}$ ) permet de prendre en compte l'autocorrélation d'ordre un des cours. Ce modèle est connu sous le nom de modèle ARCH(p) linéaire. L'intérêt de cette approche réside dans l'introduction d'une dynamique endogène au niveau de la définition de la variance. Ces modèles fournissent une spécification originale de l'hétéroscédasticité dont les causes sont inconnues.

l'ensemble de l'information disponible à la date  $(t - 1)$  est :

$$(I_t = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-p}\})$$

La variance conditionnelle à  $I_t$ , notée  $\sigma_t$ , varie au cours du temps ; elle est représentée par l'équation(3). La variance non conditionnelle des innovations, notée  $\sigma_t^2$ , est l'espérance non conditionnelle de  $\sigma_t^2$ , c.à.d,  $\sigma^2 = E(\varepsilon_{t+1}^2) = E(E_t(\varepsilon_{t+1}^2)) = E(\sigma_t^2)$ . Elle est positive si  $\alpha_i > 0$ , pour  $i = 0, 1, \dots, p$ , et stationnaire si  $\sum_{i=1}^p \alpha_i < 1$ . Ce sont les contraintes de positivité.

La variance est estimée comme moyenne mobile des carrés des innovations. Contrairement aux solutions traditionnelles qui spécifient le modèle uniquement en fonction des variables explicatives observables, les modèles ARCH prennent en compte directement la variation temporelle de la variance du processus du phénomène étudié dans l'estimation des paramètres d'intérêt. Cette variance est une fonction de sa valeur décalée et d'une combinaison linéaire d'ordre p des carrés des erreurs. Sur les données financières, ces modèles permettent de prendre en compte les regroupements de volatilité, c'est à dire, les fortes (faibles) variations de prix sont suivies par d'autres fortes (faibles) variations de prix, mais dont le signe n'est pas prévisible, voir (B.Mandelbrot,1963).

## 1.2 Le modèle GARCH(p,q) linéaire :

Il est souvent nécessaire en pratique lorsque l'on tente d'identifier un modèle ARCH(p) linéaire de retenir un grand nombre  $p$  de retards. Les modèles ARCH généralisés, introduits par Bollerslev[1986], présentent une solution alternative qui a l'avantage de retenir une structure de retards plus souple. Dans le cas d'un modèle GARCH(p,q), on remplace l'équation de la variance conditionnelle (3) par l'équation suivante :

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (4)$$

Ceci impose des conditions restrictives sur les paramètres et, ces derniers ne doivent pas être négatifs, entre autres ; dans le modèle GARCH(1,1) par exemple, la variance n'est stationnaire que si  $\alpha_1 + \beta_1 < 1$ . Dans le processus ARCH(p), la variance conditionnelle est spécifiée sous une forme autorégressive, alors que le processus GARCH(p,q) permet l'introduction d'une partie moyenne mobile que l'on peut comparer à une sorte de mécanisme d'apprentissage adaptatif.

Ces modèles visent à décrire le comportement dynamique et non statique des séries temporelles, essentiellement les cours boursiers et les séries dérivées (taux, rentabilité, etc). La variance à une date donnée est considérée comme une variable aléatoire éventuellement dépendante des valeurs d'états antérieures ou contemporaines. La raison du succès de cette modélisation est notamment son caractère intuitif, il n'est pas difficile de constater que les marchés financiers connaissent des fièvres, mais aussi des périodes de grands consensus. Comme tout nouvel outil statistique, cette classe de modèles doit être étudiée en distinguant ses divers éléments.

Le modèle dynamique sous-jacent et les procédures d'inférence permettront d'ajuster modèles et séries réelles, de faire des tests d'hypothèse et de répondre aux problèmes de prévisions. Cette approche doit aussi être étudiée en fonction des diverses utilisations potentielles au domaine financier : analyse descriptive des séries de rendements, des taux de changes, détermination de prévisions de tendances, de risque et étude des prix qui s'en déduisent, tests de théorie financière (comme les conditions d'efficience de marché, les conditions d'équilibre type CAPM), recherche de facteurs.

Ces problèmes sont évidemment classiques et étaient analysés selon des procédures

habituellement fondées sur des modélisations linéaires de la dynamique (modèles autorégressifs moyenne mobile). L'intérêt des modèles ARCH ne sera important que si l'approche apporte plus que les procédures préexistantes, soit qu'elle permette d'affiner les interprétations, soit de corriger des conclusions qui étaient en fait biaisées.

### 1.3 Modèle ARCH non linéaire et non paramétrique :

Dans leur version initiale, les modèles ARCH et GARCH ne permettent pas de tenir compte de l'hypothèse d'asymétrie. Afin de corriger ce défaut, des travaux plus récents ont proposé des extensions de cette approche de façon à tenir compte de ces effets.

Dans le modèle GARCH(p,q) décrit en (4), la variance dépend uniquement de la valeur absolue de  $u_t$  non de son signe. Cela n'est pas compatible avec le comportement empirique des prix sur le marché des actions, pour lesquels il peut exister un effet de levier. Dans le modèle GARCH(p,q) exponentiel ou EGARCH(p,q) proposé par Nelson[1990],  $\sigma_t^2$  est une fonction asymétrique du passé de :

$$\log(\sigma_t) = \alpha_0 + \alpha(|z_{t-1}| - E(|z_{t-1}|)) + \gamma z_{t-1} + \beta \log(\sigma_{t-1}) \quad (5)$$

où  $E(|z|) = \sqrt{2\pi}$  sous l'hypothèse de normalité. Contrairement au modèle GARCH(p,q) de (4), aucune restriction sur les paramètres  $\alpha_i$  et  $\beta_j$  n'est nécessaire pour assurer la positivité de la variance conditionnelle. Les différents chocs sont ainsi représentés. Le processus de volatilité est stationnaire si la contrainte  $\beta < 1$  est vérifiée. L'étude de Nelson[1991] portait sur la rentabilité des actifs financiers, elle peut bien sûr, être étendue à la prime de risque ou à la mesure de l'aversion pour le risque des actifs. Un deuxième type de non linéarité, introduit par Gouriéroux et Monfort [1992] et Zakoian [1990], a conduit aux modèles ARCH à seuil (TARCH). L'idée de départ est simple : elle consiste à écrire l'équation de la variance conditionnelle, sous forme d'une fonction linéaire par morceaux, chacun des segments étant associé aux chocs de même nature.

$$\sigma_t = \alpha_0 + \alpha|\varepsilon_{t-1}| + \gamma \Pi_{t-1}^- \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (6)$$

où  $\Pi_{t-1}^-$  est une fonction indicatrice qui vaut 1 si  $\varepsilon_{t-1} < 0$  et 0 autrement. la volatilité est positive quand les paramètres de l'équation (6) vérifient  $\alpha_0 > 0$ ,  $\alpha \geq$

0,  $\alpha + \gamma \geq 0$ , et  $\beta \geq 0$ . le processus de volatilité est stationnaire, si la contrainte  $\beta^2 + \frac{\alpha^2 + (\alpha + \beta)^2}{2} + 2\beta^2 \frac{\alpha + \gamma}{2} < 1$  est vérifié.

Une troisième spécification de l'asymétrie a été fourni par (Glosten, Jagannathan et Runkel [1993] sous le nom du modèle GJR :

$$\sigma_t^2 = \alpha_0 + \alpha \varepsilon_{t-1}^2 + \gamma \Pi_{t-1}^- \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (7)$$

La volatilité est positive quand les paramètres de l'équation(7) vérifient  $\alpha_0 > 0$ ,  $\alpha \geq 0$ ,  $\alpha + \gamma \geq 0$ , et  $\beta \geq 0$ . La volatilité engendrée par ce processus est stationnaire si la contrainte  $\beta + \alpha + \frac{\gamma}{2}$  est vérifiée(Hentschel [1995]).

Ces deux dernières formulations se distinguent très nettement par l'indicateur de dispersion retenu. Il s'agit de la variance dans le premier cas et de la volatilité dans le second. Dans le modèle TARARCH , on peut observer une discontinuité de la dérivée de la variance conditionnelle par rapport aux innovations au voisinage de zéro ; un tel comportement n'est pas possible dans le cas du GJR. Ce dernier modèle apparaît comme une généralisation naturelle du modèle GARCH traditionnel, car il porte en effet sur la variance conditionnelle, alors que le TARARCH lui porte sur la volatilité conditionnelle et l'EGARCH sur le logarithme de celle-ci. Dans ces trois modèles,  $\beta$  représente le terme autorégressif,  $\alpha$  l'effet d'un choc sur le rendement et  $\gamma$  l'effet d'asymétrie résultant d'un choc négatif.

## 1.4 Méthode d'estimation :

L'estimation des modèles ARCH est généralement fondée sur la procédure du maximum de vraisemblance. L'optimisation est effectuée en supposant que les innovations standardisées  $z_t$  sont indépendantes et identiquement distribuées, d'espérance nulle et de variance unité.

Soit  $\theta = \{\mu, b, \alpha_0, \alpha, \beta, \gamma\}$  le vecteur des paramètres à estimer, conditionnellement aux valeurs initiales, la fonction de log-vraisemblance s'écrit alors :

$$L(\varepsilon_t; \theta) = \sum_{t=1}^T \ell_t(\varepsilon_t; \theta) \quad (8)$$

avec  $\ell_t(\varepsilon_t; \theta) = \log(f(z_t; \theta)) - 0.5 \log(\sigma_t^2(\theta))$  et où  $f(z_t; \theta)$  désigne la fonction de distribution conditionnelle des innovations standardisées représentant une distribution normale. Sous cette hypothèse de normalité, l'estimateur du maximum de vraisemblance s'interprète comme un estimateur du quasi-maximum de vraisemblance. Sous les conditions de régularité appropriées, l'estimateur du quasi-maximum de vraisemblance, noté  $\hat{\theta}_{QMV}$ , est convergent et asymptotiquement normal. Dans ce cas, la distribution asymptotique s'écrit :

$$\sqrt{T}(\hat{\theta}_{QMV} - \theta_0) \rightarrow \mathcal{N}(0, A_0^{-1} B_0 A_0^{-1}) \quad (9)$$

$$\text{avec } A_0 = \frac{1}{T} \sum_{t=1}^T \frac{\partial^2 \ell_t(z_t; \theta)}{\partial \theta \partial \theta'} \text{ et } B_0 = \frac{1}{T} \sum_{t=1}^T \left( \frac{\partial \ell_t(z_t; \theta)}{\partial \theta} \right) \left( \frac{\partial \ell_t(z_t; \theta)}{\partial \theta} \right)' .$$

Toutefois on suppose que les innovations standardisées suivent une densité normale d'espérance nulle et de variance unité. La distribution marginale du processus des innovations est caractérisé par une kurtosis supérieure à celle de la loi normale. En haute fréquence, les données financières ont des distributions non conditionnelles présentant des queues plus épaisses que celles de la loi gaussienne. Afin de remédier à ce problème, Bollerslev (1987) suggéra une loi de Student pour décrire la distribution des innovations. La fonction densité de la loi de student s'écrit sous la forme :

$$f(z_t; \theta) = \frac{1}{(\nu - 2)^{1/2}} \frac{\Gamma[(\nu + 1)/2]}{\sqrt{\pi} \Gamma(\nu/2)} \left( 1 + \frac{z_t^2}{(\nu - 2)} \right)^{-\frac{(\nu+1)}{2}} \quad (10)$$

Avec  $\nu > 0$  et où  $\Gamma()$  est la fonction gamma. La loi de student est symétrique par rapport à 0 et tend vers la loi normale quand le degré de liberté  $\nu$  tend vers l'infini. Pour  $4 < \nu < \infty$ , la kurtosis conditionnelle, définie par  $E(z_t^4) = \frac{3(\nu-2)}{\nu-4}$ , est supérieure à 3. Lorsque  $\nu$  est inférieur à 4, la kurtosis conditionnelle n'est pas définie. Dans le cas où les innovations suivent une loi de student, un paramètre supplémentaire doit être estimé pour permettre d'ajuster l'excès de kurtosis de la distribution empirique. Nelson [1991] a proposé la distribution GED (generalized Error distribution), qui s'écrit sous la forme :

$$f(z_t; \theta) = \frac{\eta}{2^{-2/\eta} \lambda \Gamma(1/\eta)} \exp\left(-0.5 \left| \frac{z_t}{\lambda} \right|^\eta\right) \quad (11)$$

avec  $\lambda = [2^{-1/\eta} \frac{\Gamma(1/\eta)}{\Gamma(3/\eta)}]^{1/2}$ . Quand le paramètre  $\eta$  est égale à 2, la densité se ramène à celle de la loi normale. Lorsque  $\eta < 2$ , la distribution a des queues plus épaisses que la loi normale, et, à l'inverse, quand  $\eta > 2$ , les queues sont moins épaisses que la loi normale. La kurtosis conditionnelle, définie par  $E(Z_t^4) = \frac{\Gamma(1/\eta)\Gamma(5/\eta)}{(\Gamma(3/\eta))^2}$ , est supérieure à 3 quand  $\eta < 2$ . Quand  $\eta = 1$ , on retrouve la loi de laplace.

## 2 Ajustement empirique du CAC40

### 2.1 Les données :

Les données sont issues de la table des marchés de Paribas à Frankfurt. Elles représentent 3376 valeurs du CAC40 allant du 9 juillet 1987 au 29 Décembre 2000. Ces données ont été transformées pour permettre la comparaison des résultats avec d'autres études financières. Cette modification consiste à calculer le logarithme du différentiel des valeurs en  $t$  et  $t - 1$ , ce qui confère à la variable une structure de taux de variation, (figure 1, les lignes verticales désignent les ruptures les plus importantes dans les deux séries).

### 2.2 Les statistiques :

Dans la table 1 sont représentés les différentes caractéristiques statistiques de la série du CAC40 sur la période couvrant les années 1987 à 2000. Afin de tester l'hypothèse de normalité de la distribution de cette série, la skewness et la kurtosis ont été ajoutées. Les résultats obtenus doivent en principe se rapprocher des hypothèses couramment émises dans la théorie financière, à savoir que les cours doivent être des variables aléatoires indépendantes et identiquement distribuées. D'autre part, leur distribution n'est pas normale mais plutôt leptokurtique et asymétrique.

Moyenne	0.040912
Médiane	0.045565
Maximum	8.225436
Minimum	-10.13757
Ecart-Type	1.278657
Skewness	-0.427581
Kurtosis	8.124637
Jarque-Bera	3794.795 (0.00000)

**Table 1.** Série du TCAC40 (3375 observations)

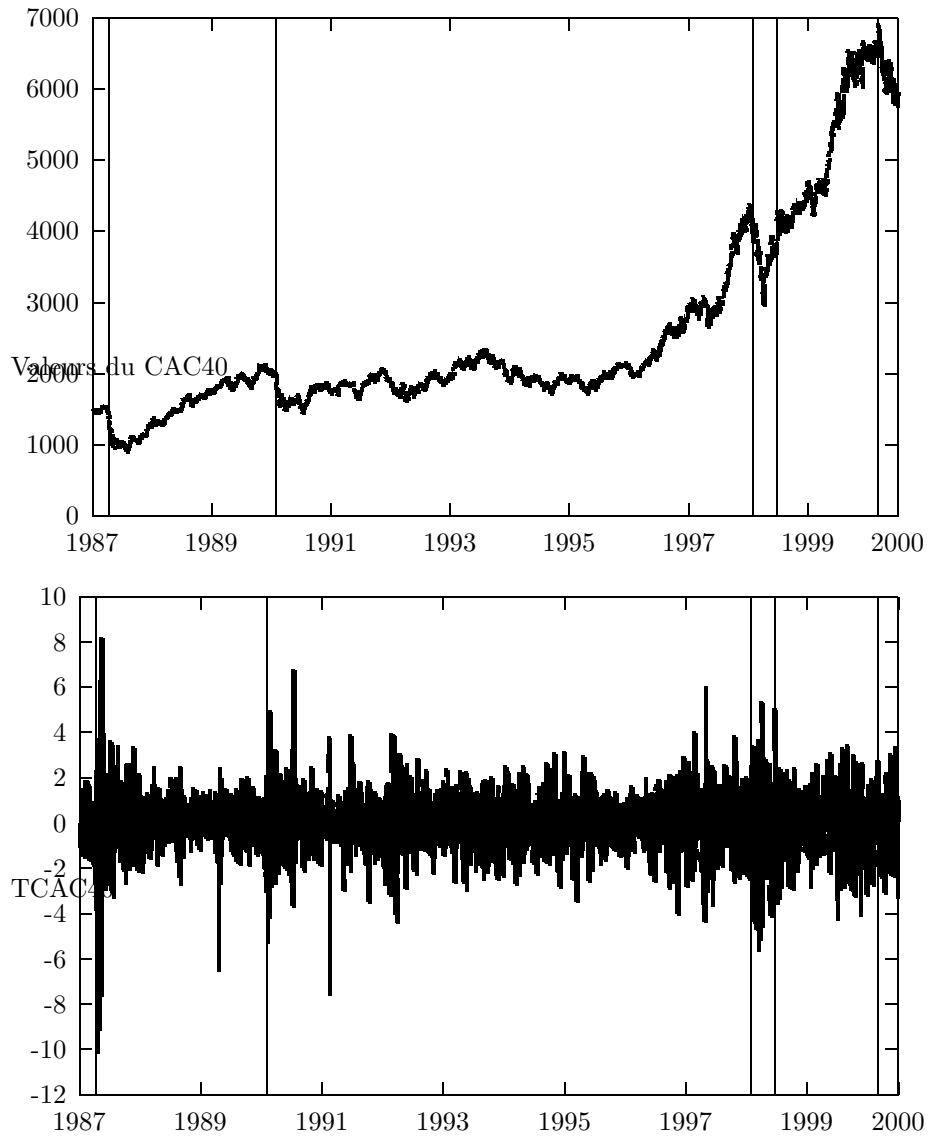


FIG. 1 – Evolution du CAC40 et du taux de variation du CAC40.

L'analyse de la skewness et de la kurtosis conduit aux conclusions usuelles dans les études des cours boursier. Elles sont différentes de 0 et 3 ( -0,427581 et 8,124637), ce qui signifie que la distribution n'est pas normale mais plutôt asymétrique avec des queues épaisses caractérisant une distribution Leptokurtique.

L'hypothèse de variables indépendantes est aussi rejetée, car la statistique Q de Ljung-box calculée avec 20 retards indique des autocorrélation de la série dès le premier retard (Q=5.7683).

Afin de vérifier ce résultat, il est nécessaire de s'intéresser aux fonction d'autocorrélation des résidus. En effet, si une série est bruit blanc strict, alors le sont également les séries déduites du carré et de la valeur absolue de celle-ci. Or nous remarquons clairement une dépendance des variables entre elles, traduites par des autocorrélations significatives durables pour les deux séries. Ceci nous conduit au rejet de l'hypothèse d'absence d'autocorrélation des cours et met en évidence la présence d'une hétéroscédasticité, confirmée également par le test ARCH.

La série TCAC étudiée ne suit ni un processus bruit blanc strict ni un processus bruit blanc, mais présente une dépendance à la fois linéaire et non linéaire sur ses valeurs passées, quant à la nature de sa distribution, bien qu'inconnue, elle est de forme leptokurtique. D'après ces résultats, le modèle d'estimation à construire doit satisfaire deux hypothèses :

1. Une dépendance linéaire due aux autocorrélations sur le retard d'ordre 1 de la série.
2. Une dépendance non linéaire due aux autocorrélations de la série TCAC sur ses premiers retards.

Afin de construire une série  $\varepsilon_t$  bruit blanc, c'est à dire non corrélée avec ses retards à partir de la série TCAC dont on dispose, la méthode utilisée est celle de la régression linéaire simple dans le cadre d'un processus autorégressif d'ordre 1. Les résidus de cette régression forment la série  $\varepsilon_t$  :

$$TCAC_t = a_0 + a_1 TCAC_{t-1} + \varepsilon_t \quad (12)$$

ainsi la série n'est plus corrélée avec ses valeurs passées. Les coefficients sont significativement différents de zéro à 5% et les simples tests de Durbin-Watson et Ljung-Box

indiquent une absence de corrélation des résidus. Afin de vérifier la totale fiabilité de ce modèle, il est nécessaire d'étudier les résidus.

La skewness et la kurtosis n'ont pas évolué, la distribution reste plutôt leptokurtique. La régression linéaire effectuée sur la série TCAC semble bien conduire à une série de résidus suivant un processus bruit blanc d'après la statistique Q, qui indique une absence d'autocorrélation. Cependant l'étude des fonctions d'autocorrélation des séries  $\varepsilon_t^2$  et  $|\varepsilon_t|$  indique que le processus n'est pas indépendant, mais qu'il existe pour ces séries des autocorrélations significatives même pour des retards assez importants. Puisque les cours du CAC40 sont non corrélés, mais leur carré l'est, une possibilité alternative de modéliser cette série est le processus ARCH. Le test pour la présence d'un effet ARCH, LM test (test du multiplicateur de Lagrange), indique clairement la présence d'un tel effet.

De manière générale un processus ARCH univarié peut être défini comme solution d'un modèle de la forme :

$$\varepsilon_t = z_t \sigma_t \quad (13)$$

où  $z_t$  est un processus iid indépendant du passé de  $\varepsilon_t$ , avec  $E(z_t) = 0$  et  $V(z_t) = 1$ , donc centré et de variance unité.  $\sigma_t$  est une fonction mesurable du passé de  $\varepsilon_t$ .

Notons  $\varepsilon_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-p}\}$  la tribu engendrée par les valeurs passées de  $\varepsilon_t$ , qui est l'ensemble d'information disponible. L'estimation du modèle est fondée sur la méthode du maximum de vraisemblance .

Le modèle estimé est donc :

$$TCAC_t = a_0 + a_1 TCAC_{t-1} + \varepsilon_t \quad (14)$$

dont les erreurs suivent théoriquement un GARCH(p,q) :

$$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^q \beta_j h_{t-1}^2 \quad (15)$$

avec  $1 \leq p \leq 4$  et  $0 \leq q \leq 4$ .

Après différents essais de spécification ARCH(1), ARCH(2),  $\dots$ , GARCH(1,1), GARCH(1,2),  $\dots$ , le modèle le plus approprié est GARCH(1,1) dont les coefficients

sont significatifs. L' équations estimé pour un GARCH(1,1) est :

$$TCAC_t = 0.58576 + 0.058187.TACAC_{t-1}$$

(3.0970)    (3.105214)

avec

$$\sigma_t^2 = 0.062525 + 0.100261.\varepsilon_{t-1}^2 + 0.861102.\sigma_{t-1}^2$$

(8.297568)    (12.4323)    (77.1121)

où les nombres entre parenthèses sont les t de Student asymptotiques habituels.

### 2.3 Les performances prévisionnelles

Si la spécification était correcte et les paramètres connus avec certitude, alors  $\sigma_t^2$  serait l'espérance conditionnelle de  $u_t^2$ . Par conséquent, la fonction de perte moyenne des erreurs quadratiques  $E\{(u_t^2 - \theta_t)^2 | u_{t-1}, u_{t-2}, \dots\}$  serait minimisée par rapport à  $\theta_t$  en prenant  $\theta_t = \sigma_t^2$  pour  $\sigma_t^2$  donnée par l'équation de la variance estimée. La valeur moyenne de  $(u_t^2 - \hat{\sigma}_t^2)^2$  est :

$$MSE = T^{-1} \sum_{t=1}^T (\hat{u}_t^2 - \hat{\sigma}_t^2)^2 \quad (16)$$

Afin d'établir des comparaisons standards entre les modèles, désignons par  $\bar{y}_t$  la moyenne d'échantillonnage du taux de variation du CAC40, (ou du rendement d'une option), et par  $s^2$  la variance non conditionnelle d'échantillonnage, soit :

$$\bar{y} = T^{-1} \sum_{t=1}^T y_t, \quad s^2 = T^{-1} \sum_{t=1}^T (y_t - \bar{y})^2$$

Ainsi, si nous faisons des prévisions de la variance de  $y_t$  dans l'échantillon, la fonction de perte est donnée par :

$$MSE = T^{-1} \sum_{t=1}^T [(y_t - \bar{y})^2 - s^2]^2$$

Certains modèles obtenus, ont donné des prévisions plus pauvre qu'un simple modèle avec variance constante.

## 2.4 La persistance

Le haut degré de persistance de la volatilité impliqué par l'équation de la variance entache les performances prévisionnelle du modèle. Afin de donner une mesure à cette persistance, remplaçons  $t$  par  $t+m$  dans l'équation de la variance. Nous obtenons pour un GARCH(1,1) :

$$E(u_{t+m}^2 | u_{t+m-1}, u_{t+m-2}, \dots) = a_0 + a_1 u_{t+m-1}^2 + b_1 E(u_{t+m-1}^2 | u_{t+m-2}, u_{t+m-3}, \dots) \quad (17)$$

Désignons par  $\sigma_{t+m|t}^2$  la prévision de la variance à l'horizon de  $m$  périodes :

$$\sigma_{t+m|t}^2 \equiv E(u_{t+m}^2 | u_t, u_{t-1}, \dots)$$

Ainsi, on a  $\sigma_{t+1|t}^2 \equiv \sigma_{t+1}^2$ . En utilisant cette notation, (17) peut être réécrite de la manière suivante :

$$\sigma_{t+m|t}^2 = a_0 + (a_1 + b_1) \sigma_{t+m-1|t}^2$$

Ainsi, la prévision à l'horizon de  $m$  périodes est tout simplement régie par une équation différentielle du premier ordre, avec un paramètre  $\lambda$  défini par :

$$\lambda = (a_1 + b_1)$$

quand il n'y a pas d'effet de levier ( $\xi = 0$ ). Quand il y a effet de levier, on a  $\lambda = (a_1 + b_1 + \xi/2)$ .

## 3 Spécification ARCH et changements de régime

Une certaine faiblesse des performances prévisionnelle des modèles ARCH peut être imputée aux changements structurels dans le processus. L'estimation d'une forte persistance via le paramètre  $\lambda$  est connue pour être non-robuste à travers des parties de l'échantillon (sous-échantillons). Diebold(1986) et Lamoureux et Lastrapes (1990) justifient la forte valeur estimée de  $\lambda$  par des changements structurels de la variance du processus. Ceci renvoie à la remarque de Perron (1989), que des changements de régimes peuvent donner l'impression de racines unités dans la caractérisation en niveau d'une série.

### 3.1 Le cadre théorique

Cela nous conduit à une spécification dans laquelle les paramètres d'un processus ARCH peuvent changer de façon occasionnelle. Désignons par  $y_t$  le vecteur des variables observables, et par  $s_t$  une variable aléatoire non-observable prenant des valeurs entières  $1, 2, \dots, k$ . Supposons que  $s_t$  puisse être décrite par une chaîne de Markov :

$$Prob(s_t = j | s_{t-1} = i, \dots, s_{t-k} = k, y_{t-1}, y_{t-2}, \dots) = Prob(s_t = j | s_{t-1} = i) = p_{ij}$$

pour  $i, j = 1, 2, \dots, K$ . Il est pratique de représenter ces probabilités par la matrice  $K \times K$  suivante :

$$P = \begin{pmatrix} p_{11} & p_{21} & \cdots & p_{K1} \\ p_{12} & p_{22} & \cdots & p_{K2} \\ \vdots & \vdots & \cdots & \vdots \\ p_{1K} & p_{2K} & \cdots & p_{KK} \end{pmatrix} \quad (18)$$

où chaque colonne somme à l'unité. La variable  $s_t$  est interprétée comme l'état ou régime où se trouve le processus à la date  $t$ . La variable  $s_t$  gouverne les paramètres de la distribution conditionnelle de  $y_t$ . Ainsi, si la densité de  $y_t$  conditionnellement à sa valeur actuelle et à  $q$  valeurs retardées est de la forme :

$$f(y_t | s_t, s_{t-1}, \dots, s_{t-q}, y_{t-1}, y_{t-2}, \dots, y_0) \quad (19)$$

alors, les méthodes développées dans Hamilton (1989) peuvent être utilisées afin d'évaluer la fonction de vraisemblance pour les données observées et inférer les régimes non-observables. En effet,  $y_t$  pourrait suivre un processus ARCH( $q$ ) dont les paramètres dépendent des réalisations non-observables de  $s_t, s_{t-1}, \dots, s_{t-q}$ . De tels modèles ont été utilisés pour l'ajustement de données sur l'inflation par Brunner (1991) ainsi que sur les rendements des bons du trésor US par Cai. Bien que (19) fournit un cadre général pour décrire le changement structurel, elle est limitée en ce qui concerne le nombre fini  $q$  de retards sur  $s$  dont pourrait dépendre la densité  $y_t$ . Les paramètres d'un ARCH( $q$ ) peuvent changer, mais ceux d'un GARCH( $p, q$ ), avec  $p > 0$  ne sont pas autorisés à changer comme cas particulier de (19), et l'objectif consiste alors à sélectionner, de manière parsimonieuse, une représentation pour les différents régimes possibles. Une spécification dans laquelle tous les paramètres changent dans un régime, impliquerait

une surparamétrisation et conduirait à des calculs numériques hors de portée. Hamilton (1989) suggèra le modèle avec changement de régime pour la moyenne conditionnelle :

$$y_t = \mu_{s_t} + \tilde{y}_t$$

où  $\mu_{s_t}$  désigne le paramètre  $\mu_1$  quand le processus est dans le régime représenté par  $s_t = 1$ ,  $\mu_2$  quand  $s_t = 2$ ,  $\dots$  etc. La variable  $\tilde{y}_t$  est supposée suivre un modèle autorégressif d'ordre  $q$  est de moyenne nulle :

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \phi_2 \tilde{y}_{t-2} + \dots + \phi_q \tilde{y}_{t-q} + \varepsilon_t$$

L'idée derrière cette spécification est que tout changement occasionnel brutal dans le niveau moyen de  $\tilde{y}_t$ , serait capté par la valeur de  $\mu_{s_t}$ .

Une extension naturelle de cette approche pour la variance conditionnelle, consiste à modéliser les résidus  $u_t$  de la manière suivante :

$$u_t = \sqrt{g_{s_t}} \cdot \tilde{u}_t$$

où  $u_t$  est supposé suivre un processus ARCH-L(q) standard.

$$\tilde{u}_t = h_t \cdot v_t$$

où  $v_t$  est une séquence *i.i.d.*,  $N(0, 1)$ , tandisque  $h_t$  est régie par :

$$h_t^2 = a_0 + \sum_{i=1}^q a_i \tilde{u}_{t-i}^2 + \xi \cdot d_{t-1} \cdot \tilde{u}_{t-1}^2$$

où  $d_{t-1} = 1$  si  $\tilde{u}_{t-1} \leq 0$  et  $d_{t-1} = 0$  pour  $\tilde{u}_{t-1} > 0$ . La variable  $\tilde{u}_t$  du modèle ARCH-L(q) est alors multipliée par la constante  $\sqrt{g_1}$  quand le processus est représenté par le régime  $s_t = 1$ , et par  $\sqrt{g_2}$  quand  $s_t = 2, \dots$ , et ainsi de suite. Le facteur  $g_1$  pour le premier état est normalisé à l'unité avec  $g_j \geq 1$  pour  $j = 2, 3, \dots, K$ . Ainsi, la modélisation des changements de régime dans le modèle, revient à modéliser les changements dans l'échelle du processus. Conditionnellement à la connaissance des régimes courants et passés, la variance des résidus  $u_t$  est :

$$\begin{aligned} E(u_t^2 | s_t, s_{t-1}, \dots, s_{t-q}, u_{t-1}, u_{t-2}, \dots, u_{t-q}) \\ &= g_{s_t} \left( a_0 + \sum_{i=1}^q a_i \frac{u_{t-i}^2}{g_{s_{t-i}}} \right) + \xi \cdot d_{t-1} \cdot \frac{u_{t-1}^2}{g_{s_{t-1}}} \\ &\equiv \sigma_t^2(s_t, s_{t-1}, \dots, s_{t-q}) \end{aligned} \quad (20)$$

où  $d_{t-1} = 1$  si  $\tilde{u}_{t-1} \leq 0$  et  $d_{t-1} = 0$  pour  $\tilde{u}_{t-1} > 0$ .

Dans l'absence d'un effet de levier ( $\xi = 0$ ), nous dirons que  $u_t$  suit un processus ARCH avec changements de régime Markovien d'ordre  $q$  à  $K$  états, ou tout simplement, un processus SWARCH( $K, q$ ). Dans le cas d'un effet de levier, ce sera un processus SWARCH-L( $K, q$ ). Pour la distribution de  $v_t$ , deux versions du modèle ont été explorées :  $v_t \sim N(0, 1)$  pour une distribution Gaussienne, et  $v_t \sim t(\nu, 1)$  pour une distribution de Student à  $\nu$  degrés de liberté et une variance unitaire.

## 4 Procédure d'estimation

La fonction Log-Vraisemblance s'écrit :

$$\mathcal{L} = \sum_{t=1}^T \log f(y_t | y_{t-1}, y_{t-2}, \dots, y_0) \quad (21)$$

qui est maximisée numériquement<sup>1</sup> par rapport aux paramètres suivants :  $\alpha, \phi, a_0, a_1, a_2, \dots, a_q, p_{11}, p_{12}, \dots, p_{KK}, g_1, g_2, \dots, g_k, \xi$  et  $\nu$ , tout en tenant compte des contraintes :  $g_1 = 1, \sum_{j=1}^K p_{ij} = 1$ , pour  $i = 1, 2, \dots, K$  et  $0 \leq p_{ij} \leq 1$  pour  $i, j = 1, 2, \dots, K$ . Pour les deux dernières contraintes, on utilise la reparamétrisation suivante :

$$\begin{aligned} p_{ij} &= \frac{\theta_{ij}^2}{1 + \sum_{\ell=1}^{K-1} \theta_{i\ell}^2} \quad \text{pour } j = 1, 2, \dots, K-1 \\ &= \frac{1}{1 + \sum_{\ell=1}^{K-1} \theta_{i\ell}^2} \quad \text{pour } j = K \end{aligned}$$

en estimant les  $\theta_{ij}$  pour  $i = 1, 2, \dots, K$  et  $j = 1, 2, \dots, K-1$  sans aucune restriction. L'appendice décrit l'inférence concernant l'état particulier où se trouve le processus à la date  $t$ . Quand cette inférence est basée sur l'information jusqu'à la date  $t$ , on l'appellera le "filtre des probabilités" :

$$p(s_t, s_{t-1}, \dots, s_{t-q} | y_{t-1}, y_{t-2}, \dots, y_0) \quad (22)$$

C'est la probabilité qu'à la date  $t$ , l'état était  $s_t$ , qu'à la date  $t-1$ , l'état était  $s_{t-1}, \dots$ , et qu'à la date  $t-q$ , l'état était  $s_{t-q}$ . Il y a  $K^{q+1}$  configurations possibles pour  $(s_t, s_{t-1}, \dots, s_{t-q})$ ,

<sup>1</sup>Voir la description de l'algorithme d'optimisation en annexe

et par conséquent, il y a  $K^{q+1}$  nombres de la forme (19) dont la somme est égale à 1 par construction.

Alternativement, l'échantillon complet des observations peut être utilisé pour construire les "probabilités lissées"

$$p(s_t|y_T, y_{T-1}, \dots, y_0) \quad (23)$$

Cette expression définit les  $K^{q+1}$  nombres pour chaque date  $t$  dans l'échantillon. Ces nombres somment à 1 par construction.

## 4.1 Le problème des prévisions

Afin de calculer la valeur  $u_{t+m}^2$  prévue de  $u_t^2$  à l'horizon  $m$ , nous considérons la situation hypothétique dans laquelle nous connaissons les valeurs de  $s_t, s_{t-1}, \dots, s_{t-q+1}$  avec certitude, et donc, nous connaissons également avec certitude, les valeurs de  $\tilde{u}_t^2 = u_\tau/\sqrt{g_s}$ , pour  $\tau = t, t-1, \dots, t-q+1$ . Pour cet ensemble d'information, la prévision de  $u_{t+m}^2$  serait :

$$\begin{aligned} & E(u_{t+m}^2 | s_t, s_{t-1}, \dots, s_{t-q+1}, \tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-q+1}) \\ &= E(g_{s_{t+m}} \tilde{u}_{t+m}^2 | s_t, s_{t-1}, \dots, s_{t-q+1}, \tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-q+1}) \\ &= E(g_{s_{t+m}} | s_t, s_{t-1}, \dots, s_{t-q+1}) \cdot E(\tilde{u}_{t+m}^2 | \tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-q+1}) \end{aligned} \quad (24)$$

car  $s_t$  est indépendante de  $v_t$  et de  $\tilde{u}_t$ ,  $\forall t$  et  $\forall \tau$ . Et puisque  $s_t$  suit une chaîne de Markov, le premier terme dans (24) peut s'écrire :

$$E(g_{s_{t+m}} | s_t, s_{t-1}, \dots, s_{t-q+1}) = \sum_{j=1}^K g_j \cdot Prob(s_{t+m} = j | s_t) \quad (25)$$

La probabilité de transition à l'horizon  $m$  peut être obtenue en multipliant la matrice dans (18) par elle-même  $m$  fois :

$$P^m = \begin{pmatrix} Prob(s_{t+m} = 1 | s_t = 1) & Prob(s_{t+m} | s_t = 2) & \dots & Prob(s_{t+m} | s_t = K) \\ Prob(s_{t+m} = 2 | s_t = 1) & Prob(s_{t+m} | s_t = 2) & \dots & Prob(s_{t+m} | s_t = K) \\ \vdots & \vdots & \dots & \vdots \\ Prob(s_{t+m} = K | s_t = 1) & Prob(s_{t+m} | s_t = 2) & \dots & Prob(s_{t+m} | s_t = K) \end{pmatrix}$$

Ainsi, si on regroupe les facteurs de commutations dans un vecteur  $g'$  de dimension  $(1 \times K)$ , soit  $g' \equiv (g_1, g_2, \dots, g_K)$ , on a :

$$E(g_{st+m} | s_t = i) = g' P^m e_i \quad (26)$$

où  $e_i$  désigne la  $i^{\text{ème}}$  colonne de la matrice unité  $(K \times K)$ .

Le second terme est simple à construire du fait que  $\tilde{u}_t$  suit un processus ARCH-L(q) standard :

$$\begin{aligned} E(\tilde{u}_{t+m}^2 | \tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-q+1}) & \quad (27) \\ &= a_0 + \sum_{i=1}^{q-1} a_i \tilde{u}_{t-i}^2 + \xi \cdot d_t \cdot \tilde{u}_t^2 \quad \text{pour } m = 1 \\ &= a_0 + (a_1 + \xi/2) \cdot \tilde{h}_{t+m-1|t}^2 + \sum_{i=2}^{m-q/t} \tilde{h}_{t+i}^2 \quad \text{pour } m = 2, 3, \dots, \end{aligned}$$

où

$$\tilde{h}_{\tau|t}^2 = \begin{cases} \tilde{u}_{\tau}^2 & \text{pour } \tau \leq t \\ E(\tilde{u}_{\tau}^2 | \tilde{u}_t^2, \tilde{u}_{t-1}^2, \dots) & \text{pour } \tau > t \end{cases} \quad (28)$$

La séquence  $(\tilde{h}_{\tau|t}^2)$  pour  $\tau = t+2, t+3, \dots$  est alors calculée par itération de (27). Rappelons que  $\tilde{u}_t = u_t / \sqrt{g_{s_t}}$ , la prévision dans (27) est une fonction de  $u_{(q)} = (u_t, u_{t-1}, \dots, u_{t-q+1})$  et de l'ensemble particulier des valeurs  $s_{(q)} = (s_t, s_{t-1}, \dots, s_{t-q+1})$  qui était :

$$E(\tilde{u}_{t+m}^2 | \tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-q+1}) = \tilde{h}_{t+m|t}^2(u_{(q)})$$

Ainsi, la prévision dans (24) peut être réécrite, avec  $\tilde{u}_{(q)} = (\tilde{u}_t, \tilde{u}_{t-1}, \dots, \tilde{u}_{t-q+1})$  :

$$E(u_{t+m}^2 | s_q, \tilde{u}_t) = (g' P^m e_i) \cdot \tilde{h}_{t+m|t}^2(s_{(q)}, u_{(q)}) \quad (29)$$

Pour des valeurs données de  $u_{(q)}$ , l'expression (29) décrit une autre manière de prévoir  $u_{t+m}^2$  pour chaque configuration possible de  $s_{(q)}$  que nous pourrions désigner par :

$$E(u_{t+m}^2 | s_{(q)}, u_{(q)}) \equiv \eta(s_{(q)}, u_{(q)}) \quad (30)$$

En pratique, nous ne connaissons pas la valeur de  $s_{(q)}$ . Cependant, par la loi des espérances itérées, nous avons :

$$\begin{aligned} \sigma_{t+m|t}^2 &= E(u_{t+m}^2 | u_{(q)}) \\ &= \sum_{s_t=1}^K \sum_{s_{t-1}=1}^K \dots \sum_{s_{t-q+1}=1}^K \eta(s_{(q)}, u_{(q)}) \times p(s_{(q)} | y_t, y_{t-1}, \dots, y_{-3}) \quad (31) \end{aligned}$$

Cela revient tout simplement à pondérer chacune des prévisions conditionnelles dans (29) par la probabilité filtre correspondant à chaque configuration afin de calculer la valeur prévue de  $u_{t+m}^2$  à l'horizon  $m$ , en se basant sur les données effectivement observées.

De cette manière, le modèle fournit une structure qui pourrait générer ce type de non-linéarité qui caractérise la volatilité de la plupart des séries financières. Friedman et Laibson (1989), et Friedman (1992) expliquent cela par le fait que les modèles ARCH standards, échouent à bien prévoir cet aspect car les chocs de petites amplitudes et les chocs d'amplitudes importantes engendrent des effets différents. Dans le contexte du modèle décrit avant, supposons que l'analyste est confiant que le marché ait été dans l'état 1 pendant les  $q$  périodes passées et que  $p_{11}$  est proche de l'unité. Si le résidu de la date  $t$  était faible, l'analyste voudrait continuer à accorder une forte probabilité à l'événement que  $s_t = 1$ , ainsi sa prévision devrait être :

$$E(u_{t+m}^2 | u_{(q)}) = a_0 + a_1(u_t^2/g_1) + a_2(u_{t-1}^2/g_1) + \dots + a_q(u_{t-q+1}^2/g_1) + \xi \cdot d_t \cdot (u_t^2/g_1)$$

L'effet marginal de  $u_t^2$  sur la prévision serait donné par  $(a_1 + \xi d_t)/g_1$ . De l'autre côté, si le résidu était suffisamment important, l'analyste serait persuadé que le régime s'est déplacé à  $s_t = 2$ , et l'effet marginal de  $u_t^2$  sur la prévision serait donné par  $(a_1 + \xi d_t)/g_2$ . Cette spécification tient compte de la non-linéarité qui émerge comme conséquence de l'inférence sur la volatilité du régime courant.

## 4.2 Le problème de la persistance

La caractérisation de la persistance de la composante ARCH du processus SWARCH peut être établie à partir de l'expression(27), où la prévision de  $u_{t+m}^2$  est régie par l'équation aux différences finies à l'ordre  $q$  suivante :

$$\tilde{h}_{t+m|t}^2 = a_0 + (a_1 + \xi/2) \cdot \tilde{h}_{t+m-1|t}^2 + a_2 \tilde{h}_{t+m-2|t}^2 + \dots + a_q \tilde{h}_{t+m-q|t}^2$$

Une solution de cette équation est de la forme :

$$\tilde{h}_{t+m|t}^2 = c_0 + c_1 \lambda_1^m + c_2 \lambda_2^m + \dots + c_q \lambda_q^m$$

où  $\lambda_1, \lambda_2, \dots, \lambda_q$  sont les valeurs propre de la matrice :

$$\begin{pmatrix} (a_1 + \xi/2) & a_2 & \cdots & a_{q-1} & a_q \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

La plus grande valeur propre est identifiée alors comme la mesure de la persistance de la composante ARCH d'un processus SWARCH-L(K,q).

### 4.3 Résultats empiriques

Une série d'ajustements a été réalisés par les modèles GARCH avec ou sans effet de levier et dont les erreurs suivent une distribution Normale ou de Student.

L'équation générique des modèles GARCH-L(p,q), avec  $0 \leq p \leq 4$  et  $0 \leq q \leq 4$  est :

1. Equation de la moyenne :

$$TCAC_t = a_0 + b.TCAC_{t-1} + \gamma.\sigma_{t-1}\varepsilon_t$$

2. Equation de la variance :

$$\sigma_t^2 = \alpha_0 + \alpha_1.\varepsilon_{t-1}^2 + \beta.\sigma_{t-1}^2 + \xi.d_{t-1}.\varepsilon_{t-1}^2$$

avec  $\lambda = (\alpha_1 + \beta + \xi/2)$

Les résultats des spécifications retenues sont donnés dans la table 2. Ces spécifications passent tous les tests standards (Fisher, Durbin-Watson, White, Goldfeld-Quandt, Critèred'Information d'akaike, Schwarz  $\dots$  etc. Certaines spécifications, telles ARCH(4,4), ARCH(4,3) aboutissent à des performances plus faibles qu'une spécification avec variance constante, et donc des prévisions plus pauvres.

Modèle	LogL	$\hat{a}_0$	$\hat{b}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\xi}$	$\hat{\lambda}$	$\hat{\nu}$	%MSE
ARCH(1)	-5526.7	.39 (1.95)	.133 (10.43)	1.29 (54.07)	.211 (13.11)	(-) (-)	(-) (-)	(-) (-)	.211 (-)	(-) (-)	-.0067 (-)
ARCH-M(1)	-5520.3	.75 (7.65)	.10 (7.05)	1.28 (53.21)	.21 (12.60)	(-) (-)	-.58 (-7.48)	(-) (-)	.21 (-)	(-) (-)	-.022 (-)
GARCH(1,1)	-5318.93	.058 (3.08)	.58 (3.10)	.062 (8.28)	.10 (12.43)	.86 (77.11)	(-) (-)	(-) (-)	.96 (-)	(-) (-)	.001 (-)
GARCH-M(1,1)	-5318.79	-.12 (-1.98)	.056 (2.95)	.06 (7.73)	.10 (12.5)	.86 (75.39)	.17 (2.31)	(-) (-)	.96 (-)	(-) (-)	-.0007 (-)
GARCH-L(1,1)	-5300.75	.037 (1.99)	.058 (3.17)	.062 (8.69)	.031 (2.99)	.88 (76.52)	(-) (-)	.093 (7.69)	(.96) (-)	(-) (-)	.001 (-)
GARCH-ML(1,1)	-5301.86	-.11 (-1.82)	.06 (3.21)	.067 (8.70)	.035 (3.20)	.87 (73.78)	.14 (1.81)	.091 (7.32)	.955 (-)	(-) (-)	.0005 (-)

**Table 2.** Comparaisons des spécifications retenues pour la série TCAC. Les nombres entre parenthèses sont les  $t(z^2)$  de Student. %MSE est le rapport de la fonction de perte du modèle rapportée à celle du modèle avec variance constante, permet de comparer les performances des modèles.

**Spécification avec changements de régime retenue :**

$$TCAC_t = a_0 + b_1 TCAC_{t-1} + \varepsilon_t$$

$$\varepsilon_t = \sqrt{g_{st}} \cdot \hat{\varepsilon}_t$$

$$\hat{\varepsilon}_t = h_t \cdot v_t$$

$v_t \sim$  iid Student t avec variance unité et  $\nu$  ddl

C'est un modèle SWARCH-L(3,2) : ARCH d'ordre 2 à 3 régimes avec effet de levier et distribution de Student.

**A-Estimation initiale :**

Une estimation de démarrage ou initiale est réalisé avec un premier régime non contraint à suivre un processus ARCH et une probabilité de transition de .992366. La valeur initiale de la variance  $h_0$  n'est pas nécessaire.

Le modèle estimée est :

$$TCAC_t = .0328 + .0603 \cdot TCAC_{t-1}$$

(3.439)                      (3.039)

$$\sigma_t^2 = .065 + .0324 \cdot \varepsilon_{t-1}^2 + .873 \cdot \varepsilon_{t-2}^2 + .101 \cdot d_{t-1} \varepsilon^2 t - 1$$

(10.124)                      (4.294)                      (3.872)                      (6.2455)

avec

$$\hat{\nu} = 7.219$$

(2.849)

$$\text{LogL} = -5488.92$$

Les variances pour le états 1 et 2 sont données par le vecteur :(4.3512,13.1465)

La matrice (transposée) de transition estimée est :

$$P = \begin{pmatrix} .992366 & .007633 & .00000 \\ .000000 & .991445 & .008554 \\ .002570 & .0143530 & .983076 \end{pmatrix}$$

**B-Estimation finale :**

L'estimation finale est obtenue avec un premier régime non contraint à être un ARCH et une variance initiale  $h_0$  égale à la moyenne des résidus au carré obtenus de l'estimation de l'équation de la moyenne (T. Bollerslev 86).

Le modèle final estimé est :

$$\begin{aligned}
 TCAC_t &= .04471 & + & .05326.TCAC_{t-1} \\
 &(3.439) & & (3.039) \\
 \sigma_t^2 &= .4471 & + & -2.69.10^{-11}.\varepsilon_{t-1}^2 & + & -.05323.\varepsilon_{t-2}^2 & + & -.04761.d_{t-1}\varepsilon^2t - 1 \\
 &(10.124) & & (4.294) & & (3.872) & & (6.2455) \\
 \text{avec} \\
 \hat{\nu} &= 17.5029 \\
 &(2.849) \\
 \text{LogL} &= -5270.5821
 \end{aligned}$$

Le vecteur des variances estimées pour les états 1 et 2 est (2.485,8.485).

La matrice (transposée) de transition estimée est :

$$P = \begin{pmatrix} .989140 & .010859 & .00000 \\ .000000 & .992594 & .007405 \\ .010313 & .021224 & .968462 \end{pmatrix}$$

L'estimation d'un quatrième régime n'a pas pu être obtenue car la matrice hessienne est trop proche d'un état de singularité. Plusieurs reparamétrage fins de l'algorithme d'optimisation étaient sans succès. On note au passage la sensibilité de l'algorithme EM (Expectation Maximization, dû à Dempster 77) aux valeurs initiales : certaines valeurs conduisent le processus à converger vers des optima locaux.

Les graphiques sur la figure 2. montrent l'évolution des probabilités (lissées) pour chaque régime. Ces probabilités donnent une indication sur l'occurrence d'appartenir à un régime donné de chaque observation. L'alternance et la persistance de

chaque régime sont clairement mises en évidence. On constate qu'un régime opère lorsque sa probabilité est maximale tandis que celles des autres régimes sont au plus bas.

Dans le modèle estimé retenu, on obtient  $\hat{\lambda} = 0.90$ . Le processus engendre donc un mouvement relativement persistant dans la volatilité du CAC40. Ainsi, pour les données journalières considérées, à raison de 5 jours ouvrables par semaine, tout choc dans la série continuera à avoir des conséquences non négligeables pendant  $((.90)^2 \cdot 2 = 0.09)$  9 jours plus tard.

Les performances prévisionnelles sont données par les graphiques de la figure 3. Les intervalles de confiance  $\pm 2\sigma_t$  montrent les périodes de forte volatilité de la série. Le choc de 1987 est nettement visible. On peut remarquer la nette amélioration des performances prévisionnelles du modèle t-SWARCH-L(3,2) par rapport aux modèles GARCH(1,1) et ARCH(2,2).

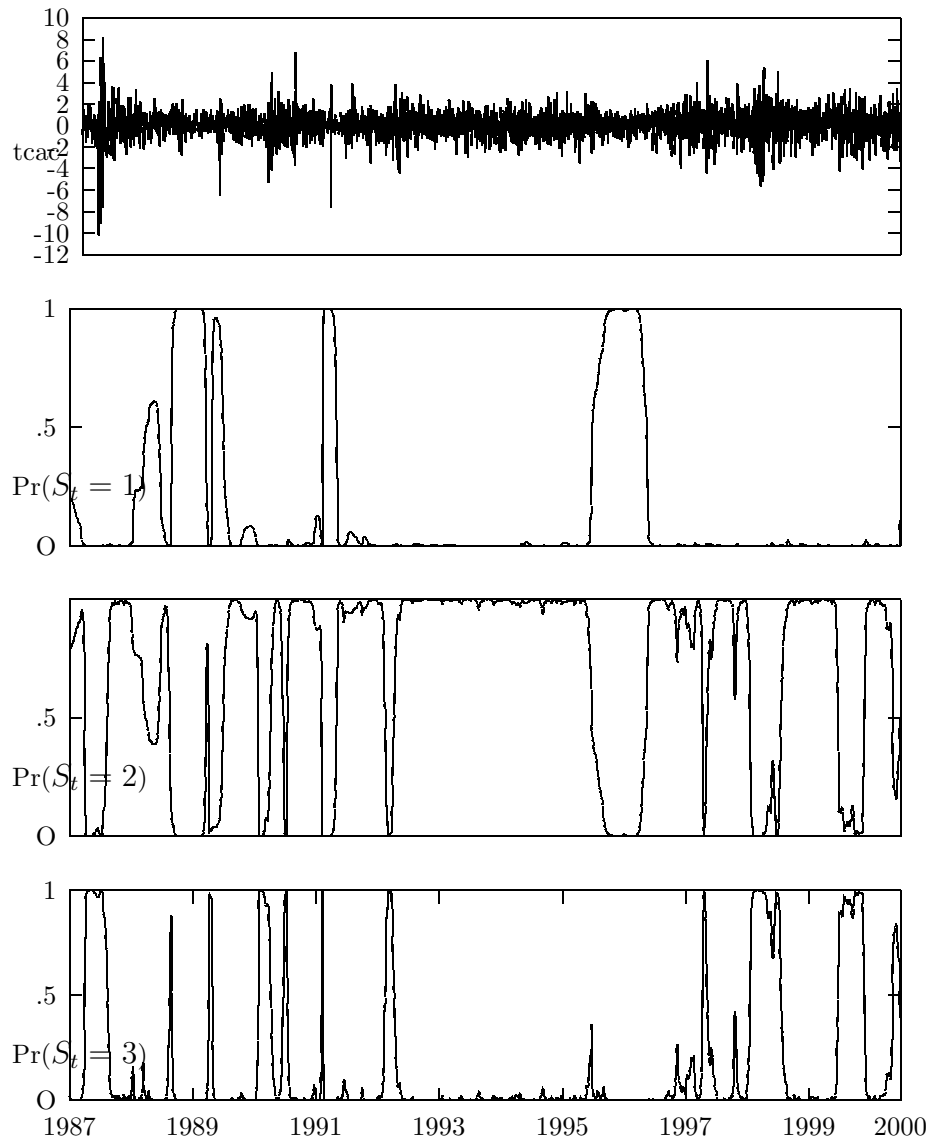


FIG. 2 – Evolution des probabilités “lissées” pour les trois régimes

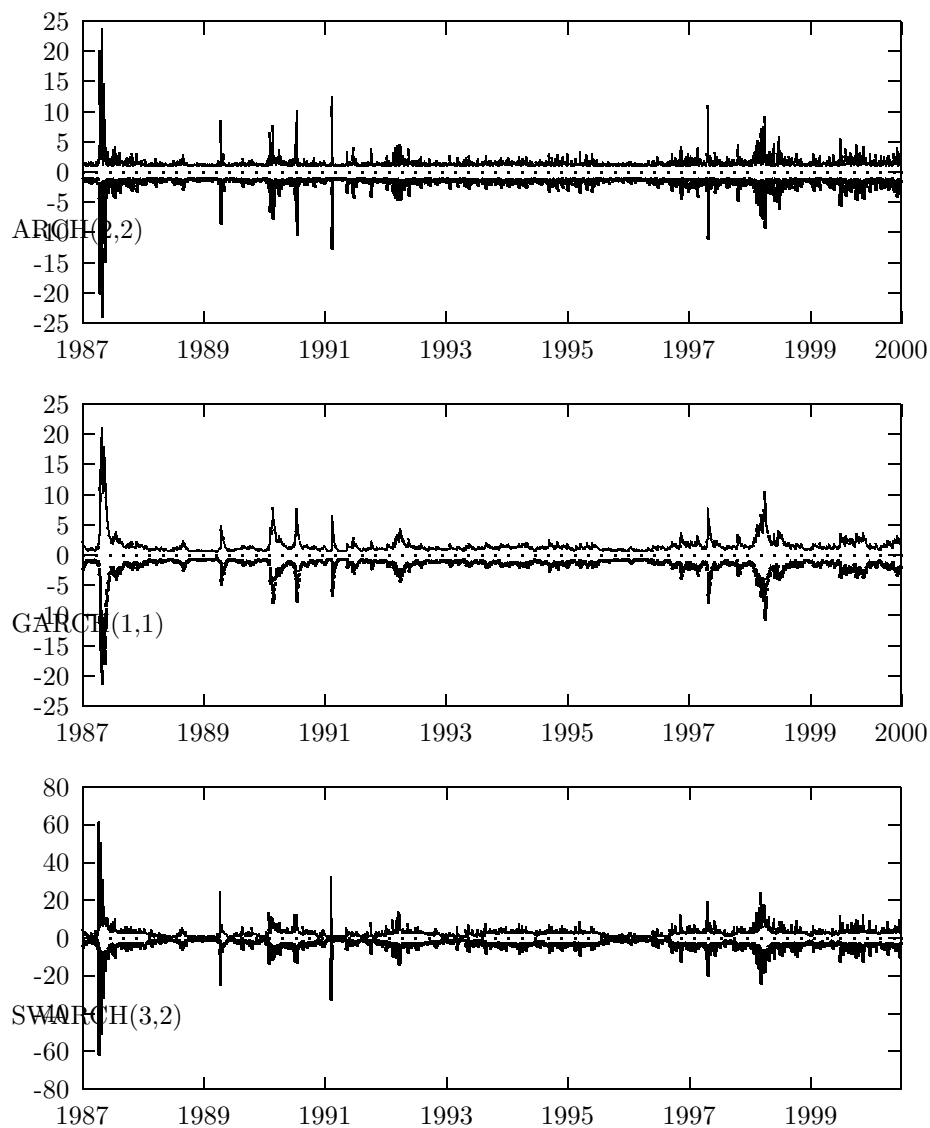


FIG. 3 – Comparaison des intervalles de confiance à 95% pour les 3 spécifications retenues. Les courbes représentent  $\pm 2\hat{\sigma}_t$  pour chaque modèle. Le  $\hat{\sigma}_t$  est obtenu de l'équation de la variance. Certaines spécifications donnent des prévisions plus pauvres que la spécification correspondante avec variance constante.

## 5 Conclusion

Les modèles SWARCH donnent des ajustements plus adéquats de la série du CAC40 étudiée, et par conséquent fournissent de meilleures performances prévisionnelles. Certes, la série étudiée est caractérisée par plus que trois régimes couvrant la période étudiée. Les périodes de décroissance correspondent à plus de volatilité que les périodes de croissance. La distribution de Student s'avère plus appropriée que la distribution Normale, et cela dans les deux configurations étudiées : dans le cas d'un ajustement sur toute la période par un GARCH, ou dans le cas d'un SWARCH à trois régimes.

Les modèles SWARCH avec changements structurels Markoviens discrets, constituent un bon compromis entre les modèles à paramètres aléatoires, difficiles à manier, et la classe de modèles SCAR (suddenly Changing AutoRegressive), proposée par Tyssdal et Tjøstheim (1988), et dans lesquels le coefficient autoregressif d'un processus AR(1) est exprimé par une chaîne de Markov.

Il est, toujours, discutable de savoir si les dynamiques de long terme des variables économiques sont déterminées par l'accumulation des chocs de long terme ou sont plutôt, le résultat d'évènements exceptionnels qui ont des effets permanents, et qui peuvent être exogènes dans leur nature.

Une approche, assez récente, est proposée par une classe de modèles (Granger et Teräsvirta, 1993), où l'on suppose que la nature intrinsèque des relations entre les phénomènes économiques, est non linéaire, et que la modélisation linéaire n'est qu'une approximation locale constante, en besoin de réactualisations et de modifications périodiques.

## Références

- [1] Akaike, H (1974), *A New Look at the Statistical Model Identification*, I.E.E.E Transaction on Automatic Control, 19(6), 716-723.
- [2] Baillie, R.T, and R.P DeGennaro (1990), *Stock Returns and Volatility*, Journal of Financial and Quantitative Analysis, 25(2), 203-214.
- [3] Bera, A.K, and S.Kim (1996), *Testing Constancy of Correlation with an Application to international Equity returns*, CIBER Working Paper 96-107.
- [4] Brendt, E.K, H.B Hall, R.E Hall, and J.A Hausman (1974), *Estimation and Inference in Nonlinear Structural Models*, Annals of Economic and Social Measurement, 4, 653-666.
- [5] Bertero, E, and C.Mayer (1990), *Structure and Performance : Global Interdependence of Stock markets around the Crach 1987*, European Economic Review, 34(6) 1115-1180.
- [6] Black, F (1976) , *Studies in Stock Price Volatility Changes*, Proceedings of the 1976 Business meeting of the Business and economic Statistics Section, American Statistical Association.
- [7] Bollerslev, T.(1986), *Generalized autoregressive Conditional Heteroskedasticity*, Journal of Econometrics, 31(3), 307-327.
- [8] Bollerslev, T.(1987), *A conditionnal Heteroskedastic time Series Models for Speculative Prices and rates of Return*, Reviews Of Economics and Statistics, 69(3), 542-547.
- [9] Bollerslev, T (1990), *Modelling the Coherence in Short-run Nominal Exchange Rates : A Multivariate Generalized ARCH Model*, Review of Economics and Statistics, 72(3), 498-505.
- [10] Bollerslev, T, Ray Chou and Kenneth F. Kroner, 1992, *ARCH modeling in Finance : A Review of the theory and empirical evidence*, Journal of Econometrics 52, 5-59.

- [11] Boyer, B.H, M.S Gibson, and M. Loretan (1997), *Pitfalls in Tests for Changes in Correlation*, Board of Governors of the Federal Reserve System, International Finance Discussion Paper 597.
- [12] Cai, J (1994), *A Markov Model of unconditional variance in ARCH*, Journal of Business and Economic Statistics.
- [13] Connolly, Robert, A 1989, *An examination of the robustness of the weekend effect*, Journal of Financial and Quantitative Analysis 24, 133-169.
- [14] Diebold Francis X, 1986, *Modeling the persistence on conditional variance : A comment*, Econometric Reviews 5, 51-56.
- [15] Diebold Francis X, Steve C Lim and C Jevons Lee, 1993, *A note on conditional Heteroskedasticity in market model*, Journal of Accounting, Auditing and Finance 8, 141-150.
- [16] Engle Robert F 1982, *Autoregressive conditionnal Heteroscedasticity with estimates of the variance of United Kingdom inflation*, Econometrica 50, 987-1007.
- [17] Engle Robert F 1991, *Statistical models of financial volatility*, Mimeo ( University of California San Dieg, CA).
- [18] Engle Robert F and Chowdhury Mustafa, 1992, *Implied ARCH models from options prices*, Journal of Econometrics 52, 289-311.
- [19] Engle Robert F and Victor K Ng 1991, *Measuring and testing the impact of news on volatility*, Mimeo, (University of California San Dieg, CA).
- [20] French Mark W and Daniel F Sichel ,1993, *Cycliacal patterns in the variance of economic activity*, Journal of Business and Economic Statistics 11, 113-119.
- [21] Friedman, Benjamin M. 1992, *Big shocks and little shocks : Security returns with nonlinear persistense of volatility*, Mimeo, ( Harvard University, Cambridge MA).

- [22] Friedman, Benjamin M. and David I Laibson, 1989, *Economic implication of extraordinary movements in stock prices*, Brookings papers of Economic Activity 2, 137-189.
- [23] Glosten, Lawrence R. Ravi Jagannathan and David Runkle, 1989, *Relationship between the expected value and the volatility of the nominal excess return on stocks*, Mimeo, ( Northwestern University, Evanston, IL).
- [24] Gouriéroux Christian and Alain Monfort 1992, *Qualitative threshold ARCH models*, Journal of Econometrics 52, 159-199.
- [25] Hamilton, James D. 1989, *A new approach to the economic analysis of non-stationary time series and the business cycle*, Econometrica 57, 357-384.
- [26] Hamilton, James D. 1994, *times series analysis*, (Princeton University Press, Princeton, NJ).
- [27] Hansen, Bruce E. 1991, *Inference when a nuisance parameter is not identified under the null hypothesis*, Mimeo, ( University of Rochester N.Y).
- [28] Hansen, Bruce E. 1992, *The likelihood ratio test under non-standard conditions : testing the Markov trend model of GNP*, Journal of Applied Econometrics 7, S61-S82.
- [29] Kim Chang-Jin 1994, *dynamic linear models with Markov-switching*, Journal of Econometrics 60, 1-22.
- [30] Lamoureux Christopher G. and William D.Lastrapes, 1990, *Persistence in variance, structural change and the GARCH model*, Journal of business and Economic Statistics 8, 225-234.
- [31] Lamoureux Christopher G. and William D.Lastrapes 1993, *Forecasting stock return variance : Toward an understanding of stochastic implied volatilities*, Review of Financial Studies 5, 293-326.
- [32] Mandelbrot B.(1963), *The variation of Certain Prices*, J.Business, 394-419.
- [33] Morgan Alison and Ieuan Morgan, 1987, *Measurement of abnormal returns from small firms*, journal of Business and Economic Statistics 5, 121-129.

- [34] Nelson ,Daniel. 1991,*Conditional Heteroskedasticity in asset returns : A new approach*, Econometrica 59,347-370.
- [35] Pagan, Adrian and G.William Schwert, 1990,*Alternative models for conditional stock volatility*, Journal of econometrics 45,267-290.
- [36] Pagan, Adrian and Aman Ullah, 1988,*The Econometric analysis of models with risk terms*, Journal of applied Econometrics 3, 87-105.
- [37] Perron, Pierre, 1989 *the great crach,the oil price shock, and tue unit root hypothesis*, Econometrica 57, 1361-1401.
- [38] Schwarz Gideaon,1978,*Estimating the dimension of model*, Annals of statistics 6, 461-464.
- [39] Schwarz G William and Paul J Seguin, 1990,*Heteroskedasticity in stock returns*, Journal of Finance 45, 1129-1155.
- [40] West, kenneth D Hali J Edison, and Dongchul Cho, 1993,*A utility based comparison of some models of foreing exchange volatility*, journal of international Economics 35,23-46.

## Annexes

### Données

Moyenne	2555.957
Médiane	1966.650
Maximum	6922.330
Minimum	893.820
Ecart-Type	1386.883
Skewness	1.611924
Kurtosis	4.626364
Jarque-Bera	1832.961
	(0.00000)

Série du CAC40 (3376 observations)

### Procédure d'évaluation de la fonction de vraisemblance

Une étape générique parmi le nombre d'itérations qui permettent de calculer la fonction de vraisemblance, utilise comme argument la probabilité conditionnelle :

$$p(s_{(q)}|y_{(-3)}) \quad (1)$$

où  $s_{(q)} = (s_t, s_{t-1}, s_{t-2}, \dots, s_{t-q})$  et  $y_{(-3)} = (y_t, y_{t-1}, y_{t-2}, \dots, y_{-3})$ . Chacun des  $K^{q+1}$  nombres représentés par (1) est multiplié par  $p_{s_t \cdot s_{t+1}}$  et par

$$f(y_{t+1}|s_{t+1}, s_t, \dots, s_{t-q+1}, y_t, y_{t-1}, \dots, y_{t-q+1})$$

pour donner les  $K^{q+2}$  nombres séparés :

$$p(s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q}, y_{t+1}|y_t, y_{t-1}, \dots, y_{-3}) \quad (2)$$

Pour une spécification Gaussienne, les calculs précédents utilisent :

$$f(y_{t+1}|s_{(q+1)}, y_{(q+1)}) = \frac{1}{\sqrt{2\pi}\sigma_{t+1}(s_{(q+1)})} \cdot \exp\left[\frac{-(y_{t+1} - \alpha - \phi y_t)^2}{2\sigma_{t+1}^2(s_{(q+1)})}\right]$$

où  $s_{(q+1)} = (s_{t+1}, s_t, \dots, s_{t-q+1})$  et  $y_{(q+1)} = (y_t, y_{t-1}, \dots, y_{t-q+1})$  et où  $\sigma_t^2(s_{(q+1)})$  est donnée par :

$$\begin{aligned} E(u_t^2 | s_t, s_{t-1}, \dots, s_{t-q}, u_{t-1}, u_{t-2}, \dots, u_{t-q}) \\ &= g_{s_t} \left( a_0 + \sum_{i=1}^q a_i \frac{u_{t-i}^2}{g_{s_{t-i}}} \right) + \xi \cdot d_{t-1} \cdot \frac{u_{t-1}^2}{g_{s_{t-1}}} \\ &\equiv \sigma_t^2(s_t, s_{t-1}, \dots, s_{t-q}) \end{aligned}$$

avec  $u_t \equiv y_t - \alpha - \phi y_{t-1}$ .

Pour la version du modèle avec la distribution de Student, nous utilisons :

$$f(y_{t+1} | s_{(q+1)}, y_{(q+1)}) = \frac{\Gamma[(\nu + 1)/2]}{\Gamma(\nu/2) \cdot \sqrt{\pi} \cdot \sqrt{\nu - 2} \cdot \sigma_{t+1}(s_{(q+1)})} \times \left[ 1 + \frac{(y_{t+1} - \alpha - \phi y_t)^2}{(\nu - 2) \cdot \sigma_{t+1}^2(s_{(q+1)})} \right]^{-(\nu+1)/2}$$

Les nombres donnés par (2) sont les sommes des densités conditionnelles de  $y_{t+1}$  :

$$f(y_{t+1} | y_{(-3)}) = \sum_{s_{t+1}=1}^K \sum_{s_t=1}^K \dots \sum_{s_{t-q}=1}^K p(s_{(q)}, y_{t+1} | y_{(-3)}) \quad (3)$$

A partir de cette expression, on peut calculer la fonction du log-vraisemblance de l'échantillon. Pour tout  $s_{(q+1)}$ , les nombres donnés dans (2) sont sommés sur les  $K$  valeurs possibles pour  $s_{t-q}$  et le résultat est alors normalisé par (3) afin d'obtenir  $p(s_{(q+1)} | y_{(-3)})$  qui est l'argument de l'étape  $t + 1$  des itérations. Ces itérations démarrent de  $p(s_0, s_{-1}, \dots, s_{-q})$  prises égales aux probabilités ergodiques, données par :

$$s_{t+1} = P \cdot s_t + v_{t+1}$$

où  $v_{t+1} = s_{t+1} - E(s_{t+1} | s_t, s_{t-1}, \dots)$ , et  $P$  est la matrice des probabilités de transition.

L'algorithme de calcul des probabilités lissées  $p(s_t | y_T, y_{T-1}, \dots, y_{-3})$  est décrit par la relation :

$$\hat{s}_{t|T} = \hat{s}_{t|t} \odot \left[ P' \cdot (s_{t+1|T} (\div) s_{t+1}) \right]$$

où  $P'$  est la transposée de la matrice des probabilités de transition et  $\odot$  et  $(\div)$  sont respectivement le produit et la division matriciels élément par élément.

## Table des matières

<b>1</b>	<b>Les modèles ARCH</b>	<b>3</b>
1.1	Le modèle ARCH(p) linéaire : . . . . .	3
1.2	Le modèle GARCH(p,q) linéaire : . . . . .	5
1.3	Modèle ARCH non linéaire et non paramétrique : . . . . .	6
1.4	Méthode d'estimation : . . . . .	7
<b>2</b>	<b>Ajustement empirique du CAC40</b>	<b>10</b>
2.1	Les données : . . . . .	10
2.2	Les statistiques : . . . . .	10
2.3	Les performances prévisionnelles . . . . .	14
2.4	La persistance . . . . .	15
<b>3</b>	<b>Spécification ARCH et changements de régime</b>	<b>15</b>
3.1	Le cadre théorique . . . . .	16
<b>4</b>	<b>Procédure d'estimation</b>	<b>18</b>
4.1	Le problème des prévisions . . . . .	19
4.2	Le problème de la persistance . . . . .	21
4.3	Résultats empiriques . . . . .	22
<b>5</b>	<b>Conclusion</b>	<b>29</b>
	PLAN	