A1.91
G
896

Tests GMM pour caractériser le biais de simultanéité de la persistence de volatilité

Rapport de Stage

en vue de l'obtention du diplôme de

Maîtrise de Finance Mathématique et Computationnelle

Miloud Boudiaf

Sous la direction des Professeurs

Bryan Campbell & Éric Renault

Centre de Documentation
Dép. de sciences économiques
Université de Montréal
C. P. 6128, Succ. "A"
Montréal, Qué., Canada, H3C 3J7

Contents

1 Introduction **ii**

2 Modèles de l'hétéroscédasticité **1**

3 Méthode généralisée des moments GMM **2**

3.1 Estimation d'un GARCH(1,1) 3

4 De la modélisation en forme réduite à la modélisation structurelle **5**

5 Dynamique de α_t **7**

6 Méthode **9**

6.1 Un modèle conforme à SR-SARV(1) 9

6.2 Test de suridentification 9

7 Expérience **11**

7.1 Données simulées 11

7.2 Données réelles 12

8 Résultats et commentaires **13**

9 Conclusion **15**

A Codes Splus **27**

B Codes C **34**

1 Introduction

L'importance de la volatilité des rendements financiers, vue comme un processus lui même volatile n'est plus à démontrer. Depuis quelques années la modélisation des processus de volatilité est au coeur d'une bonne partie de la littérature économétrique et financière. Un exemple type de modélisation auquel un grand nombre d'études a été consacré, est le modèle autoregressif de l'hétéroscédasticité conditionnelle (ARCH), introduit dans [EN], et sa généralisation (GARCH) par [BO].

Dans leur article [AR] [1], les auteurs montrent l'importance d'une formulation dite "structurelle" des processus de volatilité stochastique par opposition à une formulation dite en "forme réduite" tels les GARCH. La formulation "structurelle" est consacrée aux modèles dynamiques spécifiés par une identification de variables d'intérêt, non pas comme fonctions directes des observables, mais comme caractéristiques de certaines lois conditionnelles d'intérêt [2].

Quand la variable d'intérêt est la variance condionnelle des rendements d'actifs financiers, un de ses aspects importants dont devrait rendre compte toute identification est le phénomène de la "persistence de volatilité". Partant d'une formulation "structurelle" de la volatilité, des plus simples, qui rend compte du phénomène de persistance, à savoir un modèle autogressif d'ordre 1 de la variance condionnelle des rendements, [AR] discutent l'estimation par GMM du paramètre d'intérêt, le coefficient de la regression. Cette estimation exploite des conditions de moments induits par le modèle de la volatilité choisi, et l'idée selon laquelle les rendements (innovations) au carré fournissent un "estimateur sans biais du facteur de volatilité". Dans le cas des ARCH les conditions de moments utilisées

[1] Dont quelques développements sont à la base de ce travail de stage

[2] Ceci est une paraphrase de [AR]

induisent souvent un biais négatif (biais de simultanéité) du paramètre d'intérêt, voir e.g. les expériences Monte Carlo de [MR].

C'est pour illustrer ce fait que [AR] partent d'une formulation "structurelle" qui généralise le modèle GARCH(p,q). Ils montrent que la caractérisation du biais de simultanéité, revient à caractériser le coefficient de la régression affine condionnelle au passé de la variance conditionnelle sur les rendements au carré. L'ultime raison de ce travail est une contribution pour caractériser ce coefficient.

Cette dernière motivation ayant exigé, en fait le moins d'énergie, le travail de préparation des expériences notamment, constitue la majeur partie de ce stage. Aussi, nous en rendons compte dans ce rapport. La section 2 rappelle les modèles autorégressifs de l'hétéroscédasticité, en particulier le modèle GARCH(1,1) souvent invoqué dans le texte. La section 3 récapitule la méthode des moments généralisés ou GMM et son utilisation pour l'estimation des paramètres d'un GARCH(1,1).

2 Modèles de l'hétéroscédasticité

Les modèles autorégressifs de l'hétéroscédasticité conditionnelle de la volatilité des rendements financiers (ARCH) ont été introduits par Engle(1982) ([EN]) pour tenir compte de deux decouvertes. Le phénomène dit "volatility clustering" qui se résume en la persistance de la volatilité (de grandes amplitudes de la variance sont suivies par de grandes amplitudes et des petites par des petites) révélée par la présence de corrélation dans les carrés des résidus des taux de rendements dans les données financières. Et le phénomène des "fat tails" ou queues épaisses des distributions des rendements. En notant ϵ_t le (résidus) taux de rendement et σ^2_{t-1} sa variance condiotionnelle à l'information du passé $I(t-1)$, Engle introduit son modèle ARCH(q) en exprimant linéairement σ^2_{t-1} en fonction des q "lag"de ϵ^2_t

$$\sigma^2_t = \omega + \sum_{i=1}^{q} \alpha_i \epsilon^2_{t-i} \tag{1}$$

avec $\omega, \alpha_i \geq 0$. On remarque, en posant $\nu_t := \epsilon^2_t - \sigma^2_{t-1}$ que ϵ^2_t suit un modèle autorégressif d'ordre q (AR(q)) avec résidus ν_t, ce qui peut expliquer la persistance de la volatilité.

L'utilisation des modèles ARCH a montré qu'il fallait utiliser de longs AR pour expliquer les données. C'est pour cette raison que Bollerslev (1986) ([BO]) généralisa le modèle ARCH(p) par GARCH(p,q) qui s'écrit:

$$\sigma^2_t = \omega + \sum_{i=1}^{q} \alpha_i \epsilon^2_{t-i} + \sum_{i=1}^{p} \beta_i \sigma^2_{t-i} \tag{2}$$

avec $\omega, \alpha_i, \beta_i \geq 0$. De même que pour les ARCH on peut montrer que $\epsilon^2_t \sim ARMA(max(p,q), q)$. En pratique il est souvent suffisant d'estimer un modèle GARCH(1,1).

Une multitude de modèles de la volatilité ont ensuite été introduits. Nous ferons allusion dans ce travail au modèle GARCH asymétrique, introduit par Engle et Ng (1993) ([])pour expliquer l'impact des nouvelles sur la volatilité. C'est défini par l'équation (2) où on remplace ϵ^2_{t-i} par $(\epsilon_{t-i} + \delta_i)^2$.

3 Méthode généralisée des moments GMM

Nous suivons [HA] pour ce survol des GMM qui reprend lui même les développements de Hansen (1982) ([HL]). Soient w_t un vecteur $(h \times 1)$ de variables aléatoires observées au temps t, θ un vecteur $(a \times 1)$ de paramètres à estimer et $\mathbf{h}(\theta, w_t)$ une fonction de $\mathbb{R}^a \times \mathbb{R}^h$ à valeurs dans $\mathbb{R}^r$ telle que la vraie valeur de θ, θ_0 vérifie

$$E\{\mathbf{h}(\theta_0, w_t)\} = 0 \tag{3}$$

Cette équation vectorielle représente les r conditions d'orthogonalité ou conditions de moments.

Soient $\mathcal{Y}_T \equiv (w_1', \cdots, w_T')$ un échantillon de taille T de toutes les observations et $\mathbf{g}(\theta, \mathcal{Y}_T)$ la moyenne échantillonale de $\mathbf{h}(\theta, w_t)$:

$$\mathbf{g}(\theta, \mathcal{Y}_T) = (1/T) \sum_{t=1}^{T} \mathbf{h}(\theta, w_t) \tag{4}$$

g est encore appellé le moment généralisé.

L'idée de GMM est de choisir θ tel que le moment échantillonal $\mathbf{g}(\theta, \mathcal{Y}_T)$ soit aussi proche que possible du moment théorique égal à zéro. L'importance des conditions d'orthogonalité n'étant pas la même, des poids y sont affectés. L'estimateur GMM $\hat{\theta}$ est défini par la valeur de θ qui minimise la fonction scalaire:

$$Q(\theta, \mathcal{Y}_T) = [\mathbf{g}(\theta, \mathcal{Y}_T)]' W_T [\mathbf{g}(\theta, \mathcal{Y}_T)], \tag{5}$$

où $\{W_T\}_{T=1}^{\infty}$ est une suite de matrices de poids définies positives et qui peuvent dépendre des données $\mathcal{Y}_T$. En pratique la minimisation de la fonction objective $Q(\theta, \mathcal{Y}_T)$ se fait numériquement par des méthodes d'optimisation classiques. Pour ce faire nous disposons de bons optimiseurs dans tous les logiciels (qui se respectent) dans le marché. Il n'est donc pas question d'implanter encore ce genre de procédures d'optimisation. Un point important, cependant dans les GMM est le choix de la matrice de poids. Cette dernière est souvent fonction

(non connue) des paramètres θ et la question qui se pose est quelle est la matrice optimale. En fait étant donné un estimateur $\hat{\theta}$ plusieurs estimateurs $\hat{S}_T(\hat{\theta})$ de $[W_T]^{-1}$ ont été proposés dans la littérature (basé sur l'hypothèse que le processus $\{\mathbf{h}(\theta_0, w_t)\}_{t=-\infty}^{\infty}$ est stationnaire). La procédure est souvent itérative pour estimer $\hat{\theta}_{GMM}$. À partir d'une matrice de poids arbitraire un estimateur initial $\hat{\theta}_T^{(0)}$ est obtenu (en optimisant (5)). Ce dernier est utilisé pour obtenir un estimateur initial $\hat{S}_T^{(0)} = \hat{S}_T(\hat{\theta}_T^{(0)})$ qui va servir à optimiser (5) avec $W_T = \hat{S}_T^{(0)}$ et obtenir un nouvel estimateur GMM $\hat{\theta}_T^{(1)}$. Cette procédure est répétée jusqu'à ce que $\hat{\theta}_T^{(j)} \cong \hat{\theta}_T^{(j)}$.

3.1 Estimation d'un GARCH(1,1)

Soit à estimer par GMM, un modèle GARCH(1,1) sur un échantillon (de taille T) observé à partir d'un processus ϵ_t de génération de données:

$$\epsilon_t = \sigma_{t-1} u_t, \quad u_t \sim iid(0,1) \tag{6}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_t^2 + \beta\sigma_{t-1}^2 \tag{7}$$

où σ_{t-1}^2 est la variance de ϵ_t conditionnelle à l'information passée $I(t-1)$. L'objectif est d'estimer $\theta_0 \in \Theta$ où $\theta_0' = (\omega, \alpha, \beta)$ et Θ est une espace de paramètres compact. Pour cela définissons le vecteur

$$\mathbf{r}_t = [\epsilon_t, (\epsilon_t^2 - \sigma_{t-1}^2)]' \tag{8}$$

Les conditions d'orthogonalité s'ecrivent comme $E\mathbf{r}_t = 0$.

Pour une fonction de variable instrumentale (que l'on se donne) posons

$$\mathbf{g}_t = \mathbf{F}_t'\mathbf{r}_t \tag{9}$$

alors l'estimateur GMM du vecteur de paramètres θ est la solution de

$$\text{argmin}_{\theta\in\Theta} \left[T^{-1}\sum_{t=1}^{T}\mathbf{g}_t\right]' W_T \left[T^{-1}\sum_{t=1}^{T}\mathbf{g}_t\right] \tag{10}$$

pour un choix approprié de la matrice de poids W_T.

Pour estimer des modèles GARCH(1,1) par GMM dans nos expériences nous avons suivi [SJ] qui propose la fonction objective à minimiser suivante:

$$Q_T = T^{-1}\left[\sum_{t=1}^{T} \mathbf{g}_t\right]' \left(\sum_{t=1}^{T} \Lambda_t\right)^{-1} \left[\sum_{t=1}^{T} \mathbf{g}_t\right] \tag{11}$$

où $\Lambda_t = \mathbf{g}_t\mathbf{g}_t'$ est une matrice de poids qui dépend des paramètres. La procédure n'est donc pas itérative, d'où notre intérêt pour cette méthode. Les éléments du moment généralisé sont donnés par

$$\begin{aligned} g_{it} &= \frac{1}{\Delta}\left(\frac{\partial\sigma_{t-1}^2}{\partial\theta_i}\right)\sigma_{t-1}^{-2}\left[\left(\frac{\epsilon_t}{\sigma_{t-1}}\right)v_3 - \left(\left(\frac{\epsilon_t^2}{\sigma_{t-1}^2}\right) - 1\right)\right] \\ \Lambda_{ijt} &= \frac{1}{\Delta^2}\left(\frac{\partial\sigma_{t-1}^2}{\partial\theta_i}\right)\sigma_{t-1}^{-2}\left[\left(\frac{\epsilon_t}{\sigma_{t-1}}\right)v_3 - \left(\left(\frac{\epsilon_t^2}{\sigma_{t-1}^2}\right) - 1\right)\right]^2 \sigma_{t-1}^{-2}\left(\frac{\partial\sigma_{t-1}^2}{\partial\theta_i}\right) \end{aligned}$$

où $v_l = E[u_t^l]$ et $\Delta = (v_4 - 1) - v_3^2$.

4 De la modélisation en forme réduite à la modélisation structurelle

La modélisation en forme réduite de la volatilité, terminologie non standard utilisée par [AR], signifie que la dynamique de la volatilité est définie uniquement à partir de caractéristiques deterministes en fonction des innovations de rendements ϵ_t. L'exemple type de modèles en forme réduite de la volatilité est le modèle autoregressif de l'hétéroscédasticité conditionnelle ARCH. [AR] discutent les risques de modèle de la volatilité, mais en particulier de la formulation en forme réduite. Parmi ces risques, ils montrent celui de l'estimation du paramètre de persistance en utilisant l'auto-corrélation empiriques des ϵ_t^2 dont le processus ARMA est induit par la formulation en forme réduite. Or cet estimateur n'en est pas un de l'auto-corrélation théorique dans les cas les plus couraments estimés où le moment d'ordre quatre des rendements est infini. Après cela la considération d'une modélisation structurelle, c'est-à-dire d'une dynamique de la volatilité non deterministe, trouve naturellement son intérêt. Un des modèles les plus simples qui rend compte du phénomène de persistance de la variance condionnelle σ_t^2 est le modèle auto-régressif :

$$\sigma_{t+1}^2 = \omega + \gamma\sigma_t^2 + \nu_{t+1}, \quad 0 < \gamma < 1 \tag{12}$$

Écrivons que les rendements (innovations) au carré fournissent un "estimateur sans biais du facteur de volatilité":

$$\epsilon_{t+1}^2 = \sigma_t^2 + \eta_{t+1}, \qquad E_t(\eta_{t+1}) = 0 \tag{13}$$

En intégrant (12) dans (13) on obtient:

$$\epsilon_{t+1}^2 = \omega + \gamma\epsilon_t^2 + +\nu_t - \gamma\eta_t + \eta_{t+1} \tag{14}$$

De l'equation (14) on voit qu'on ne peut espérer, en général que:

$$E(\epsilon_{t+1}^2 - \omega + \gamma\epsilon_t^2 | I(t)) = 0 \tag{15}$$

à moins de supposer

$$\nu_t = \gamma\eta_t, \qquad \forall t$$

Cette dernière condition (étrange postulat comme expliqué dans [AR]) est pourtant necessaire et suffisante pour que l'equation (12) se réduise à un modèle ARCH(1).

Alors que l'equation structurelle (12) induit que

$$E(\epsilon_{t+1}^2 - \omega + \gamma\epsilon_t^2 | I(t-1)) = 0 \tag{16}$$

ce sont les conditions de moments (15) qui sont utilisées, en général, dans l'estimation du paramètre d'intérêt γ par GMM. Cela induit souvent un biais "de simultanéité". [AR] montrent que le signe de ce biais est celui de $(\alpha_t - \gamma)$, où α_t est le coefficient de la regression affine, conditionnelle à $I(t)$ de σ_{t+1}^2 sur ϵ_{t+1}^2.

5 Dynamique de α_t

D'après la section précédente la caractérisation du coefficient α_t trouve son intérêt par celui que porte le paramètre γ qui mesure la persistence de la volatilité. On se pose donc le problème de caractériser la dynamique de α_t. Pour ce faire on partira naturellement de sa définition :

$$\alpha_t = \frac{Cov_t[\sigma_{t+1}^2, \epsilon_{t+1}^2]}{Var_t[\epsilon_{t+1}^2]} \tag{17}$$

Notons par κ_t la kurtosis conditionnelle des innovations de rendements normalisées $u_t := \frac{\epsilon_t}{\sigma_{t-1}}$. On a

$$Var_t[\epsilon_{t+1}^2] = (\kappa_t - 1)\sigma_t^4 \tag{18}$$

en effet, on a par définition

$$\begin{aligned} Var_t[\epsilon_{t+1}^2] &= E_t(\epsilon_{t+1}^4) - \left(E_t\epsilon_{t+1}^2\right)^2 \\ &= \sigma_t^4 E_t(u_{t+1}^4) - \sigma_t^4 \left(E_t u_{t+1}^2\right)^2, \end{aligned} \tag{19}$$

ensuite u_t est de moyenne condionnelle nulle et de variance condionnelle constante égale à 1 par construction. Donc on a bien que la kurtosis condionnelle de u_t est égale à $E_t(u_{t+1}^4)$ qui, remplacée dans (19) par κ_t, prouve (18). Reportons cette dernière dans (17) nous aurons

$$\alpha_t = \frac{Cov_t[\sigma_{t+1}^2, \epsilon_{t+1}^2]}{(\kappa_t - 1)\sigma_t^4} \tag{20}$$

Cette expression de α_t suggère de supposer (à défault de mieux et dans un premier temps) κ_t constante. C'est une hypothèse suggérée par le cas des GARCH. En effet dans ce cas nous avons $u_t \sim iid(0,1)$ donc $E_t[u_{t+1}^4] = E[u_{t+1}^4]$.

Le problème se réduit alors à la caractérisation de la dynamique de

$$\frac{Cov_t[\sigma_{t+1}^2, \epsilon_{t+1}^2]}{\sigma_t^2}$$

On se pose la question, par exemple, de savoir s'il existe une constante c telle que

$$Cov_t[\sigma_{t+1}^2, \epsilon_{t+1}^2] = c\sigma_t^4 \quad ? \tag{21}$$

En écrivant la covariance comme l'espérance du produit moins le produit des espérances, puis en remplaçant σ_{t+1}^2 dans (21) par sa définition (variance conditionnelle à $I(t+1)$ de ϵ_{t+2}) et utilisant le théorème des espérances itérées on arrive à:

$$E_t[\epsilon_{t+2}^2 \epsilon_{t+1}^2] = E_t[\epsilon_{t+2}^2] E_t[\epsilon_{t+1}^2] + c\sigma_t^4 \tag{22}$$

En réutilisant les mêmes arguments et aussi en utilisant l'équation (12), le premier terme à droite dans (22) peut être remplacé par

$$\omega + \gamma\sigma_t^2$$

et le deuxième terme par σ_t^2.

En appellant encore c la somme $\gamma + c$ on peut réecrire notre problème comme:

Existe-t-il une constante c tel que

$$E_t[\epsilon_{t+2}^2 \epsilon_{t+1}^2] = \omega\sigma_t^2 + c\sigma_t^4 \tag{23}$$

6 Méthode

Nous adoptons la démarche suivante pour répondre à la question posée dans l'équation (23):

1– Estimation de modèle de volatilité conforme à SR-SARV(1)

2– Estimation et test GMM sur l'existence de la constante c.

6.1 Un modèle conforme à SR-SARV(1)

Les modèles SR-SARV(p) introduits par [MR] (nous renvoyons à cet article pour la définition) trouvent leur intérêt parceque ce sont des formulations structurelles de la volatilité. On a, en particulier, un processus d'innovations de rendements ϵ_t dont la variance conditionnelle au passé $I(t-1)$, σ^2_{t-1} vérifie l'equation (12), est un processus SR-SARV(1). Dans la perspective de ce rapport de stage, seuls des modèls GARCH(1,1), qui en sont des sous-classes, ont été considérés. En fait il ne sera question d'estimation d'un GARCH(1,1) que dans le cas où nous considérons des données réelles où nous ne disposons pas du nécessaire (paramètre ω et série des "non-observables" σ^2_t) pour répondre à la question (23). Nous y répondons en effectuant un test de "suridentification".

6.2 Test de suridentification

Rappel Considérons le modèle linéaire

$$y_t = cz_t + u_t \tag{24}$$

où la variable explicative est endogène $E(z_t u_t) \neq 0$ et soit x_t un vecteur $(r \times 1), r > 1$ de variables explicatives corrélés avec z_t mais non avec u_t.

Alors un estimateur $\hat{c}$ par variables instrumentales, x_t est en fait un estimateur GMM de c, où les conditions de moments sont données par

$$E[x_t(y_t - \hat{c}z_t)] = 0 \tag{25}$$

La procédure d'estimation que nous utiliserons est itérative telle qu'expliquée dans la section 3. Cela revient à minimiser la fonction objective donnée par

$$T[\mathbf{g}(\hat{c}; \mathcal{Y}_T)]' \hat{S}_T^{-1} [\mathbf{g}(\hat{c}; \mathcal{Y}_T)] \tag{26}$$

où $\mathcal{Y}_T$ représente un un échantillon de taille T de toutes les observations

$$(y_1, z_1, x_1', \cdots, y_T, z_T, x_T')$$

et

$$\mathbf{g}(\hat{c}; \mathcal{Y}_T) = (1/T) \sum_{t=1}^{T} x_t (y_t - \hat{c} z_t)$$

Nous utiliserons comme estimateur de la matrice de poids l'estimateur de Newey-West:

$$\hat{S}_T = \hat{\Gamma}_{0,T} + \sum_{\nu=1}^{q} \{1 - [\nu/(q+1)]\} (\hat{\Gamma}_{\nu,T} + \hat{\Gamma}'_{\nu,T}) \tag{27}$$

où

$$\hat{\Gamma}_{\nu,T} = (1/T) \sum_{t=\nu+1}^{T} \hat{u}_t \hat{u}_{t-\nu} x_t x_{t-\nu}.$$

Nous savons tester l'hypothèse que ce modèle est bien spécifié. En effet, sous cette hypothèse la statistique

$$T[\mathbf{g}(\hat{c}_{GMM}; \mathcal{Y}_T)]' \hat{S}_T^{-1} [\mathbf{g}(\hat{c}_{GMM}; \mathcal{Y}_T)] \tag{28}$$

est asymptotiquement un $\chi^2(r-1)$ (voir par e.g. [HA]).

Réécrivons maintenant, la condition (d'orthogonalité) (23) sous forme d'une hypothèse

$\mathbf{H_0}$: $\exists\, \mathbf{c}$ tel que

$$E[x_t(\epsilon_{t+2}^2 \epsilon_{t+1}^2 - \omega \epsilon_{t+1}^2 - c\sigma_t^4)] = 0 \quad \forall x_t \in I(t) \tag{29}$$

On est bien dans les conditions d'un test de suridentification.

7 Expérience

Nous avons considéré deux types de données pour mener les expériences. Des données simulées à partir d'un modèle GARCH(1,1) et des données réelles représentant des rendements journaliers d'actifs financiers. Le premier point de la méthode (section6) n'ayant plus d'intérêt dans le cas des données simulées sinon de vérifier nos procédures informatiques. Nous ne donnerons donc que les résultats de l'estimation des GARCH(1,1) des données réelles. Voici comment nous avons procédé dans les deux cas:

7.1 Données simulées

On se donne les paramètres d'un GARCH(1,1), $\Theta = (\omega, \alpha, \beta)$. On simule à partir de là une série de données ϵ_t en utilisant le processus de génération de données décrit par les equations (6) et (7)

On se donne différents instruments $x_t \in I(t)$ et on estime par GMM une constante c telle que (29) soit vrai. On teste (Test de suridentification) alors $\mathbf{H_0}$. Les résultats de ce test sont présentés dans le tableau 2 (tous les résultats seront commentés dans la section 8) . Pour les obtenir, nous avons utilisés les valeurs suivantes des paramètres du GARCH(1,1) ($\omega = 0.01, \alpha = 0.05, \beta = 0.9$). Nous avons simulé des séries de longeurs 1000 et répété le scénario 200 fois.

Remarquons que l'on est dans un cas où la kurtosis conditionnelle des innovations est constante. Aussi pouvons nous réécrire l'hypothèse nulle et tester de façon équivalente, l'hypothèse

$\tilde{\mathbf{H}}_0$: $\exists\, c$ telle que

$$E[x_t(\epsilon_{t+2}^2\epsilon_{t+1}^2 - \omega\epsilon_{t+1}^2 - c\epsilon_{t+1}^4)] = 0 \quad \forall x_t \in I(t) \tag{30}$$

en effet, $E_t[\epsilon_{t+1}^4] = \sigma_t^4 E_t[u_{t+1}^4] = \sigma_t^4\kappa$. Les résultats de ce test sont présentés dans le tableau 4.

7.2 Données réelles

Les données réelles utilisées représentent les log-rendements journaliers, durant les 10 dernières années, des actions de deux compagnies, souvent utilisées pour estimer des modèles GARCH. Il s'agit de IBM et INTEL. Les données ont été téléchargées du site web de Yahoo. Elles couvrent la période du 20 mai 1991 au 20 août 2002, soit des longueurs de 2839 pour IBM et 2838 pour Intel. Les log-rendements r_t ont été calculés à partir des prix à la fermeture P_t (ajustés pour tenir compte, notamment, des dividendes) selon la formule:

$$r_t = \log(\frac{P_t}{P_{t-1}})$$

Le tableau 1 présente quelques statistiques descriptives. Les deux séries présentent des "skewness" négatives, mais elles sont non significatives (variances asymptotiques $= 6/T = 0.002115$). Elles présentent également de forts excés de "kurtosis", surtout dans le cas de Intel (35.701). Les histogrammes dans la figure 9 le confirment aussi. Pour le reste des statistiques les deux séries présentent les mêmes caractéristiques.

Nous appliquons donc notre méthode à ces données. Nous estimons, dans une premiere étape, un modèle GARCH(1,1) sur les log-rendements centrés, *i.e.* $\epsilon_t = r_t - \overline{r_t}$. L'estimation se fait par GMM. Nous avons suivi pour cela [SJ] tel qu'expliqué dans la section 3.1. Les résultats de ces estimations sont présentés dans les tableaux 5 et 6. Dans une seconde étape et pour différents instruments $x_t \in I(t)$ nous estimons par GMM une constante c telle que (29) soit vrai. Nous testons alors $\mathbf{H_0}$. Les résultats de ces tests sont aussi dans les tableaux 5 et 6.

8 Résultats et commentaires

Les résultats de l'expérience sur les données simulées sont donc présentés au tableau 2. Rappelons que l'on est dans un cas où l'hypothèse $\mathbf{H_0}$ que nous testons est vraie, puisque tous les arguments utilisés pour y arriver (formulation structurelle et hypothèse de kurtosis conditionnelle constante) sont vérifiés par les GARCH(1,1). La constante c que nous estimons vaut nous la connaissons et elle vau $\alpha(\kappa - 1) + \gamma = \alpha\kappa + \beta = 1.05$. Pour différents instruments $x_t \in I(t)$ le tableau 2 présente la valeur (en moyenne) estimée de c, et pour différents seuils de tests, les fréquences de rejets de $\mathbf{H_0}$.

Nous remarquons que l'estimateur $\hat{c}$ est de moins en moins "bon" à mesure que l'on augmente les instruments. C'est normal puisque cela augmente le nombre des conditions de moments dans les GMM, et de ce fait la difficulté à satisfaire ces conditions.

Les fréquences de rejets de $\mathbf{H_0}$ sont bien de l'ordre des seuils de tests utilisés (1%, 5% et 10%) quand nous prenons comme instruments les valeurs passées des innovations ϵ_t et ϵ_{t-1}. Elles augmentent sans exploser dés que nous considérons les puissances des valeurs passées des innovations, pour se stabiliser autour de 14% (moins quand on considère les puissansces de ϵ_{t-1}) au seuil de 5%. Comme le test est asymptotique, nous avons repris les cas $x_t = (1, \epsilon_t, \epsilon_t^2)$ et $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_{t-1}, \epsilon_{t-1}^2)$ avec des longeurs de séries T de plus en plus grandes, voir résultats dans le tableau 3. Si l'estimation de la constante c est meilleure, les fréquences de rejets restent, pour le moins insensibles.

Une partie de l'explication résiderait dans le fait qu'en considérant ces puissances nous augmentons le nombre de conditions de moments dans le cadre des GMM. Ensuite ajouter aux instruments des puissances de ϵ_t pour représenter l'information $I(t)$ serait moins opportun que l'ajout des "lag" ϵ_{t-1}.

Les fréquences de rejets de $\tilde{\mathbf{H}}_0$ (dans le tableau 4) qui devraient êtres les mêmes que celles de $\mathbf{H_0}$ restent en deçà des seuils de tests, et pour tous les instru-

ments. Moins gênant à la limite, ces derniers résultats nous rappellent que l'on teste, en fait l'hypothèse jointe $\kappa_t = \mathrm{C}^{ste}$ et $\mathbf{H_0}$.

Les résultats de l'expérience sur les données réelles sont présentés dans les tableaux 5 pour IBM et 6 pour Intel. Le test de $\mathbf{H_0}$ est repris avec les mêmes différents instruments que dans le cas des données simulées. Une première remarque concerne les estimateurs des paramètres du GARCH. Ces derniers sont estimés avec moins de variance asymptotique dans le cas d'IBM (par exemple $\sigma_{asympto}(\beta) = 0.01745$) que dans le cas d'Intel ($\sigma_{asympto}(\beta) = 0.0455$) (deux à trois fois moins). Ceci peut expliquer les différences dans les résultats des tests. Dans le cas de Intel, nous avons de grandes "P-value" (entre 0.15 et 0.7). On ne rejette donc jamais le test. Pour IBM on rejette dans les cas où les instruments contiennent les puissances des valeurs passées des innovations. Ce sont les mêmes cas où nous avons eu des fréquences "pas trés attendues" avec les données simulées.

9 Conclusion

Dans ce rapport nous avons rappelé les modèles de volatilité de l'hétéroscédasticité conditionnelle ainsi que la méthode généralisés des moments GMM, c'est à dire le cadre économétrique et financier de ce travail de stage.

Nous avons exposé le problème qui a motivé ce travail, la caractérisation du biais de "simultanéité"obtenu lors de l'estimation par GMM du paramètre de la persistence de la volatilité. Cela revient à caractériser le coefficient α_t de la régression affine, conditionnelle à l'information passée, de la variance conditionnelle à cette même information, sur les carrés des rendements. Nous avons rappelé l'utilité de partir d'une formulation structurelle de la volatilité pour répondre à cette question.

Nous avons supposé, comme dans le cadre des modèles GARCH, que la kurtosis conditionnelle des innovations standardisées est constante, puis à partir d'un modèle AR(1) de la variance condionnelle au passé des rendements financiers nous avons dérivé un test de l'hypothèse que le coefficient α_t est constant.

Les résultats de ce test ont une conclusion qui se résume par le fait que la propriété de base des GARCH que nous avons testée (α_t constant) n'a pas été remise en cause par ce test sur les données considérées. Ils suggèrent de relâcher l'hypothèse que la kurtosis conditionnelle est constante, mais ils suggèrent aussi de mener le test sur d'autres données, à considérer des GARCH asymétriques, voir d'autres modèles plus généraux où la variance conditionnelle ne coincide pas avec sa régression affine conditionnelle au passé sur les carrés des rendements comme c'est le cas des GARCH. Nous avons considéré, par exemple des GARCH(1,1) asymétriques comme définis dans [EG]. Cela nous a amené à considérer d'autres tests, toujours pour caractériser le coefficient α_t, comme tester l'existence d'une constante α telle que $E_t[(\epsilon_{t+2}^2 - \alpha\epsilon_{t+1}^2)\epsilon_{t+1}] = 0$. Test qui n'a pas d'intérêt (statistiquement parlant) si on n'a pas de "skewness", $E[\epsilon_{t+1}^3] = 0$.

Des questions techniques sont restées aussi en suspens , comme par exemple comment simuler un modèle où la variance conditionnelle ne coincide pas avec sa régression affine conditionnelle au passé sur les carrés des rendements. Nous avons aussi effectué plusieurs tests (de moyennes par exemple) pour vérifier nos procédures.

Table 1: Statistiques descriptives des log-rendements journaliers d'IBM et Intel, période du 20/05/91 au 20/08/02

Statistique	IBM	INTEL
Moyenne	4.04×10^{-4}	4.21×10^{-4}
StDev	0.0215	0.0259
Skewness	-0.0387	-1.507
Kurtosis	8.587	35.701
Min	-0.1689	-0.4288
Max	0.1236	0.1732

Figure 1: Évolution et histogramme des log-rendements de IBM et Intel

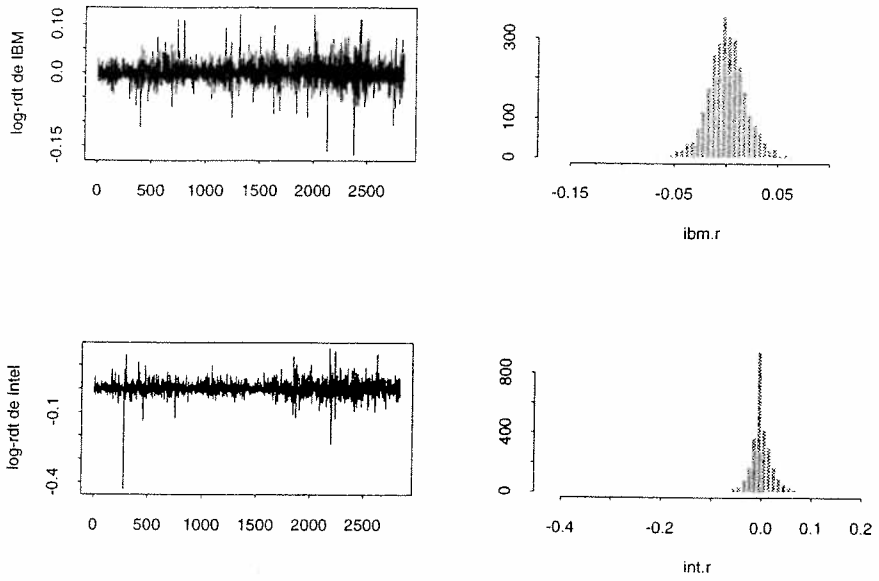

Table 2: Test de H_0; 1000 pas de temps et 200 replications sont utilisés. Les paramètres du GARCH(1,1) sont $(\omega = 0.01, \alpha = 0.05, \beta = 0.9)$

Seuil de test	1%	5%	10%
Instruments $x_t = (1, \epsilon_t)$		$\hat{c} = 0.9793(0.1135)$	
Nbre de rejets	0%	4%	8%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2)$		$\hat{c} = 0.9249(0.1013)$	
Nbre de rejets	6%	16%	25%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3)$		$\hat{c} = 0.9039(0.1007)$	
Nbre de rejets	6%	14.5%	23.5%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4)$		$\hat{c} = 0.8656(0.0923)$	
Nbre de rejets	4%	14%	23%
Instruments $x_t = (1, \epsilon_t, \epsilon_{t-1})$		$\hat{c} = 0.9414(0.1079)$	
Nbre de rejets	0.5%	2%	6%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_{t-1}, \epsilon_{t-1}^2)$		$\hat{c} = 0.8774(0.0922)$	
Nbre de rejets	4%	11%	18.5%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3)$		$\hat{c} = 0.82849(0.0840)$	
Nbre de rejets	5%	10%	18.5%
Instruments $x_t = (1, \epsilon_t, \cdots, \epsilon_t^4, \epsilon_{t-1}, \cdots, \epsilon_{t-1}^4)$		$\hat{c} = 0.77698(0.0741)$	
Nbre de rejets	3.5%	12.5%	18%

Table 3: fréquences des rejets de H_0 pour différentes tailles d'échantillons et 200 replications

Seuil de test	1%	5%	10%	$\hat{c}(StDev)$
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2)$				
$T = 1000$	6%	16%	25%	0.9249(0.1013)
$T = 3000$	6%	17%	25.5%	1.00138(0.06838)
$T = 6000$	4.5%	15%	24%	1.0225(0.0508)
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_{t-1}, \epsilon_{t-1}^2)$				
$T = 1000$	4%	11%	18.5%	0.8774(0.0922)
$T = 3000$	5%	13.5%	23%	0.9690(0.06308)
$T = 6000$	5%	18%	24.5%	0.9964(0.04758)

Table 4: Test de $\tilde{H}_0$; 1000 pas de temps et 200 replications sont utilisés

Seuil de test	1%	5%	10%
Instruments $x_t = (1, \epsilon_t)$		$\hat{c} = 0.34689(0.02848)$	
Nbre de rejets	0.5%	2.5%	7%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2)$		$\hat{c} = 0.34531(0.02614)$	
Nbre de rejets	0.5%	2%	6%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3)$		$\hat{c} = 0.3424(0.02525)$	
Nbre de rejets	0%	0.5%	3%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4)$		$\hat{c} = 0.3415(0.0242)$	
Nbre de rejets	0%	0.5%	2%
Instruments $x_t = (1, \epsilon_t, \epsilon_{t-1})$		$\hat{c} = 0.3472(0.0274)$	
Nbre de rejets	0%	0%	5.5%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_{t-1}, \epsilon_{t-1}^2)$		$\hat{c} = 0.3462(0.027)$	
Nbre de rejets	0.5%	2%	4.5%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3)$		$\hat{c} = 0.3442(0.0228)$	
Nbre de rejets	0%	0%	3%
Instruments $x_t = (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3, \epsilon_{t-1}^4)$		$\hat{c} = 0.3442(0.02143)$	
Nbre de rejets	0%	0%	3%

Table 5: Estimation d'un GARCH(1,1) et Test de suridentification sur les données IBM

$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
1.09×10^{-5}	0.07088	0.9088
(5.7×10^{-6})	(0.01124)	(0.01745)
Instruments $= (1, \epsilon_t)$		
$\hat{c}$	1.015(0.1285)	
Statistique (P-value)	1.275(0.2588)	
Instruments $= (1, \epsilon_t, \epsilon_t^2)$		
$\hat{c}$	0.8035(0.1089)	
Statistique (P-value)	9.289(0.00961)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3)$		
$\hat{c}$	0.5926(0.0803)	
Statistique (P-value)	14.41(0.0024)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4)$		
$\hat{c}$	0.588(0.07967)	
Statistique (P-value)	14.52(0.0058)	
Instruments $= (1, \epsilon_t, \epsilon_{t-1})$		
$\hat{c}$	1.017(0.1273)	
Statistique (P-value)	1.288(0.5252)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_{t-1}, \epsilon_{t-1}^2)$		
$\hat{c}$	0.8019(0.1075)	
Statistique (P-value)	10.23(0.0366)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3)$		
$\hat{c}$	0.5665(0.07786)	
Statistique (P-value)	16.26(0.0124)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3, \epsilon_{t-1}^4)$		
$\hat{c}$	0.565(0.0777)	
Statistique (P-value)	16.47(0.03608)	

Table 6: Estimation d'un GARCH(1,1) et Tests de suridentification sur les données d'INTEL

$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
3.18×10^{-5}	0.0789	0.875
(2.82×10^{-5})	(0.022)	(0.0455)
Instruments $= (1, \epsilon_t)$		
$\hat{c}$	0.81(0.167)	
Statistique (P-value)	1.08(0.2986)	
Instruments $= (1, \epsilon_t, \epsilon_t^2)$		
$\hat{c}$	0.8533(0.1702)	
Statistique (P-value)	1.637(0.4411)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3)$		
$\hat{c}$	0.7348(0.1436)	
Statistique (P-value)	5.241(0.155)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4)$		
$\hat{c}$	0.7294(0.1424)	
Statistique (P-value)	5.371(0.2513)	
Instruments $= (1, \epsilon_t, \epsilon_{t-1})$		
$\hat{c}$	0.7935(0.1438)	
Statistique (P-value)	1.144(0.5643)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_{t-1}, \epsilon_{t-1}^2)$		
$\hat{c}$	0.848(0.1465)	
Statistique (P-value)	2.512(0.6425)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3)$		
$\hat{c}$	0.7865(0.1356)	
Statistique (P-value)	5.36(0.4985)	
Instruments $= (1, \epsilon_t, \epsilon_t^2, \epsilon_t^3, \epsilon_t^4, \epsilon_{t-1}, \epsilon_{t-1}^2, \epsilon_{t-1}^3, \epsilon_{t-1}^4)$		
$\hat{c}$	0.7879(0.1344)	
Statistique (P-value)	5.452(0.7084)	

References

[AI] A. ALAMI IDRISSI, *Estimation de la persistance de la volatilité* , Thèse de Doctorat, 1999.

[AR] A. ALAMI, E. RENAULT, *Risque de modèle de volatilité* , Cahiers Scientifiques du CIRANO 2001s-06.

[BO] T. BOLLERSLEV, *Generalized Autoregressive Conditional Heteroskedasticity* , Journal of Econometrics 31 (1986) 307-327.

[EG] R. F. ENGEL, V. K. NG, *Measuring and Testing the Impact of News on Volatility*, The Journal of finance, 48 (1993), 1749-1778.

[EN] R. F. ENGEL, *Autoregressive Condional Heteroskedasticity With Estimates of the Variance of United Kingdom Inflation*, Econometrica, 50, 987-1007.

[HA] J.D. HAMILTON, *Time Series Analysis*, Princeton University Press 1994.

[HL] L.P. HANSEN, *Large Sample Properties of Generalized Method of Moments Estimators*, Econometrica 50:1029-54, 1982.

[MR] N. MEDDAHI, E. RENAULT, *Temporal Aggregation of Volatility Models* , Cahiers Scientifiques du CIRANO 2000s-22.

[SJ] J. SKOGLUND, *A simple efficient GMM estimator of GARCH models* , Working Paper Series in Economics and Finance No. 434, Février 2001.

ANNEXE

L'aspect informatique a constitué une grande part de ce travail, aussi et pour toute fin utile nous donnons une brève description des principales procédures écrites (en Splus), ainsi que leurs codes. Ces fonctions font appel à d'autres fonctions qui souvent font appel, à leur tour, à des procédures codées en language C que nous avons écrites. Nous reproduisons donc aussi les codes des fonctions qui nous semblent les plus importantes.

Pour ceux qui pourraient accéder au réseau du CIRANO, toutes les procédures ont été regroupées dans les deux fichiers `outils.ssc` et `outils.c` auquels on peut accéder par le chemin `/usagers/boudiafm/stage`.

Description des fonctions

ga11.sim : cette fonction simule un GARCH(1,1) (asymétrique). Cela prend comme arguments requis *theta* un vecteur (ou liste) des paramètres ω, α, β et optionnellemnt δ. Comme arguments optionnels *len* longueur des séries, égal 1000 par défaut, *distr* loi des innovations standardisées. Sont valables "g"(gausienne par défaut) et "t"de Student auquel cas rajouter l'argument *df* pour les degrés de liberté. Cela retourne une liste avec deux composantes: $et pour les innovations et $sigma.t pour la volatilité.

ga11.mv : estime par maximum de vraissemblance un GARCH(1,1) asymétrique. Cela utilise la fonction *garch* du module GARCH de Splus qui ne contient pas le modèlde asymétrique de [EG]. Cela prend un argument, la série des innovations et retourne une liste de quatre composantes: $coefficients, $objective la valeur de la vraissemblance à l'optimum, $et la séries des innovations et $sigma.t pour la volatilité.

g11.gmm : estime par GMM (à la manière de [SJ]) un GARCH(1,1) (asymétrique). cela prend comme argument requis une série d'innovations. Optionnelle-

ment un "flag" *asym* égal F par défaut, et un autre *trace* pour afficher les étapes de la procédure. Cela retourne une liste avec trois composantes: $parametres, $serie et $var.cond pour les innovations et leurs variances conditionnelles.

tsls : estime par GMM un modèle linéaire et donne le résultat d'un test de suridentiffication. cela prend trois arguments requis *y,z, x* des séries de variables endogènes, explicatives et instrumentales tel que $dim(x) > dim(z)$. Cela retourne une liste de six composantes: $c la constante c telle que $E[x'(y-cz)] = 0$, $std.hetero et $std.homo les écart-type asymptotiques de c selon qu'on veuille prendre en compte ou pas l'hétéroscédasticité. $statistique, $pvalue du test de suridentiffication et enfin $residus pour les résidus de la regression.

A Codes Splus

ga11.sim simule un GARCH(1,1) (asymétrique)

```
function(theta, len = 1000, distr = "g", df = (len - 1))
{
theta <- unlist(theta)
u <- if(distr == "t") rt(len, df = df) else rnorm(len)
sig2 <- rep(0, len)
sig2 <- fsig2.asym(theta, u, sig2)
sigma.t <- sqrt(sig2)
et <- c(u[1] * sqrt(theta[1]/(1 - theta[2] - theta[3])), sigma.t[ -
len] * u[-1])
list(et = et, sigma.t = sigma.t)
}
```

fsig2.asym appelle une fonction en C pour calculer la variance conditionnelle étant donné la séries des rendements standardisés, les paramètres et un vecteur pour y stocker le résultat.

```
function(theta, u, sig2)
{
theta <- unlist(theta)
if(length(theta) < 4)
theta <- c(theta, 0)
taille <- length(u)
.C("fsig2_asym",
theta = as.double(theta),
sig2 = as.double(sig2),
u = as.double(u),
taille = as.integer(taille))$sig2
}
```

ga11.mv estime un GARCH(1,1) (asymétrique) par maximum de vraissemblance

```
function(eta)
{
len <- length(eta)
vareta <- var(eta)
start <- garch( ~ -1,  ~ garch(1, 1), series = eta, trace = F)$coef
delta <- (var(eta[ - len], eta[-1]^2) * var(eta^2))/(2 * var(eta[ -
len]^2, eta[-1]^2) * var(eta))
start <- c(start, delta)
names(start) <- c("A", "ARCH(1)", "GARCH(1)", "LEV")
mv <- nlminb(start = start, obj = moinslogMV, eta = eta, vareta =
vareta)
sig2 <- rep(0, len)
sig2 <- f.sig2(mv$parameters[1:3], eta + mv$parameters[4], sig2)
sigma.t <- sqrt(sig2)
list(coefficients = as.matrix(mv$parameters, ncol = 1), objective = mv$
objective, series = eta, sigma.t = sigma.t)
}
```

moinslogMV appelle la fonction *S_logMV* pour calculer la log-vraissemblance.

```
function(para, eta, vareta)
{
resultat <- 0
delta <- if(length(para) > 3) para[4] else 0
.C("S_logMV",
theta = as.double(para),
taille = as.integer(length(eta)),
eta = as.double(eta + delta),
vareta = as.double(vareta),
resultat = as.double(resultat))$resultat
}
```

f.sig2 fait la même chose que *fsig2.asym* mais prend les rendements (non standardisés)

```
function(theta, eta, sig2)
{
theta <- theta[1:3]
vareta <- var(eta)
taille <- length(eta)
.C("fsig2",
theta = as.double(theta),
sig2 = as.double(sig2),
eta = as.double(eta),
vareta = as.double(vareta),
taille = as.integer(taille))$sig2
}
```

g11.gmm estime un GARCH(1,1) par GMM à la manière de [SJ]

```
function(eta, asym = F, trace = T)
{
len <- length(eta)
vareta <- var(eta)
if(trace)
cat("\n -- Premiere etape:  estimation par ML   ---------------\n"
)
mv11 <- if(!asym) garch( ~ -1,  ~ garch(1, 1), series = eta, trace = F)
 else ga11.mv(eta)
emv11 <- mv11$coef
if(trace)
cat("\n ---------------------- MLE ----------------------------\n "
)
std <- if(!asym) sqrt(diag(cov(mv11))) else NA
if(trace) {
if(!asym)
print(cbind(emv11, std = std), digits = 3)
else print(emv11, digits = 3)
```

```
}
# --calculer la serie z des innovations des rdts standardisees en MLE ----
z <- eta[-1]/mv11$sigma.t[ - len]
v4 <- mean(z^4)
v3 <- mean(z^3)
#---- MINIMISATION de g11.gmm.obj la fct objectif de GMM ---
#     en partant de theta0=MLE
if(trace) cat("\n -- Deuxieme  etape:
minimisation de la fct obj GMM (en partant de MLE) ... \n")
theta0 <- emv11
#-------- PAS ASYMM-------------
if(!asym) {
gmm <- nlminb(start = theta0, obj = g11.gmm.obj, eta = eta,
vareta = vareta, len = len, v3 = v3, v4 = v4, trace = F,
control = nlminb.control(iter.max = 100, x.tol = 1e-06),
low = c(0, 0, 0), up = c(vareta, 1, 1))
}
#-------- ASYM-------------
if(asym) {
gmm <- nlminb(start = theta0, obj = g11.gmm.obj, eta = eta,
vareta = vareta, len = len, v3 = v3, v4 = v4, trace = F,
asym = T, low = c(0, 0, 0, -1), up = c(vareta, 1, 1,
1))
}
if(!asym)
names(gmm$parameters) <- c("Cte.in.var", "ARCH(1)", "GARCH(1)")
if(asym)
names(gmm$parameters) <- c("Cte.in.var", "ARCH(1)", "GARCH(1)",
"LEV")
# ----- var. cond. et  innov. stand. en gmm$par ------------
sig2 <- rep(0, len)
if(!asym)
sig2 <- if(all(!is.na(gmm$param))) f.sig2(gmm$param, eta, sig2)
 else mv11$sig^2
if(asym)
```

```
sig2 <- if(all(!is.na(gmm$param))) f.sig2(gmm$param, eta + gmm$
param[[4]], sig2) else mv11$sig^2
z <- eta[-1]/sqrt(sig2[ - len])
v4 <- mean(z^4)
v3 <- mean(z^3)
#-------- grad.sig2-------------
grad.sig2 <- rep(0, 4 * len)
if(all(!is.na(gmm$param)))
grad.sig2 <- f.gradsig2(gmm$par, eta, sig2, grad.sig2)
grad.sig2 <- matrix(grad.sig2, nrow = len)
if(!asym)
grad.sig2 <- grad.sig2[, 1:3]
std <- NA
if(any(grad.sig2 != 0)) {
Vgmm <- (v4 - 1 - v3^2) * inverse(t(grad.sig2) %*% grad.sig2) *
mean(sig2^2)
std <- sqrt(diag(Vgmm))
}
coef <- cbind(gmm$par, std = std)
if(trace) {
cat("\n ------------------ ESTIMATEUR  GMM    -----------------\n"
)
print(coef, digits = 3)
}
out <- list(parametres = coef, serie = eta, var.cond = sig2)
out
}
```

g11.gmm.obj calcule la fonction objective à minimiser dans la fonction *g11.gmm*

```
function(theta, eta, vareta, len, v3, v4, asym = F)
{
theta <- if(length(theta) < 4) c(unlist(theta), 0) else unlist(theta)
sig2 <- rep(0, len)
# ---------  calcul de sig2 ---------
```

```
sig2 <- if(!asym) f.sig2(theta[1:3], eta, sig2) else f.sig2(theta[
1:3], eta + theta[4], sig2)
# ---------   calcul de z= eta/sig2  ---------
z <- eta[-1]/sqrt(sig2[ - len])
del <- v4 - 1 - v3^2
# ---------   calcul du gradient de sigma^2 ----------------
grad.sig2 <- rep(0, 4 * len)
grad.sig2 <- f.gradsig2(theta, eta, sig2, grad.sig2)
grad.sig2 <- matrix(grad.sig2, nrow = len)
if(!asym)
grad.sig2 <- grad.sig2[, 1:3]
C <- (v3 * z - z^2 + 1)/(del * sig2[ - len])
if(any(is.na(C)))
return(10000000000.)
g <- t(grad.sig2[ - len,  ]) %*% C
#
L <- t(grad.sig2[ - len,  ] * C) %*% (grad.sig2[ - len,  ] * C)
#
#cat("\n Theta= ", format(theta, digit=4))
if(any(is.na(L)) || any(is.na(g))) return(1e+20)
##if(any(is.na(L))||any(is.na(g)))stop(" Probleme avec inversion de matrice des
poids ou bien avec le calcul des moments generalises")
#if(any(is.inf(L))stop(" Probleme avec inversion de matrice des poids")
if(any(is.inf(L))) return(1e+20)
if(any(L == 0))
return(1e+20)
as.numeric(t(g) %*% inverse(L) %*% g)
}
```

f.gradsig2 calcule le gradient de la variance conditionnelle par rapport au paramètres.

```
function(theta, eta, sig2, gradsig2)
{
if(length(theta) < 4)
```

```
theta <- c(theta, 0)
vareta <- var(eta)
eta <- eta + theta[4]
taille <- length(eta)
.C("fgradsig2",
theta = as.double(theta),
sig2 = as.double(sig2),
gradsig2 = as.double(gradsig2),
eta = as.double(eta),
vareta = as.double(vareta),
taille = as.integer(taille))$gradsig2
}
```

B Codes C

S_logMV est appelée par *moinslogMV* pour calculer la log-vraissemblance

```
void S_logMV( double *theta, long *taille, double *eta, double
*vareta, double *resultat){

  double sig2[*taille];
  double  tmp[*taille];
  double logsig2[*taille];
  long t;

  sig2[0]= theta[0] + theta[1]* eta[0]*eta[0] + theta[2]*(*vareta);
  tmp[0]= eta[0]*eta[0]/(2*(*vareta));

  logsig2[0]= log(sig2[0]);

  for(t=1;t<(*taille);t++)
    {
    sig2[t]= theta[0] + theta[1]* eta[t]*eta[t] + theta[2]*sig2[t-1];
    tmp[t]=eta[t]*eta[t]/(2*sig2[t-1]);
    logsig2[t]=log(sig2[t]);
    }

  *resultat=(*taille)* log(2*PI)/2.0 + somm((*taille),logsig2)/2.0
+ somm((*taille), tmp );
}
```

fsig2 est appelée par la fonction Splus *f.sig2*.

```
void fsig2(double *theta, double *sig2, double *eta, double *vareta,
long* taille)
{
  long t;
  sig2[0]= theta[0] + theta[1]* eta[0]*eta[0] + theta[2]*(*vareta);
  for(t=1;t<(*taille);t++)
```

```
    sig2[t]= theta[0] + theta[1]* eta[t]*eta[t] + theta[2]*sig2[t-1];

}
```

fsig2_asym est appelée par la fonction Splus *fsig2.asym*

```
void fsig2_asym(double *theta, double *sig2, double *u, long* taille)
{
  long t;
  double sig0=pow(theta[0]/(1- theta[1]- theta[2]), 0.5);
  sig2[0]= theta[0]+ theta[1]*(sig0*u[0]+ theta[3] )*( sig0*u[0]
+ theta[3] )+ theta[2]*sig0*sig0;
  for(t=1;t<(*taille);t++)
    sig2[t]= theta[0] + theta[1]*( sqrt(sig2[t-1])*u[t] + theta[3])
*( sqrt(sig2[t-1])*u[t] + theta[3]) +theta[2]*sig2[t-1];

}
```

fgradsig2 est appelée par la fonction Splus *f.gradsig2*

```
void fgradsig2(double *theta, double *sig2,double *gradsig2, double *eta,
 double *vareta, long* taille)
{
  long t;
  gradsig2[0]=1.0 + theta[2]/(1-theta[1]-theta[2]);
  gradsig2[*taille]=eta[0]*eta[0]+(pow(theta[3],2)*(1-theta[1]-theta[2])
+(theta[0]+theta[1]*pow(theta[3],2)))*theta[2]/pow(1-theta[1]-theta[2],2);
  gradsig2[2*(*taille)]=(*vareta)+(theta[0]+theta[1]*pow(theta[3],2))
*theta[2]/pow(1-theta[1]-theta[2],2);
  gradsig2[3*(*taille)]=2*theta[1]*eta[0]*eta[0]+theta[1]*theta[2]/
pow(1-theta[1]-theta[2],2);

  for(t=1;t<(*taille);t++){
    gradsig2[t]= 1.0+ theta[2]*gradsig2[t-1];
    gradsig2[t+(*taille)]= eta[t]*eta[t] + theta[2]*gradsig2[t-1+(*taille)];
    gradsig2[t+2*(*taille)]= sig2[t-1]+ theta[2]*gradsig2[t-1+2*(*taille)];
```

```
    gradsig2[t+3*(*taille)]= 2*theta[1]*eta[t]*eta[t]+ theta[2]*
gradsig2[t-1+3*(*taille)];

  }
}
```