Université de Montréal

Analyse des rendements et des volatilités des indices S&P 500 et TSE 300

Par
Meriem Aït-Ouyahia

Directeur de recherche : Nour Meddahi

Département des sciences économiques
Faculté des arts et des sciences

Rapport de recherche présenté à la faculté des études supérieures
en vue de l'obtention du grade de
Maîtrise ès sciences (M.Sc)

Décembre, 1999

Résumé

Ce travail est consacré à la modélisation et à l'estimation de deux indices boursiers, le Standards and Poors 500 (S&P500) et le Toronto Stock Exchange 300 (TSE 300) qui représentent le marché des États Unis et du Canada respectivement. Il présente une modélisation de ces séries de la façon la plus usuelle en finance : un modèle autorégressif pour les rendements et un modèle GARCH pour le risque. Deux méthodes d'estimation sont mises en œuvre : la première, couramment employée, est une estimation jointe de la moyenne et de la variance, dite aussi méthode du quasi maximum de vraisemblance. La seconde est la méthode des moments généralisés. Celle ci permet une estimation séparée de la moyenne et de la variance, ce qui la rend robuste à la forme de l'hétéroscédasticité. Les résultats de ce travail tendent à montrer que ces deux estimations sont aussi efficaces mais que les valeurs des coefficients sont différentes : on peut donc penser qu'avec la méthode du quasi maximum de vraisemblance, la moyenne peut être contaminée par une variance mal spécifiée. Par ailleurs, étendue au contexte multivarié, l'analyse montre qu'il y a causalité en moyenne entre les indices étudiés.

Abstract

This work deals with the estimation of two stock indices, the Standards and Poors 500 (S&P500) and the Toronto Stock Exchange (TSE 300), that represent the stock market in the United States an Canada respectively. Two models are used : the autoregressive model for the mean and the GARCH one for the variance. Two methods of estimation are also implemented. The first one is the commonly used quasi maximum likelihood estimation (QMLE) that estimates jointly both the mean and the variance. The second one is the general method of moments (GMM) that allows a separate estimation of the mean and the variance and therefore is robust in terms of heteroscedasticity. The results of this work indicate that those two estimations are equally efficient even though the values of the coefficients are different. One can therefore think that the first method provides a contaminated mean when the variance is misspecified. In addition, in a multivariate context, the result of this work shows that there is causality in mean between the two indices.

Table des matières

1) Introduction.....7

2) Modélisation de la moyenne et de la variance.....10

2.1) Le rendement.....10

2.2) Le risque.....11

2.2.1) La modélisation du risque.....12

2.2.2) La structure des ARMA du carré des résidus.....12

2.2.3) Généralisation aux modèles GARCH et EGARCH.....13

3) Les données : le S&P500 et le TSE 300.....15

4) Analyse jointe de la moyenne et de la variance.....18

5) Analyse robuste de la moyenne vis à vis de l'hétéroscédasticité.....24

5.1) La méthode des moments généralisés.....24

5.2) Estimations.....26

5.2.1) La méthode des moments généralisés juste-identifié.....26

5.2.2) La méthode des moments généralisés sur-identifié.....27

5.2.3) Suridentification et correction de l'hétéroscédasticité.....27

5.3) Les Tests.....31

5.3.1) Le J test de Hansen.....31

5.3.2) Le test de Box Pierce.....31

5.3.3) Le I test de Cumby et Huizinga.....32

6) Analyse de la variance.....33

7) Analyse jointe bivariée de la moyenne et de la variance :présentation générale.....37

7.1) Modèle VAR et causalité.....37

7.2) Modèle GARCH bivarié.....39

7.2.1) Modélisation.....39

7.2.2) Contraintes de positivité et conditions de stabilité.....40

7.2.3) Modèles contraints.....42

8) Analyse jointe bivariée de la moyenne et de la variance : cas particulier.....43

9) Analyse robuste bivariée de la moyenne vis à vis de l'hétéroscédasticité.....46

10) Conclusion.....51

Bibliographie.....52

Annexe 1 : Valeurs de la fonction de vraisemblance, du critère d'Akaike et de Schwartz pour le S&P500.....55

Annexe 2 : Valeurs de la fonction de vraisemblance, du critère d'Akaike et de Schwartz pour le TSE 300.....56

Annexe 3 : Estimation par le quasi maximum de vraisemblance du S&P500 et du TSE300.....57

Annexe 4 : Estimation de la moyenne du S&P500 et du TSE300 par les MCO.....59

Annexe 5 : Estimation de la moyenne du S&P500 et du TSE300 par les GMM.....61

Annexe 6 : Estimation de la moyenne du S&P500 et du TSE300 par les GMM1.....63

Annexe 7 : Le l test.....65

Annexe 8 : Valeurs de la fonction de vraisemblance, du critère d'Akaike et de Schwartz pour le S&P500 estimé de manière bivariée.....66

Annexe 9 : Valeurs de la fonction de vraisemblance, du critère d'Akaike et de Schwartz pour le TSE 300 estimé de manière bivariée.....67

Annexe 10 : Estimation de la moyenne bivariée du S&P500 et du TSE300 par les MCO.....68

Annexe 12 : Estimation de la moyenne bivariée du S&P500 et du TSE300 par les GMM.....70

Liste des tableaux et graphiques

Graphique 1 : S&P500.....17
Graphique 2 : TSE 300.....17
Graphique 3 : Corrélogramme des résidus aux carrés du S&P500.....22
Graphique 4 : Corrélogramme des résidus aux carrés du TSE 300.....23
Tableau 1 : Statistiques du S&P500 et du TSE 300.....15
Tableau 2 : AR(5)-GARCH(2,2) pour le S&P500.....20
Tableau 3 :AR(1)-GARCH(1,2) pour le S&P500.....20
Tableau 4 :AR(3)-GARCH(2,2) pour le TSE 300.....21
Tableau 5 :AR(1)-GARCH(2,2) pour le TSE 300.....21
Tableau 6 : Estimations de la moyenne du S&P500.....30
Tableau 7 : Estimations de la moyenne du TSE 300.....30
Tableau 8 : Valeur de la fonction de vraisemblance pour le S&P500 et le TSE 300.....35
Tableau 9 : Valeur du critère d'Akaike pour le S&P500 et le TSE 300.....35
Tableau 10 : Valeur du critère de Schwartz pour le S&P500 et TSE 300.....35
Tableau 11 : GARCH(1,2) pour le S&P500.....36
Tableau 12 : GARCH(1,2) pour le TSE 300.....36
Tableau 13 : AR(4)-GARCH(1,2) pour le S&P500.....37
Tableau 14 : AR(3)-GARCH(1,2) pour le TSE 300.....37
Tableau 15 : VAR(4)-GARCH(1,2) pour le S&P500.....44
Tableau 16 : VAR(1)-GARCH(1,2) pour le S&P500.....44
Tableau 17 : VAR(4)-GARCH(2,2) pour le TSE 300.....45
Tableau 18 : VAR(1)-GARCH(1,2) pour le TSE 300.....45
Tableau 19 : Estimation de la moyenne bivariée du S&P500 par les GMM1.....47
Tableau 20 : Estimation de la moyenne bivariée du TSE 300 par les GMM1.....48
Tableau 21 : Statistiques du test de Wald et du ratio de vraisemblance pour le S&P500.50
Tableau 22 : Statistiques du test de Wald et du ratio de vraisemblance pour le TSE 300.50

1) Introduction

L'incertitude est au cœur des développements de la théorie économique moderne, et plus particulièrement de la théorie financière. L'incertitude est, notamment, un des déterminants les plus importants dans la formation des prix des actifs financiers. C'est une des raisons pour lesquelles le développement de nouvelles techniques économétriques des séries temporelles a été très fertile ces vingt dernières années.

Tout commence en 1963, lorsque Mandelbrot constate deux faits sur les données financières. Premièrement, elles ont des queues de distribution plus épaisses que celles de la loi normale, c'est à dire que les valeurs loin de la moyenne apparaissent avec des probabilités plus élevées que pour une loi normale; le coefficient d'aplatissement étant élevé. Deuxièmement, des regroupements de volatilité, « clustering effect », sont remarqués dans les faits empiriques. Les fortes (faibles) variations de la variance sont suivies par d'autres fortes (faibles) variations, mais dont le signe n'est pas prévisible.

Puis, dans les années 70, on assiste au développement des modèles de séries temporelles. Ces modèles permettent de décrire les moments conditionnels du premier ordre. Ils sont utiles mais insuffisants pour les données financières. Ils décrivent les rendements, mais ils ne donnent aucune indication sur le risque car ils supposent que la variance est constante. Or les prix spéculatifs ont une volatilité variable au cours du temps.

Les modèles de type ARCH (AutoRegressive Conditional Heteroscedastic), qui émergent avec Engle au début des années 80, permettent de mieux comprendre cette dynamique des prix des actifs financiers, et en particulier de leur variance. En effet, ils

font une distinction entre la distribution non conditionnelle et la distribution conditionnelle. Tandis que la première ne varie pas avec le temps, la distribution conditionnelle a une moyenne et une variance variables au cours du temps.

Dans notre étude, nous nous intéressons à l'estimation de deux indices boursiers. Les prix de ces indices sont très volatiles, il faudra donc le prendre en compte dans l'estimation. Ces deux indices sont le Standards and Poors 500 (S&P500) et le Toronto Stock Exchange 300 (TSE 300). Nous les avons choisis parce qu'ils incluent les plus grandes compagnies et reflètent le marché des États Unis et du Canada respectivement. En outre, ils ont l'avantage d'avoir une structure de marché et des règlements similaires.

Nous modélisons les données de façon standard, soit un modèle AutoRégressif (AR) pour la moyenne et un modèle GARCH pour la variance. Pour estimer le rendement de ces deux séries, nous utilisons deux méthodes qui permettent de ne pas spécifier une fonction de densité pour les résidus : la méthode du quasi maximum de vraisemblance et celle des moments généralisés.

La méthode du quasi maximum de vraisemblance, une méthode couramment employée, estime de manière jointe la moyenne et la variance à l'aide d'une fonction de vraisemblance conditionnellement normale. Même si la véritable loi des données n'est pas conditionnellement normale, l'estimateur est convergent mais il faut corriger son écart type pour supprimer l'hétéroscédasticité.

De son côté, la méthode des moments généralisés est une estimation de la moyenne seule. Elle permet donc une estimation robuste à la forme de l'hétéroscédasticité. En effet, la moyenne n'est pas contaminée par une variance qui peut

être mal spécifiée. Si l'on compare les deux méthodes, on constate que les estimateurs sont aussi efficaces avec la méthode des moments généralisés mais, par contre, les coefficients diffèrent. On peut donc penser qu'avec la méthode d'estimation jointe la variance fausse l'estimation de la moyenne.

Pour estimer la variance, la méthode des moments généralisés n'a pas été employée pour des raisons de temps et de complexité. C'est donc, la méthode du quasi maximum de vraisemblance qui est retenue. Mais cette fois, nous ne faisons pas d'estimation jointe, car nous nous servons des carrés des résidus des modèles, sélectionnés pour la moyenne par la méthode des moments généralisés, comme variable interne au modèle GARCH.

Par ailleurs, nous remarquons qu'avec la croissance de l'intégration financière internationale, les produits échangés sur ces marchés sont de plus en plus liés. En octobre 1987, le krack boursier a montré que, par exemple, les mouvements des prix des actifs étaient fortement corrélés. Ainsi nous pouvons donc nous poser la question du lien en moyenne et en variance entre le S&P500 et le TSE 300 et étendre notre étude à une analyse bivariée. Nous nous aperçevons que le rendement du S&P500 cause celui du TSE 300 et que le rendement du TSE 300 cause également celui du S&P500.

En pratique, l'estimation des ARCH multivariés est plus compliquée que celle des ARCH univariés. En effet, ces modèles ont beaucoup trop de paramètres à estimer. Plusieurs études cherchent à restreindre de façon plausible les modèles ARCH multivariés pour réduire le nombre de paramètres. Ceci est assez contraignant car les restrictions doivent, entre autres, garantir que la matrice de covariance est définie

positive. Et il est aussi important, dans notre cas, que la causalité de Granger en variance puisse être appliquée en fonction de la paramétrisation choisie, c'est à dire que les informations passées d'une variable prédisent la variance conditionnelle de l'autre variable. Le modèle GARCH-BEKK tient compte de ces restrictions, mais il présente des problèmes de convergence. Faute de temps, nous simplifions le problème en estimant de façon jointe, par la méthode du quasi maximum de vraisemblance, un VAR séparé dont la variance suit un GARCH univarié.

La section 2 est consacrée à un rappel sur la modélisation de la moyenne et de la variance des séries temporelles. Ensuite, la section 3 est dédiée aux données du S&P500 et du TSE 300 qui ont servi au travail empirique. Les sections 4 et 5 présentent la méthode du quasi maximum de vraisemblance et celle des moments généralisés respectivement et leur application aux séries précédentes. Enfin, la section 6 présente plus particulièrement l'analyse de la variance par la méthode du quasi maximum de vraisemblance, l'analyse bivariée étant détaillée dans les sections 7,8 et 9.

2) Modélisation de la moyenne et de la variance

2.1) Le rendement

Les modèles de séries temporelles ont tout d'abord permis de modéliser les rendements variables dans le temps. Notamment, avec la méthodologie de Box-Jenkins (1976) pour spécifier, estimer et valider des modèles ARIMA (AutoRegressif Integrated

Moving Average). Pour le rendement des séries financières, nous utilisons usuellement des modèles AutoRégressif d'ordre p, AR(p), tel que :

$y_t = b_0 + \sum_{i=1}^{p} b_i y_{t-i} + \varepsilon_t$, où ε_t est un Bruit Blanc de moyenne 0 et de variance constante : $BB(0, \sigma^2)$. Nous réécrivont le modèle de la façon suivante :

$b(L) y_t = b_0 + \varepsilon_t$, où $b(L) = 1 - b_1 L - ... b_p L^p$, L étant l'opérateur de retard. Le processus sera stationnaire si le polynôme $b(L) = 1 - b_1 L - ... b_p L^p$ a toutes ses racines à l'extérieur du cercle unité.

2.2) Le risque

L'introduction des modèles ARCH, par Engle, a ouvert la possibilité de modéliser la variance variable au cours du temps[1]. Cette modélisation prend en compte l'existence de queues plus épaisses que celles de la loi normale, constatée dans les données financières. Le coefficient d'aplatissement associé à la distribution marginale est toujours supérieur à 3, qui est la valeur correspondant à la loi normale. La distribution est dite leptokurtique. Elle tient aussi compte des regroupements de volatilité, « clustering effect ». C'est à dire que les fortes (faibles) variations de la variance sont suivies par d'autres fortes (faibles) variations, mais dont le signe n'est pas prévisible.

[1] Ce n'est pas en contradiction avec l'hypothèse de stationnarité.

2.2.1) La modélisation du risque

Le premier modèle introduit par Engle, en 1982, est le modèle ARCH(q). Pour modéliser la variance, nous nous intéressons aux résidus du processus AR vu à la section 2.1.

$\varepsilon_t = \sqrt{h_t}.u_t$.

On suppose que u_t est indépendant et identiquement distribué (i.i.d.) de moyenne 0 et de variance 1 : $u_t = i.i.d.(0,1)$.

La moyenne conditionnelle de ε_t est :

$E(\varepsilon_t / I_{t-1}) = 0.$ I_{t-1} représente l'information passée, c'est à dire la sigma algèbre engendrée par $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}...\}$.

Sa variance conditionnelle varie avec le temps, tel qu'elle est exprimée linéairement en fonction des carrés des valeurs passées du processus :

$$Var(\varepsilon_t / I_{t-1}) = h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \quad .$$

Le modèle précédent est appelé ARCH(q). Pour assurer la positivité de h_t, il faut que $\alpha_i \geq 0$ et $\omega > 0$.

2.2.2) La structure des ARMA du carré des résidus

Si nous élèvons les innovations aux carrés :

$\varepsilon_t^2 = h_t u_t^2 = h_t(1 + (u_t^2 - 1)) = h_t + h_t(u_t^2 - 1)$,

nous pouvons réecrire le modèle de la façon suivante : $\varepsilon_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \eta_t$, avec $\eta_t = h_t(u_t^2 - 1)$. Ainsi, nous remarquons que le carré de l'innovation est un AR(q).

Nous pouvons réécrire le AR(q) en fonction d'un opérateur de retard comme nous l'avons vu lors de la modélisation des rendements. $\rho(L)\varepsilon_t^2 = \omega + \eta_t$, où $\rho(L) = 1 - \sum_{i=1}^{q} \alpha_i L^i$. Pour que le processus soit stationnaire, il faut que les racines de la fonction $\rho(L)$ soient de module supérieur à 1. La moyenne non conditionnelle de ε_t^2 est $\frac{\omega}{1 - \sum_i \alpha_i}$ pour qu'elle existe et soit positive il faut que $\sum_i \alpha_i < 1$.

2.2.3) Généralisation : GARCH et EGARCH

En 1986, Bollerslev a présenté les modèles GARCH. C'est un modèle ARCH avec une partie moyenne mobile en plus de la forme autorégressive.

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)h_t$$

Pour que le processus soit bien défini, il faut que tous les paramètres de la représentation AR(∞), soient non négatifs. Pour le GARCH(1,1) ceci est équivalent à $\alpha_1, \beta_1 \geq 0$. (Voir Nelson et Cao, 1992).

Pour que le processus soit stationnaire en covariance il faut que $\alpha_1 + \beta_1 < 1$.

Si nous remplacons h_t par $\varepsilon_t^2 - \eta_t$, où $\eta_t = h_t(u_t^2 - 1)$, nous pouvons réécrire :

$$\varepsilon_t^2 - \eta_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j (\varepsilon_{t-j}^2 - \eta_{t-j})$$

$$\Rightarrow \varepsilon_t^2 = \omega + \sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^{p} \beta_j \eta_{t-j} + \eta_t .$$

Dans ce cas-ci, le carré de l'innovation suit un ARMA(max(p,q),p). Il faut, tout de même prendre garde car la terminologie GARCH est symétrique par rapport à la terminologie ARMA usuelle au sens où le degré p apparaît comme un degré moyenne mobile et non nécessairement comme un degré autorégressif. Comme pour les rendements, il est en général suffisant de ne retenir qu'un petit nombre de paramètres, même pour modéliser un processus de variance très long. Nous voyons, donc, que ces modèles ont l'avantage de retenir une structure de retards plus riche que les ARCH.

En 1976, Black s'est aperçu que le comportement empirique des prix sur le marché des actions pouvait comporter un effet de levier. C'est à dire que la variance a tendance à croître (décroître) lorsque ε_{t-i} est négatif (positif). Le modèle GARCH(p,q) exponentiel ou EGARCH(p,q) introduit par Nelson en 1990, est en accord avec ces constatations empiriques.

$$\log h_t = \omega + \sum_{i=1}^{q} \alpha_i (\phi u_{t-i} + \gamma [| u_{t-i} | - E | u_{t-i} |]) + \sum_{i=1}^{p} \beta_i \log h_{t-i} .$$

Contrairement au GARCH(p,q), aucune restriction sur les paramètres α_i *et* β_i n'est nécessaire pour assurer la positivité de la variance conditionnelle.

Remarquons que nous pouvons estimer ces modèles par le quasi maximum de vraisemblance ou encore par la méthode des moments généralisés, comme nous le verrons ultérieurement. Théoriquement, nous pouvons facilement passer à un contexte

multivarié (voir la partie sur l'analyse bivariée de la variance). L'estimation des modèles ARCH multivariés s'apparente à celle des ARCH univariés. Cependant, en pratique, l'estimation est plus compliquée car il y a beaucoup de paramètres à estimer en plus.

3) Les données : le S&P500 et le TSE 300

Les données sont le taux de rendement du Standards and Poors 500 et du Toronto Stock Exchange 300 obtenu par Datastream. Ce sont des données journalières, datant du 28 juin 1974 jusqu'au 28 juin 1999. Dans le tableau 1, nous présentons les statistiques des deux séries.

Tableau 1 : Statistiques du S&P500 et du TSE 300.

Séries	**S&P500**	**TSE300**
Nombre d'observations	6483	6483
Moyenne	0.0534%	0.0423
Écart type	0.9613%	0.7776
Minimum	-22.82%	-11.99
Maximum	8.72%	8.66
Coefficient d'asymétrie	-2.24	-1.04
Coefficient d'aplatissement	54.39	18.78

Nous remarquons que les données sont très variables. Cela se traduit par le fait que les coefficients d'aplatissements des deux séries sont élevés. La distribution est dite leptokurtique. En outre, si nous regardons les coefficients d'asymétrie, nous voyons qu'ils sont différents de zéro, qui est la valeur correspondant à la loi normale.

Les graphiques 1 et 2 permettent d'avoir une vue d'ensemble sur les observations. Les données semblent stationnaires. Nous le vérifions avec un test de Dickey-Fuller (1979). La statistique de DF teste la présence de racine unitaire dans le processus et nous permet ainsi de savoir si les données sont stationnaires.

Soit $y_t = \mu + \varphi y_{t-1} + \varepsilon_t$, où ε_t est i.i.d de moyenne 0 et de variance σ^2, et y_t représente l'une ou l'autre des deux séries.

La statistique de DF permet de tester l'hypothèse nulle que $\varphi = 1$ contre l'hypothèse alternative que. Elle est telle que :

$$t_T = \frac{\hat{\varphi}_T - 1}{\hat{\sigma}_{\hat{\varphi}_T}}, \qquad \text{Où } \hat{\sigma}^2_{\hat{\varphi}_T} = s_T^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} T & \sum y_{t-1} \\ \sum y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

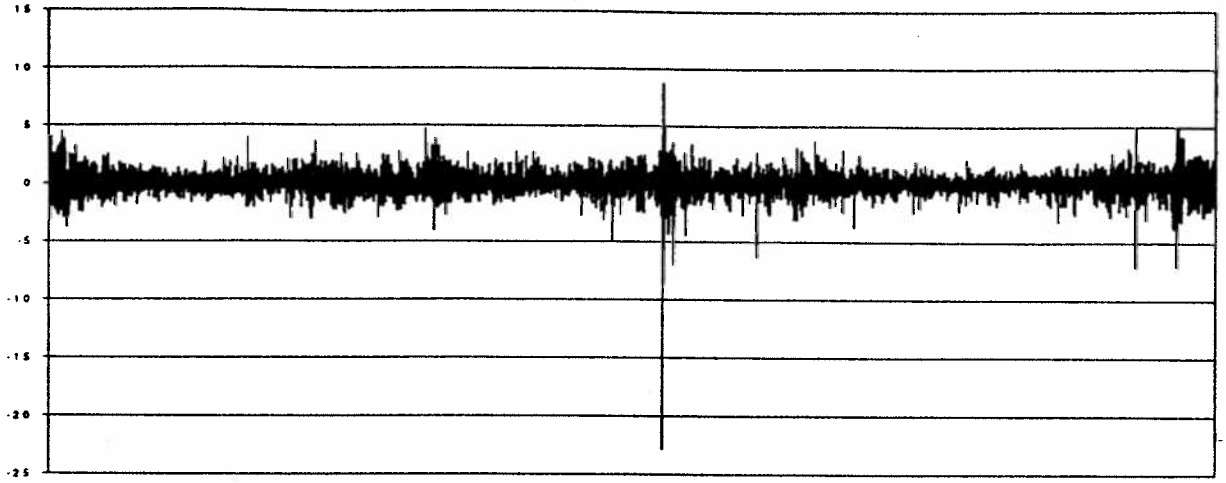
et $s_T^2 = (T-2)^{-1} \sum_{t=1}^{T} (y_t - \hat{\mu}_T - \hat{\varphi}_T y_{t-1})^2$.

Il est facile de montrer que $s_T^2 \xrightarrow{p} \sigma^2$ (voir Hamilton,1994). Sous l'hypothèse nulle, y_t suit une marche aléatoire, et n'est en conséquence pas stationnaire. La valeur critique à 5% de la table de Dickey-Fuller est de –2.86. La statistique t_T pour le S&P500 donne une valeur de –75.46 et pour le TSE une valeur de –65.78[2]. Pour nos deux séries l'hypothèse de racine unitaire est rejetée, puisque la valeur critique est supérieure aux valeurs des deux statistiques. Les données sont donc stationnaires.

[2] Pour être rigoureux nous aurions dû corriger la corrélation avec le test ADF ou celui de Phillips-Perron (PP). Mais la statistique donne des valeurs éloignées de la valeur critique, ce qui laisse supposer que les données seraient stationnaires même en corrigeant la corrélation.

Graphique 1 :

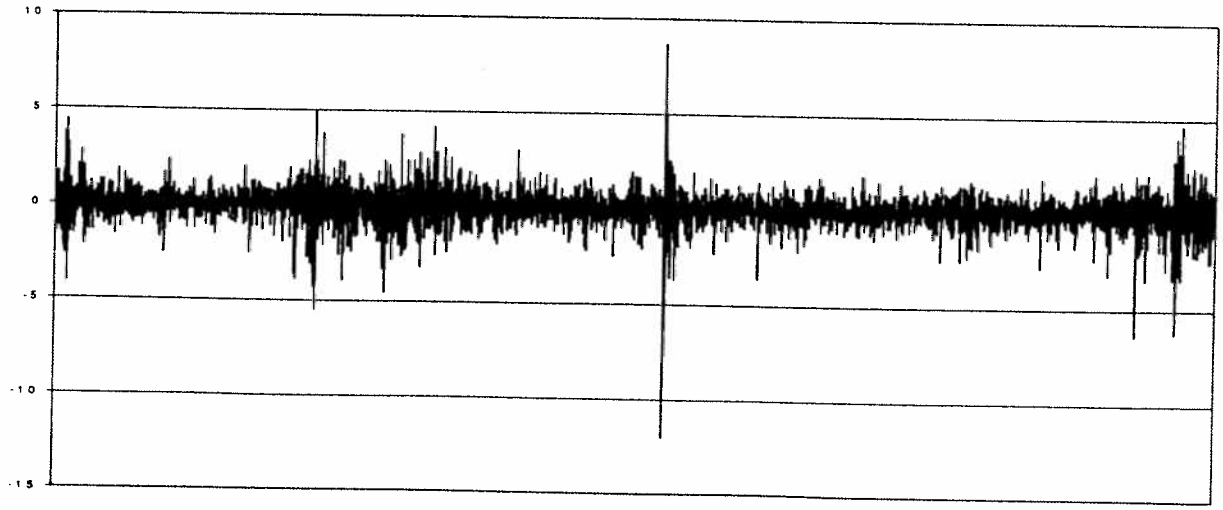
S&P500



28 juin 1974-28 juin 1999

Graphique 2 :

TSE 300



28 juin 1974-28 juin 1999

4) Analyse jointe de la moyenne et de la variance

Nous pouvons considérer plusieurs formes pour la moyenne conditionnelle. Les plus usuelles en Finance sont les formes affines des valeurs passées de y_t, soit des formes AutoRégressives d'ordre p, AR(p). Nous choisissons donc pour caractériser la moyenne le modèle suivant :

$y_t = \mu_t + \varepsilon_t$, avec successivement $\mu_t = b_0 + \sum_{i=1}^{p} b_i y_{t-i}$.

Nous choisissons d'aller jusqu'à cinq retards. Pour chaque modèle, nous maximisons la vraisemblance selon que le carré des innovations est une constante, un ARCH(1), un ARCH(2), un ARCH(3), un ARCH(4), un GARCH(1,1) un GARCH(1,2), un GARCH(2,1) ou un GARCH(2,2). Pour estimer les modèles de type ARCH, nous déterminons l'estimateur du quasi maximum de vraisemblance basé sur l'hypothèse de loi conditionnelle normale. Si $l_t(y;\theta)$ désigne la vraisemblance associée à Y_t et conditionnelle au passé, la fonction de vraisemblance de $Y_1,...,Y_T$ conditionnelle à Y_0 est :

$$L(y;\theta) = \prod_{t=1}^{T} l_t(y;\theta).$$

L'estimateur optimal est celui qui est solution du problème de maximisation suivant :

$$\max_{\theta} \log L(y;\theta)$$

Cet estimateur $\hat{\theta}_T$, appelé estimateur du quasi maximum de vraisemblance, dépend à la fois de la vraie distribution sous-jacente et de celle utilisée pour calculer la vraisemblance qui est ici la loi normale. Sous diverses conditions de régularité, (voir Gallant, 1987,

Gouriéroux-Monfort, 1989, Gouriéroux-Monfort-Trognon, 1984, White, 1981), cet estimateur est convergent même si la véritable loi sous-jacente n'est pas conditionnellement normale, c'est à dire si elle ne correspond pas à la loi utilisée pour calculer la vraisemblance. En outre, cet estimateur est asymptotiquement normal et sa matrice de covariance asymptotique est robuste à l'hétéroscédasticité.

$$V_{as}\left[\sqrt{T}(\hat{\theta}_T - \theta)\right] = J^{-1} I\, J^{-1}$$

avec $J = \underset{0}{E}\left[-\frac{\partial^2 Log\ l_t(Y,\theta)}{\partial\theta\ \partial\theta'}\right]$

et $I = \underset{0}{E}\left[\frac{\partial \log l_t(Y,\theta)}{\partial\theta}\frac{\partial \log l_t(Y,\theta)}{\partial\theta'}\right]$.

L'estimation par le quasi maximum de vraisemblance est la méthode la plus usuellement employée pour les modèles de type ARCH.

Pour sélectionner le meilleur modèle, nous utilisons les deux critères les plus usuels : le critère d'Akaike et celui de Schwartz adaptés aux modèles ARCH. Ils permettent de corriger la fonction de vraisemblance du nombres de paramètres. Le critère d'Akaike est défini de la manière suivante :

AIC = 2*log($\hat{\vartheta}$)-2*k.

Quant au critère de Schwartz, il est tel que :

SIC = 2* log($\hat{\vartheta}$)-k*log(T).

Oú $\log(\hat{\vartheta}) = -\frac{1}{2}\sum_{t=1}^{T}\log(h_t) - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{h_t}$ est la valeur de la fonction de vraisemblance et

h_t est la variance du modèle. Nous calculons les critères précédents indice par indice pour

chaque modèle. Nous trouvons les valeurs de la fonction de vraisemblance, ainsi que les critères d'Akaike et Schwartz dans les annexes 1 et 2 pour le S&P500 et le TSE 300 respectivement. Les deux critères nous amènent à des conclusions différentes. Selon le critère d'Akaike, le meilleur modèle du S&P500 est un AR(5) dont les erreurs suivent un GARCH(2,2). Tandis que le critère de Schwartz préfère un AR(1) dont les erreurs suivent un GARCH(1,2). Nous ne pouvons distinguer quel est le meilleur critère des deux[3]. Les coefficients des modèles sélectionnés sont inscrits dans les tableaux 2 et 3. Nous remarquons que les coefficients 2, 3 et 5 du modèle AR(5) choisi par le critère d'Akaike ne sont pas significatifs. Les coefficients des modèles GARCH(2,2) et GARCH(1,2) sélectionnés pour la variance par les deux critères sont, par contre, significatifs. (Voir section 6, pour les contraintes de positivité des modèles GARCH).

Tableau 2 : AR(5)-GARCH(2,2) pour le S&P500.

$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	$\hat{b}_4$	$\hat{b}_5$
0.0591	0.0835	0.0013	-0.0081	-0.0378	-0.0107
(0.0098)	(0.0147)	(0.0135)	(0.0131)	(0.0132)	(0.0123)

$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.0198	0.0698	0.0498	-0.0357	0.8962
(0.0026)	(0.0020)	(0.0054)	(0.0122)	(0.0087)

Tableau 3 : AR(1)-GARCH(1,2) pour le S&P500.

$\hat{b}_0$	$\hat{b}_1$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0543	0.0827	0.0067	0.1062	-0.0618	0.9487
(0.0098)	(0.0148)	(0.0009)	(0.0036)	(0.0056)	(0.0033)

[3] Nous remarquons, dans la littérature ARCH, qu'une préférence est donnée au critère de Schwartz. Pourtant ceci n'est pas justifié, car la littérature statistique dit qu'il est moins bon si le modèle est mal spécifié.

Pour le TSE 300, également, les deux critères nous amènent à des choix différents. Le premier critère préfère un AR(3) avec des erreurs GARCH(2,2), tandis que le critère de Schwartz sélectionne un AR(1) avec des erreurs GARCH(2,2). Ces estimations sont présentées dans les tableaux 4 et 5. Les modèles choisis par les deux critères sont significatifs, sauf le coefficient 2 du AR(3) choisi par le critère d'Akaike.

Tableau 4 : AR(3)-GARCH(2,2) pour le TSE 300.

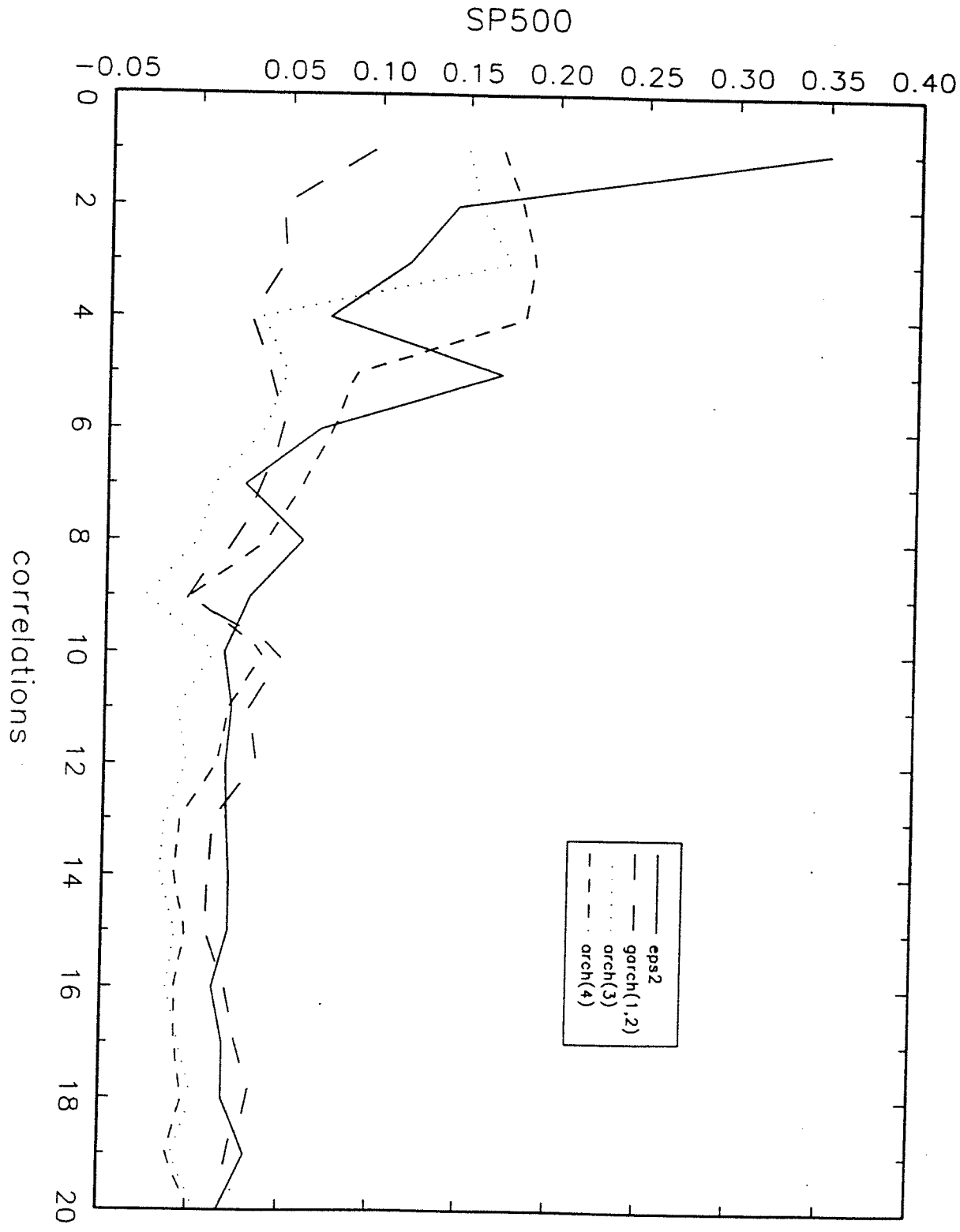
$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.0404	0.2577	-0.0086	0.0334	0.0336	0.1267	0.1211	-0.1429	0.8411
(0.0075)	(0.0146)	(0.0128)	(0.0131)	(0.0027)	(0.0040)	(0.0043)	(0.0069)	(0.0060)

Tableau 5 : AR(1)-GARCH(2,2) pour le TSE 300.

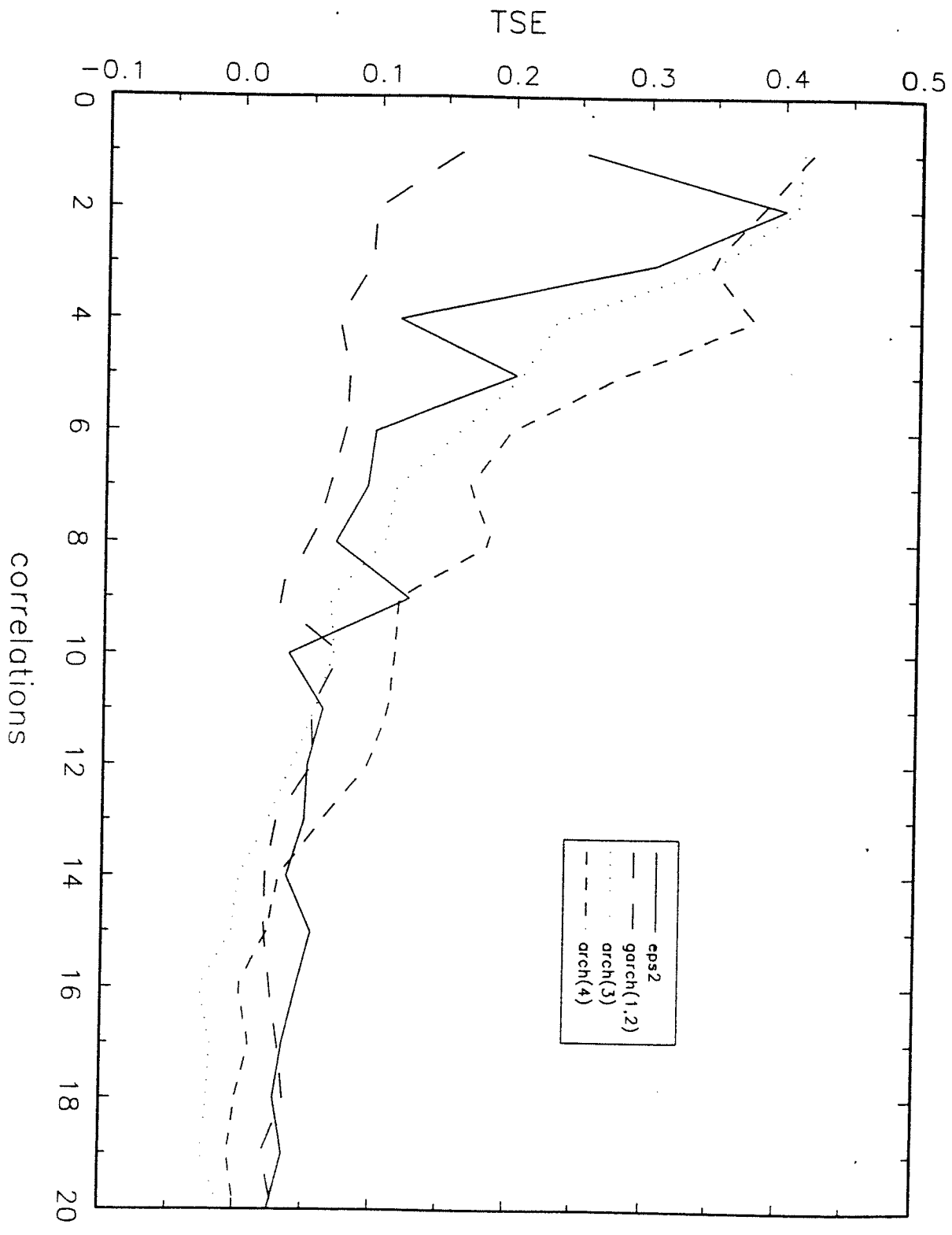
$\hat{b}_0$	$\hat{b}_1$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.0419	0.2573	0.0323	0.1246	0.1181	-0.1355	0.8416
(0.0075)	(0.1380)	(0.0025)	(0.0036)	(0.0043)	(0.0076)	(0.0067)

La corrélation du carré des erreurs peut aider à sélectionner des modèles possibles de la moyenne et de la variance. Pour la moyenne, nous simulons des modèles AR(p), p allant de 1 à 5. Les coefficients utilisés pour la simulation sont ceux résultant de l'estimation précédente par le quasi maximum de vraisemblance. Nous comparons ces simulations aux autocorrélations des données. Ceci ne nous donne aucune indication car les coefficients des différents AR sont trop proches et les simulations sont pratiquement identiques. Pour la variance, nous comparons le corrélogramme des résidus aux carrés des deux séries avec ceux des modèles de type ARCH simulés ; soit un ARCH(1), ARCH(2), ARCH(3), ARCH(4), GARCH(1,1), GARCH(1,2), GARCH(2,1) et GARCH(2,2). Sur les graphiques 3 et 4, nous ne gardons que les modèles ARCH simulés

Graphique 3 :


SP500
−0.05
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0
2
4
6
8
10
12
14
16
18
20
correlations
eps2
garch(1,2)
arch(3)
arch(4)

Graphique 4 :


TSE
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
12
14
16
18
20
correlations
eps2
garch(1,2)
arch(3)
arch(4)

qui sont proches du corrélogramme des résidus des séries, soit un GARCH(1,2), un ARCH(3) et un ARCH(4). Avec le critère de Schwartz nous avons effectivement trouvé un GARCH(1,2) pour le S&P500. Par contre tous les autres choix se portaient sur un GARCH(2,2) qui selon nos simulations, ne semble pas s'approcher du corrélogramme des résidus des deux séries.

La méthode du quasi maximum de vraisemblance, qui est la méthode la plus couramment employée, nous a permis d'arriver à deux possibilités de modèles pour les deux indices. Le critère d'Akaike choisit un AR(5) avec des erreurs GARCH(2,2) pour le S&P500 et un AR(3) avec également des erreurs GARCH(2,2) pour le TSE 300. Tandis que le critère de Schwartz préfère un AR(1) pour la moyenne des deux séries. Les erreurs de la série S&P500 suivent un GARCH(1,2), et celle du TSE 300 un GARCH(2,2). Comme nous ne savons pas lequel des deux critères favoriser, nous allons estimer nos données par une autre méthode.

5) Analyse robuste de la moyenne vis à vis de l'hétéroscédasticité

5.1) La méthode des moments généralisés

Dans la partie précédente, nous avons analysé de façon jointe la moyenne et la variance. Le fait de supposer la forme de la variance peut fausser l'estimation de la moyenne. En effet, si la variance choisie est mal spécifiée alors l'estimation de la moyenne peut être contaminée. Pour cela, nous recommencons l'estimation de la

moyenne avec la méthode des moments généralisés (GMM : General Method of Moments). Cette méthode permet d'estimer de manière séparée la moyenne et la variance. Nous avons ainsi une méthode robuste à la forme de l'hétéroscédasticité.

Soit θ le vecteur des coefficients à estimer. Supposons que θ_0 est la vraie valeur de θ, et que cette vraie valeur est caractérisée par la propriété : $E(f_t(\theta_0)) = 0$. Nous appellons aussi l'équation précédente les conditions d'orthogonalité. Si $E(y_t / I_{t-1}) = \mu_t(\theta)$, alors $E(z_t(y_t - \mu_t(\theta))) = E(f_t(\theta))$, où z_t représente les variables instrumentales. Nous faisons une approximation de $E(f_t(\theta))$ par $g_T(\theta) = T^{-1}\sum_{t=1}^{T} f_t(\theta)$. L'idée des GMM est de choisir θ de manière à ce que $g_T(\theta)$ soit aussi proche que possible de zéro. Si le nombre de paramètres (k) à estimer est égal au nombre de conditions d'orthogonalité (n), ce qui est le cas pour la méthode des moindres carrés ordinaires (MCO), alors $\hat{\theta}_T$ est la valeur de θ qui est solution de $g_T(\theta) = 0$. Tandis que s'il y a plus de conditions que de paramètres (n > k), $\hat{\theta}_T$ est la valeur de θ qui minimise la forme quadratique $Q_T(\theta) \equiv g_T(\theta)' W_T g_T(\theta)$, où W_T est une matrice définie positive qui satisfait $\lim_{T\to\infty} W_T = W_0$, avec probabilité un, et avec W_0 une matrice définie positive. Les deux matrices W_T et W_0 sont appelées matrice de poids. La proximité de zéro du $i^{ème}$ élément de $g_T(\theta)$ dépend du poids donné à la $i^{ème}$ condition d'orthogonalité par la matrice de poids. Soit S la variance asymptotique de $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} f_t(\theta_0)$ (voir Hamilton 1994) : $S = \lim_{T\to\infty} T.E\{(g_T(\theta_0))(g_T(\theta_0))'\}$.

La valeur optimale de la matrice W_T, dans la fonction objective, est donnée par S^{-1}, l'inverse de la matrice de variance asymptotique. Si le vecteur $f_t(\theta_0)$ n'est pas corrélé, nous pouvons calculer S de la manière suivante :

$$\hat{S}_T = (1/T)\sum_{t=1}^{T}(f_t(\hat{\theta}_T))(f_t(\hat{\theta}_T))' \xrightarrow{p} S .$$

Notons qu'avant d'estimer θ nous avons besoin de connaître la matrice S, et nous savons que celle-ci dépend de θ. Nous allons donc procéder en plusieurs étapes. Tout d'abord, nous estimons une valeur initiale de $\hat{\theta}_T^{(0)}$ avec une matrice de poids arbitraire tel que par exemple $W_T = I_n$, où I_n est une matrice identité. Cette estimation de θ est, ensuite, utilisée pour estimer une valeur initiale $\hat{S}_T^{(0)}$. La fonction objective est donc minimisée avec $W_T = (\hat{S}_T^{(0)})^{-1}$ et donne un nouvel estimateur $\hat{\theta}_T^{(1)}$. Nous continuons l'itération du processus jusqu'à ce que $\hat{\theta}_T^{(j)} \cong \hat{\theta}_T^{(j+1)}$.

5.2) Estimations

5.2.1) La méthode des moments généralisés juste-identifié

Nous régressons, tout d'abord, le modèle par la méthode des moindres carrés ordinaires qui est la méthode la plus classique et la plus simple. On a vu précédemment que les GMM juste-identifiés sont équivalents aux MCO, la différence étant qu'avec les GMM les écarts types sont robustes à l'hétéroscédasticité. Nous savons également que le nombre d'instruments est égal au nombre de paramètres lorsque la méthode des GMM est juste identifiée. Les instruments z_t sont égaux aux variables explicatives, et donc dans

notre cas à des variables endogènes retardées. Par exemple pour un AR(1), $z_t = (1, y_{t-1})$. Les résultats pour le S&P500 et le TSE 300 sont donnés dans l'annexe 4. Nous remarquons, pour le S&P500, que seule la constante est significative. Pour le TSE 300, les coefficients sont significatifs seulement jusqu'au AR(1). Ces résultats ne sont pas très satisfaisants, nous allons donc essayer de les améliorer.

5.2.2) La méthode des moments généralisés sur-identifié

Pour améliorer l'estimation précédente, nous augmentons le nombre d'instruments. Quand le nombre de conditions d'orthogonalité excède le nombre de paramètres à estimer, le modèle est dit sur-identifié. Nous utilisons pour chaque modèle trois instruments de plus que de paramètres. Par exemple, pour un AR(1), la variable instrumentale est $z_t = (1, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4})$. Les résultats de l'estimation du S&P500 et du TSE 300 se trouvent dans l'annexe 5. Nous constatons que l'estimation ne s'est pas beaucoup améliorée, (voir tableaux 6 et 7), les écarts types ne diminuent que légèrement. Pour le S&P500, la constante est significative, le retard 1 dans les modèles AR(4) et AR(5) l'est également. Tandis que pour le TSE 300, la constante ainsi que les coefficients 1 et 3 sont significatifs.

5.2.3) Suridentification et correction de l'hétéroscédasticité

Les résultats obtenus par les GMM juste-identifiés et sur-identifiés ne nous satisfont pas. Nous voulons avoir une estimation encore plus précise, c'est à dire que

nous voulons que les écarts-types soient plus efficaces. Pour cela, nous utilisons de nouveau les GMM en essayant de corriger l'hétéroscédasticité de manière similaire à la méthode des Moindres Carrés Généralisés (MCG), car celle-ci est habituellement utilisée quand les données contiennent de l'hétéroscédasticité. La méthode des MCG consiste à diviser le modèle par la variance conditionnelle. Dans notre cas nous avons, $y_t = \mu_t + \varepsilon_t$. Or nous avons vu que $Var(\varepsilon_t / I_{t-1}) = Var(\sqrt{h_t} u_t / I_{t-1}) = h_t$, où I_{t-1} est l'information passée. La variance conditionnelle n'est pas forcément égale à h_t, mais c'est une supposition valable. Nous choisissons, parmi les instruments admis, un GARCH(1,1) pour spécifier h_t. Avec les résidus estimés de chacun des modèles régressés par les GMM (voir annexe 5), Nous estimons des GARCH(1,1) par la méthode du quasi maximum de vraisemblance, que nous notons $\hat{h}_t$. Ainsi, nous utilisons la méthode des GMM en divisant chaque instrument par $\hat{h}_t$. Par exemple pour un AR(1), nous avons $z_t = (\frac{1}{\hat{h}_t}, \frac{y_{t-1}}{\hat{h}_t}, \frac{y_{t-2}}{\hat{h}_t}, \frac{y_{t-3}}{\hat{h}_t}, \frac{y_{t-4}}{\hat{h}_t})$. Nous noterons, désormais, cette estimation GMM1. Dans les tableaux 6 et 7, nous pouvons voir que l'estimation des GMM n'améliore que légèrement les résultats, en effet l'estimation des GMM et des MCO donnent des résultats très proches. Par contre la méthode des GMM1 améliore nettement l'estimation. C'est l'estimation la plus précise. En effet, les écarts types sont les plus petits des trois estimations. C'est donc à partir de cette estimation, que nous sélectionnons le meilleur modèle. Dans l'annexe 6 qui correspond à l'estimation du S&P500 par les GMM1, nous constatons que les coefficients des AR(2) et AR(3) ne sont pas significatifs, mais que celui du AR(4) l'est. De plus, lorsque nous ajoutons un retard 5 au modèle, il reste stable.

Le coefficient 5 n'apporte donc rien au modèle. Par ailleurs, la suridentification du AR(4) est bonne puisque la p-value du J test (voir la section sur les tests) est supérieure à 0.05. Nous sélectionnons donc pour la moyenne du S&P500 un modèle AR(4). Quant au TSE 300, nous voyons dans l'annexe 6 que le coefficient du deuxième retard n'est pas significatif. Nous remarquons également que celui du quatrième et du cinquième retards ne le sont pas non plus, et que l'ajout de ces variables n'apporte rien au modèle. Le J test rejette le BB et le AR(2). Nous sélectionnons donc pour le TSE 300 un modèle AR(3). Nous remarquons aussi que nous avons supposé que la variance était un GARCH(1,1) et cela a amélioré le modèle. (Nous n'avons pas eu besoin d'un GARCH(2,2), qui est le meilleur modèle sélectionné pour la variance, par le quasi maximum de vraisemblance.) Ceci renforce l'idée que la variance est mal spécifiée et donc que la moyenne estimée par le quasi maximum de vraisemblance peut être contaminée. Si nous comparons maintenant les GMM1 avec les modèles sélectionnés par le quasi maximum de vraisemblance, dans les tableaux 2 et 3 pour le S&P500 et 4 et 5 pour le TSE 300, nous remarquons que les écarts types sont très proches, et que pourtant les coefficients sont différents. Nous en déduisons que nous ne pouvons pas avoir des estimateurs aussi efficaces avec les GMM qu'avec la méthode standard, et qui sont en plus robustes à la forme de l'hétéroscédasticité. Il est donc plus avantageux d'utiliser la méthode des GMM. En effet, l'estimation de la moyenne ne risque pas d'être contaminée par une variance mal spécifiée.

Par ailleurs, nous constatons que le critère d'Akaike sélectionne pour le S&P500 un AR(5) comme étant le meilleur modèle alors que le coefficient du retard 5 n'est pas significatif. Il faut également noter que la méthode des moments généralisés nous amène

à un AR(4) pour le S&P500. De plus, pour le TSE 300 le critère d'Akaike a préféré un AR(3), et c'est aussi la sélection que nous faisons avec les GMM1. Il semble donc, que ce dernier est plus fiable que le critère de Schwartz qui préfère un AR(1) pour modéliser la moyenne des deux séries de données. Ceci est en accord avec la littérature statistique (voir note 3 p11). Pourtant, cette constatation est à relativiser, puisque nous ne comparons pas les mêmes estimations : dans un cas, l'estimation est jointe, alors que dans l'autre elle est séparée.

Tableau 6 : Estimations du S&P500.

Méthode d'estimation	$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	$\hat{b}_4$
MCO	0.0547	0.0651	-0.0267	-0.0121	-0.0302
	(0.0133)	(0.0366)	(0.0345)	(0.0302)	(0.0183)
GMM	0.0510	0.0822	-0.0264	-0.0041	-0.0318
	(0.0130)	(0.0359)	(0.0330)	(0.0290)	(0.0167)
GMM1	0.0418	0.1066	0.0050	-0.0068	-0.0296
	(0.0100)	(0.0143)	(0.0135)	(0.0133)	(0.0132)

Tableau 7 : Estimations du TSE 300.

Méthode d'estimation	$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$
MCO	0.0334	0.2040	-0.0256	0.0384
	(0.0103)	(0.0324)	(0.0377)	(0.0315)
GMM	0.0334	0.2030	-0.0340	0.0530
	(0.0100)	(0.0322)	(0.0346)	(0.0265)
GMM1	0.0265	0.2966	-0.0046	0.0342
	(0.0075)	(0.0148)	(0.0147)	(0.0138)

5.3) Les tests

5.3.1) Le J test de Hansen

Quand le nombre de conditions d'orthogonalité excède le nombre de paramètres à estimer, le modèle est sur-identifié. On peut tester la validité de la suridentification avec la statistique J : J = $TQ_T(\hat{\theta})$, où $Q_T(\hat{\theta}) = g(\hat{\theta})'W_T g(\hat{\theta})$, est la valeur optimale de la fonction objective. Hansen (1982) montre que T fois la valeur minimale de la fonction objective suit une distribution asymptotique $\chi^2(h-k)$. Où h est le nombre d'instruments et k le nombre de paramètres. La p-value est la probabilité d'accepter le modèle. Ce test a tendance à accepter trop facilement les modèles. C'est pour cela que nous utilisons deux autres statistiques.

5.3.2) Le test de Box-Pierce

S'il n'existe pas de corrélation entre les résidus du modèle, celui-ci est bien spécifié. Nous utilisons la statistique de Box-Pierce (1970) pour savoir si les résidus sont « blanchis ».

Soit $r_k = \dfrac{\sum_{t=k+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-k}}{\sum_{t=1}^{T} \hat{\varepsilon}_t^2}$ la fonction d'autocorrélation, avec $r_k = [r_1, ..., r_K]'$.

Sous l'hypothèse nulle que les résidus suivent un bruit blanc, la statistique est la suivante :

$$Q_K = T\sum_{k=1}^{K} r_k^2 .$$

Sous cette hypothèse, Q_K suit asymptotiquement une $\chi^2(K-(p+q))$, où p est l'ordre du processus autorégressif et q est le degré moyenne mobile. Cependant cette statistique n'est pas valide lorsqu'il y a présence d'erreurs conditionnellement hétéroscédastiques dans les données.

5.3.3) Le l test de Cumby et Huizinga

L'hétéroscédasticité peut fausser un test sur la corrélation lorsqu'on l'ignore. Pour cela, un autre test, le l test, adapté aux modèles dont les erreurs sont conditionnellement hétéroscédastiques, est employé. Ce test permet au modèle de contenir des erreurs qui suivent une moyenne mobile d'ordre q ≥ 0. Nous suivons Cumby et Huizinga (1992) qui montrent comment une estimation convergente de la matrice de variance-covariance des résidus peut être utilisée pour tester l'hypothèse nulle que les autocorrélations d'ordre supérieur à q sont égales à zéro, contre l'hypothèse alternative que la corrélation des erreurs est différente de zéro quand les retards sont supérieurs à q≥ 0.

Soit $y_t = X_t\delta + \varepsilon_t$. X_t est un vecteur (1×k) de variables explicatives.

δ est un vecteur (k×1) de paramètres inconnus.

Z_t est un vecteur (h×1) de variables instrumentales.

Nous supposons que : $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ et $E(\varepsilon_t\varepsilon_{t-n})/\sigma_\varepsilon^2 = \rho_n$.

Le test s'écrit, $\begin{cases} H_0 : \rho = (\rho_{q+1},\ldots,\rho_{q+s}) = 0 \\ H_a : \rho \neq 0 \end{cases}$

Soit $\hat{V}_r, \hat{B}, \hat{V}_d, \hat{C}$ *et* $\hat{D}$ des estimations convergentes de V_r, B, V_d, C *et* D (voir annexe 7).

La statistique s'écrit : $l_{q,s} = T\hat{r}'\left[\hat{V}_r + \hat{B}\hat{V}_d\hat{B}' + \hat{C}\hat{D}'\hat{B}' + \hat{B}\hat{D}\hat{C}'\right]^{-1}\hat{r}$.

Sous l'hypothèse nulle, $l_{q,s} \xrightarrow{A} \chi^2(s)$.

Dans le cas où les variables instrumentales sont des variables endogènes retardées, ce qui est le cas ici, l'hypothèse nulle est une hypothèse jointe concernant la corrélation des résidus et la validité des instruments. Ce test peut donc être une alternative à la statistique J proposée par Hansen.

La méthode des moments généralisés permet donc une estimation séparée de la moyenne et de la variance. Ceci amène à une estimation de la moyenne non contaminée par la forme de la variance, alors que cela pourrait être le cas avec l'estimation jointe de la moyenne et de la variance par le quasi maximum de vraisemblance. Les rendements sont estimés par la méthode des GMM suridentifiés avec une correction de l'hétéroscédasticité. L'estimation et les tests nous amènent à choisir un AR(4) pour le S&P500 et un AR(3) pour le TSE 300. Il nous faut, ensuite, estimer le risque des deux séries.

6) Analyse de la variance

Nous voulons maintenant modéliser la variance conditionnelle pour chacune des deux séries. Avec les travaux empiriques déjà faits sur le sujet, nous nous aperçevons que les données financières exigent des modèles ARCH avec des retards relativement longs.

De plus, si nous regardons les modèles sélectionnés lors de l'estimation jointe de la moyenne et de la variance, nous voyons que la variance conditionnelle retenue suit un GARCH dont le coefficient β est élevé. Tout comme pour la modélisation ARMA, le β est le coefficient de la moyenne mobile. Or si celui-ci est grand, nous savons que l'ordre p du processus autorégressif AR(p) sera élevé. Si nous faisons le parallèle avec le modèle ARCH(q), nous en déduisons qu'il y aura beaucoup de paramètres à estimer avec celui-ci. Nous préfèrons donc estimer des modèles GARCH(p,q). Pour cela, on se sert des erreurs du modèle sélectionné pour la moyenne, par la méthode des moments généralisés corrigée de l'hétéroscédasticité (GMM1).

Soit pour le S&P500 :

$$\varepsilon_{1t} = y_{1t} - 0.0418 - 0.1066 y_{1t-1} - 0.0050 y_{1t-2} + 0.0068 y_{1t-3} + 0.0296 y_{1t-4}.$$

Et pour le TSE 300 :

$$\varepsilon_{2t} = y_{2t} - 0.0265 - 0.2966 y_{2t-1} + 0.0046 y_{2t-2} - 0.0342 y_{2t-3}.$$

Le modèle GARCH(p,q) est le suivant :

$$h_{jt} = \omega_j + \sum_{i=1}^{q} \alpha_{ji} \varepsilon_{jt-i}^2 + \sum_{i=1}^{p} \beta_{ji} h_{jt-i} \quad \text{avec j=1,2.}$$

Nous estimons ce modèle par la méthode du quasi maximum de vraisemblance. Mais cette fois l'estimation n'est pas jointe puisque nous estimons la variance seule à partir des résidus présentés ci-dessus. La valeur de la fonction de vraisemblance pour les GARCH estimés est inscrite dans le tableau 8, la valeur de du critère d'Akaike et du critère de Schwartz dans les tableaux 9 et 10 respectivement. Les nombres qui sont en gras représentent les valeurs optimales.

Tableau 8 : Valeurs de la fonction de vraisemblance pour le S&P500 et le TSE 300.

Valeur de la vraisemblance	S&P500	TSE 300
GARCH(1,1)	-2124.80	-395.93
GARCH(2,1)	-2119.22	-392.60
GARCH(1,2)	-2114.40	-388.61
GARCH(2,2)	-2119.10	-394.06

Tableau 9 : Valeurs du critère d'Akaike pour le S&P500 et le TSE 300.

AIC	S&P500	TSE 300
GARCH(1,1)	-4249.60	-797.86
GARCH(2,1)	-4246.44	-793.20
GARCH(1,2)	-4236.80	-785.22
GARCH(2,2)	-4248.20	-798.12

Tableau 10 : Valeurs du critère de Schwartz pour le S&P500 et le TSE 300.

SIC	S&P500	TSE 300
GARCH(1,1)	-4261.03	-803.29
GARCH(2,1)	-4253.68	-800.44
GARCH(1,2)	-4244.04	-792.46
GARCH(2,2)	-4257.25	-807.17

Le modèle retenu, selon tous les critères, pour les deux séries est un GARCH(1,2). Ces modèles sont présentés dans les tableaux 11 et 12. Nous remarquons que les coefficients du modèle choisi pour les deux séries sont significatifs. De plus, Les contraintes de positivité sont respectées. En effet si nous suivons Nelson et Cao (1992), un modèle GARCH(1,2) est positif si les contraintes $\begin{cases} a)\ \omega \geq 0 \\ b)\ 0 \leq \beta_1 < 1 \\ c)\ \beta_1\alpha_1 + \alpha_2 \geq 0 \\ d)\ \alpha_1 \geq 0 \end{cases}$ sont réalisées. Ce qui est le cas ici.

Tableau 11 : GARCH(1,2) pour le S&P500.

$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0069	0.1053	-0.0602	0.9477
(0.0010)	(0.0036)	(0.0056)	(0.0034)

Tableau 12 : GARCH(1,2) pour le TSE.

$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0118	0.1658	-0.0724	0.8874
(0.0010)	(0.0049)	(0.0070)	(0.0053)

Notons que dans les études déjà effectuées sur la modélisation de la variance conditionnelle d'actifs financiers, le modèle GARCH(1,2) est celui que l'on retrouve le plus souvent pour des données journalières et de longues périodes.

Nous avons estimé le risque avec la méthode du quasi maximum de vraisemblance, en employant des modèles GARCH. Cette estimation n'est pas jointe avec la moyenne car nous avons utilisé les résidus des modèles de rendements sélectionnés, soit un AR(4) et un AR(3) pour le S&P500 et le TSE 300 respectivement. Nous trouvons que la variance conditionnelle des deux indices suit un GARCH(1,2). Si nous comparons ces résultats avec l'estimation jointe par le quasi maximum de vraisemblance vue à la section 4, nous remarquons que les modèles choisis sont différents. (En effet, le modèle le plus sélectionné pour la variance de l'estimation jointe est un GARCH(2,2), sauf le critère de Schwartz pour le S&P500 qui préfère un GARCH(1,2)). La différence des ordres p et q dans le choix des modèles GARCH est peut être due au fait que nous n'utilisons pas la même moyenne. Nous comparons, maintenant, les deux estimations jointe et séparée avec la même moyenne et la même variance (voir les tableaux 13 et 14). Nous avons déjà vu

que les deux estimations de la moyenne ne donnaient pas exactement les mêmes résultats, il est donc normal que les estimations de la variance ne soient pas identique. Cependant elles sont quand même proches.

Tableau 13 : AR(4)-GARCH(1,2) pour le S&P500.

$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	$\hat{b}_4$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0569	0.0820	0.0002	-0.0053	-0.0384	0.0069	0.1044	-0.0590	0.9474
(0.0098)	(0.0148)	(0.0130)	(0.0127)	(0.0132)	(0.0009)	(0.0043)	(0.0063)	(0.0035)

Tableau 14 : AR(3)-GARCH(1,2) pour le TSE 300.

$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0380	0.2554	-0.0122	0.0376	0.0123	0.1660	-0.0697	0.8836
(0.0075)	(0.0147)	(0.0132)	(0.0129)	(0.0011)	(0.0065)	(0.0085)	(0.0057)

7) Analyse bivariée jointe de la moyenne et de la variance : présentation générale

7.1) Modèle VAR et causalité

Jusqu'ici nous avons employé pour la modélisation de la moyenne univariée des modèles autorégressifs. Si nous passons à un contexte multivarié, nous pouvons utiliser des Vecteurs AutoRégressifs d'ordre p, VAR(p). Ils s'écrivent de la façon suivante :

$$y_t = b + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \ldots + \Phi_p y_{t-p} + \varepsilon_t$$

avec b un vecteur (n×1) où n est le nombre de composantes et Φ_j est une matrice (n×n) de coefficients autorégressifs avec j = 1,2,...p. Le vecteur ε_t est de dimension (n×1) et il est tel que : $E(\varepsilon_t) = 0, \quad E(\varepsilon_t \varepsilon_\tau^{'}) = \begin{cases} \Omega & pour\ t = \tau \\ 0 & autrement \end{cases}$

avec Ω une matrice (n×n) symétrique et définie positive.

Dans notre étude nous nous limitons au cas bivarié puisque nous avons deux indices. Pour avoir une idée d'un modèle bivarié, nous simplifions à un VAR(1). Nous obtenons :

$$y_t = \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix}, \quad b = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \quad \Phi = \begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Cette modélisation va nous permettre de voir le lien qui existe entre le taux de rendement américain et le taux de rendement canadien. En effet, les modèles VAR peuvent être utilisés pour tester la causalité au sens de Granger (1980). Dans le cas bivarié, nous considère deux ensembles d'informations, $I_t : y_{1t-j}, j \geq 0$ *et* $J_t : y_{1t-j}, y_{2t-j}, j \geq 0$. y_{2t} cause y_{1t+1} en moyenne ($y_{2t} \xrightarrow{G} y_{1t+1}$) si $E[y_{1t+1} / I_t] \neq E[y_{1t+1} / J_t]$.

En pratique, si nous prennons l'exemple du VAR(1) bivarié décrit ci dessus, y_{1t} ne cause pas y_{2t} si b_{21} est significativement égal à zéro, et y_{2t} ne cause pas y_{1t} si b_{12} est significativement égale à zéro. Notons que y_{1t} peut causer y_{2t} sans que y_{2t} ne cause y_{1t}. La causalité permet de voir si l'information passe entre les deux variables. Plus le coefficient est élevé et plus l'information passe rapidement.

7.2) GARCH bivarié

7.2.1) Modélisation

Lors de la modélisation de la variance de manière univarié nous avons modélisé un GARCH(p,q). Nous voudrions faire la même chose de façon bivariée, mais nous allons voir que cela est beaucoup plus difficile car les modèles GARCH bivariés ont beaucoup trop de paramètres à estimer. En théorie, un processus GARCH multivarié a le même type de formulation pour chaque élément de la matrice de variance-covariance conditionnelle qu'un GARCH univarié. Si les composantes du processus sont notées $(\varepsilon_{lt}, l = 1,...,n)$, et si $H_t = Var(\varepsilon_t / \varepsilon_{t-1})$ a pour éléments $h_{k\,l\,t}$, nous pouvons écrire :

$$\begin{cases} E(\varepsilon_t / \varepsilon_{t-1}) = 0, Var(\varepsilon_t / \varepsilon_{t-1}) = H_t, \\ h_{k\,l\,t} = C_{k\,l} + \sum_{i=1}^{q}\left[\sum_{k',l'} \alpha_{k\,l\,k'\,l'\,i}\varepsilon_{k',t-i}\varepsilon_{l',t-i}\right] + \sum_{i=1}^{p}\left(\sum_{k',l'} \beta_{k\,l\,k'\,l'} h_{k',l',t-i}\right) \end{cases}$$

Les divers coefficients $c, \alpha \; et \; \beta$ peuvent être contraints sans perte de généralité, en notant que les matrices H_{t-i} et $\varepsilon_{t-i}\,\varepsilon_{t-i}^{'}$ sont symétriques. Ces contraintes sont :

$$\begin{cases} C_{k\,l} = C_{l\,k} \\ \alpha_{k\,l\,k'\,l'\,i} = \alpha_{l\,k\,k'\,l'\,i}, \alpha_{k\,l\,k'\,l'\,i} = \alpha_{k\,l\,l'\,k'\,i}, \\ \beta_{k\,l\,k'\,l'\,i} = \beta_{l\,k\,k'\,l'\,i}, \beta_{k\,l\,k'\,l'\,i} = \beta_{k\,l\,l'\,k'\,i}. \end{cases}$$

Prenant en compte ces contraintes de symétrie, nous pouvons noter que le modèle dépend de $\frac{n(n+1)}{2} + (p+q)\left[\frac{n(n+1)}{2}\right]^2$ paramètres indépendants. Par exemple, pour un GARCH(1,1) si le nombre de composantes passe de 1 à 2, le nombre total de paramètres

passe de 3 à 21. Nous constatons que le nombre de paramètres à estimer augmente considérablement lorsque nous passons d'un contexte univarié à un contexte multivarié.

7.2.2) contraintes de positivité et conditions de stabilité

Pour définir les propriétés d'un GARCH multivarié, nous pouvons réécrire sous forme matricielle :

$$vech(H_t) = vech\, C + \sum_{i=1}^{q} \underline{A_i} vech(\varepsilon_{t-i}\varepsilon_{t-i}^{'}) + \sum_{i=1}^{p} \underline{B_i} vech(H_{t-i}),$$

où $vech(H_t)$ désigne l'opérateur consistant à présenter sous forme vectorielle la portion inférieure de la matrice symétrique H et où $\underline{A_i}, \underline{B_i}$ sont les matrices adéquates, dont les éléments sont respectivement fonctions des α et β. Nous pouvons l'écrire autrement, sans vectorialiser la matrice de covariance H_t. Soient les matrices A_{kli}, B_{kli} d'éléments :

$\alpha_{klk'l'i}, k' = 1,...,n, l' = 1,...,n.$
$\beta_{klk'l'i}, k' = 1,...,n, l' = 1,...,n.$

respectivement, et les matrices A_i et B_i de taille $n^2 \times n^2$ admettant les éléments blocs A_{kli}, B_{kli}. On a :

$$h_{kl,t} = C_{kl} + \sum_{i=1}^{q} \varepsilon_{t-i}^{'} A_{kli} \varepsilon_{t-i} + \sum_{i=1}^{p} E\left[\varepsilon_{t-i}^{'} B_{kli} \varepsilon_{t-i} / \varepsilon_{t-i-1}\right].$$

Introduisons le produit tensoriel $\otimes$ de matrices, qui à deux matrices C et D fait correspondre la matrice C $\otimes$ D, admettant les éléments blocs $C_{ij}D$. On a :

$$H_t = C + \sum_{i=1}^{q} (Id_n \otimes \varepsilon_{t-i}^{'}) A_i (Id_n \otimes \varepsilon_{t-i}) + \sum_{i=1}^{p} E\left[(Id_n \otimes \varepsilon_{t-i}^{'}) B_i (Id_n \otimes \varepsilon_{t-i}) / \varepsilon_{t-i-1}\right].$$

Cette équation de récurrence de la matrice de covariance conduit à une solution symétrique. Mais il faut également que H_t soit définie positive. Les conditions suffisantes pour la positivité sont que les matrices $C, A_i, i = 1,..., q, B_i, i = 1,..., p$, sont définies positives.

Nous contraignons aussi le modèle a être stable. Pour cela nous introduisons le vecteur des innovations sur le carré du processus :

$u_t = vech(\varepsilon_t \varepsilon_t^{'}) - vech(H_t)$.

Si nous remplacons dans l'écriture vectorielle vue plus haut $vech\ H_t$ par son expression en fonction de ces innovations, nous obtenont :

$$vech(\varepsilon_t \varepsilon_t^{'}) = vech\ C + \sum_{i=1}^{Max(p,q)} (A_i + B_i) E\ vech(\varepsilon_{t-i} \varepsilon_{t-i}^{'}) + u_t - \sum_{i=1}^{p} B_i u_{t-i}\ .$$

Nous avons vu que le processus ε_t est centré et non corrélé car $E(\varepsilon_t / \varepsilon_{t-1}) = 0$. Il sera stationnaire à l'ordre deux, si sa variance peut être prise indépendante du temps. Or nous déduisons de l'équation précédente que :

$$E\ vech(\varepsilon_t \varepsilon_t^{'}) = vech\ C + \sum_{i=1}^{Max(p,q)} (A_i + B_i) E\ vech(\varepsilon_{t-i} \varepsilon_{t-i}^{'})\ .$$

Cette équation admet une solution convergente vers une limite constante, lorsque les racines de l'équation :

$\det\left[Id - \sum_{i=1}^{Max(p,q)} (A_i + B_i) z^i \right] = 0$, sont strictement à l'extérieur du disque unité.

Cette condition est suffisante, mais non nécessaire. La matrice de covariance non conditionnelle est alors donnée par :

$$vech\ E(\varepsilon_t\varepsilon_t') = \left[Id - \sum_{i=1}^{Max(p,q)} (A_i + B_i) \right]^{-1} vech\ C\ .$$

7.2.3) Modèles contraints

Les modèles décrits précédemment sont difficilement utilisables car ils ont beaucoup trop de paramètres à estimer. Dans la littérature, diverses contraintes ont été imposées pour réduire le nombre de paramètres, comme par exemple les modèles diagonaux (Bollerslev-Engle-Wooldridge (1988)), où chaque élément de la matrice de covariance $h_{jk,t}$ ne dépend que de son propre passé ou du passé de $\varepsilon_{j,t}\varepsilon_{k,t}$. De ce fait, ce modèle ne peut nous aider à détecter la causalité. Le modèle BEKK (Baba-Engle-Kraft-Kroner (1987)) semble le plus approprié à notre étude puisqu'il tient compte de la corrélation entre les variables. Dans le cas d'un GARCH(1,1) bivarié, le modèle BEKK est le suivant :

$$H_t = C'C + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}' \begin{bmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1}\varepsilon_{2,t-1} \\ \varepsilon_{2,t-1}\varepsilon_{1,t-1} & \varepsilon_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}' H_{t-1} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Nous pouvons réécrire, si nous enlèvons l'indice de temps et les termes GARCH :

$$h_{11} = c_{11} + a_{11}^2\varepsilon_1^2 + 2a_{11}a_{21}\varepsilon_1\varepsilon_2 + a_{21}^2\varepsilon_2^2$$
$$h_{12} = c_{12} + a_{11}a_{12}\varepsilon_1^2 + (a_{21}a_{12} + a_{11}a_{22})\varepsilon_1\varepsilon_2 + a_{21}a_{22}\varepsilon_2^2$$
$$h_{22} = c_{13} + a_{12}^2\varepsilon_1^2 + 2a_{12}a_{22}\varepsilon_1\varepsilon_2 + a_{22}^2\varepsilon_2^2.$$

Nous estimons un GARCH-BEKK(1,1) bivarié par la méthode du quasi maximum de vraisemblance, mais il ne converge pas. Nous savons que ceci est courant car il est

difficile d'estimer des GARCH bivariés. (En plus nous manquons de temps). Nous cherchons donc une autre manière de simplifier l'estimation.

8) Analyse bivariée jointe de la moyenne et de la variance : modèle particulier

Comme nous l'avons vu, l'estimation de la variance bivariée est difficile. Pour simplifier le modèle, nous estimons un VAR séparé de manière à pouvoir estimer la variance conditionnelle de chacun des deux indices. (Ainsi, nous estimons de façon jointe la moyenne de manière bivariée et la variance de façon univariée par la méthode du quasi maximum de vraisemblance). Soit y_{1t} représente l'indice S&P500 et y_{2t} l'indice TSE 300. Nous commencons par estimer le VAR(p) séparé suivant : $y_{1t} = b_{10} + \sum_{j=1}^{p} b_{1j} y_{1t-j} + \sum_{j=1}^{p} b_{2j} y_{2t-j}$, selon que sa variance soit un modèle ARCH(q) avec q allant de 1 à 4, ou un GARCH(p,q) avec p et q allant de 1 à 2. Puis nous estimons l'autre partie du VAR : $y_{2t} = b_{20} + \sum_{j=1}^{p} b_{1j}^{*} y_{1t-j} + \sum_{j=1}^{p} b_{2j}^{*} y_{2t-j}$ en supposant l'une après l'autre les mêmes variances que précédemment. Comme dans le cas univarié, nous calculons les critères de Schwartz et d'Akaike de manière à sélectionner le meilleur modèle. Les valeurs sont inscrites dans les annexes 8 et 9 pour le S&P500 et le TSE 300 respectivement. Pour le S&P500, d'après le premier critère le modèle est un VAR(4) avec des erreurs GARCH(1,2). Selon le critère de Schwartz, le meilleur modèle est un VAR(1) dont les erreurs suivent un GARCH(1,2). Les coefficients de ces deux modèles

sont inscrits dans les tableaux 15 et 16 respectivement. Les coefficients du VAR(4) sélectionnés par le critère d'Akaike ne sont pas tous significatifs. Seuls le retard 1 de la variable TSE et le retard 4 du S&P500 le sont. Dans le VAR(1) choisi par le critère de Schwartz le coefficient de la variable retardée endogène n'est pas significatif non plus. Nous remarquons également que le modèle préféré pour la variance est, avec les deux critères, un GARCH(1,2) ; c'est aussi le modèle que nous avons choisi quand nous avons estimé la variance de façon univarié (voir section 6).

Tableau 15 : VAR(4)-GARCH(1,2) pour le S&P500 :

$\hat{b}_{10}$	$\hat{b}_{11}$	$\hat{b}_{21}$	$\hat{b}_{12}$	$\hat{b}_{22}$	$\hat{b}_{13}$	$\hat{b}_{23}$
0.0530	0.0124	0.1579	-0.0171	-0.0045	-0.0156	0.0236
(0.0098)	(0.0171)	(0.0174)	(0.0155)	(0.0180)	(0.0151)	(0.0187)

$\hat{b}_{14}$	$\hat{b}_{24}$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
-0.0384	-0.0134	0.0065	0.1085	-0.0641	0.9488
(0.0159)	(0.0175)	(0.0009)	(0.0044)	(0.0062)	(0.0034)

Tableau 16 : VAR(1)-GARCH(1,2) pour le S&P500 :

$\hat{b}_{10}$	$\hat{b}_{11}$	$\hat{b}_{21}$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0495	0.0146	0.1505	0.0063	0.1099	-0.0664	0.9499
(0.0097)	(0.0171)	(0.0171)	(0.0009)	(0.0035)	(0.0053)	(0.0032)

Pour le TSE 300, le critère d'Akaike préfère un VAR(4) dont les erreurs suivent un GARCH(2,2) tandis que le critère de Schwartz sélectionne un VAR(1) dont les erreurs suivent un GARCH(1,2). Les coefficients sont inscrits dans les tableaux 17 et 18. On constate que comme dans le cas où la variable dépendante est le S&P500, le critère d'Akaike préfère un VAR(4) qui contient peu de coefficients significatifs. Les retards 1

du S&P500 et du TSE le sont, ainsi que le retard 4 du S&P500. Pour le VAR(1) sélectionné par le critère de Schwartz tous les coefficients sont significatifs.

Tableau 17 : VAR(4)-GARCH(2,2) pour le TSE 300.

$\hat{b}_{20}$	$\hat{b}^*_{11}$	$\hat{b}^*_{21}$	$\hat{b}^*_{12}$	$\hat{b}^*_{22}$	$\hat{b}^*_{13}$	$\hat{b}^*_{23}$
0.0396	0.0426	0.2255	-0.0070	-0.0035	0.0198	0.0238
(0.0075)	(0.0109)	(0.0167)	(0.0100)	(0.0161)	(0.0107)	(0.0159)

$\hat{b}^*_{14}$	$\hat{b}^*_{24}$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
-0.0372	0.0178	0.0008	0.1592	-0.1520	1.6769	-0.6856
(0.0115)	(0.0147)	(0.0001)	(0.0070)	(0.0065)	(0.0236)	(0.222)

Tableau 18 : VAR(1)-GARCH(1,2) pour le TSE 300.

$\hat{b}_{20}$	$\hat{b}^*_{11}$	$\hat{b}^*_{21}$	$\hat{\omega}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$
0.0388	0.0424	0.2211	0.0129	0.1638	-0.0630	0.8784
(0.0075)	(0.0106)	(0.0157)	(0.0011)	(0.0061)	(0.0073)	(0.0056)

Si nous comparons nos résultats avec ceux de Karolyi (1995), qui modélise les mêmes séries mais sur une période différente, nous nous aperçevons qu'il choisit un VAR(1) lorsqu'il suppose que la variance suit un GARCH univarié. Or c'est le modèle que sélectionne le critère de Schwartz. Ceci remet en question la préférence que nous avions accordé au critère d'Akaike dans la section 5.2.3. De façon identique au cas univarié, nous voulons faire une analyse de la moyenne sans spécification de la variance avant de se prononcer sur la modélisation choisie.

9) Analyse bivariée robuste des rendements vis à vis de l'hétéroscédasticité.

Nous estimons les modèles VAR séparés avec la méthode des moments généralisés sans faire aucune hypothèse sur la forme de la variance, de la même manière que dans le cas univarié. Nous utilisons, d'abord, les GMM juste-identifiés. Les résultats sont exposés dans l'annexe 10. Nous estimons, ensuite, les deux séries par les GMM sur-identifiés. Nous utilisons deux instruments de plus que de variables explicatives. Par exemple pour un VAR(1), $z_t = (1, y_{1t-1}, y_{2t-1}, y_{1t-2}, y_{2t-2})$. Les estimations sont présentées dans l'annexe 11. Finalement, nous améliorons l'estimation en corrigeant l'hétéroscédasticité. Pour cela nous divisons chaque instrument par la variance conditionnelle. Nous choisissons, parmi les instruments admis, un GARCH(1,1) pour spécifier h_t. Avec les résidus estimés de chacun des modèles régressés par les GMM (voir annexe 11), nous estimons des GARCH(1,1) par la méthode du quasi maximum de vraisemblance, que nous notons $\hat{h}_t$. Ainsi, nous utilisons la méthode des GMM en divisant chaque instrument par $\hat{h}_t$. Par exemple pour un VAR(1), nous avons $z_t = (\frac{1}{\hat{h}_t}, \frac{y_{1t-1}}{\hat{h}_t}, \frac{y_{2t-1}}{\hat{h}_t}, \frac{y_{1t-2}}{\hat{h}_t}, \frac{y_{2t-2}}{\hat{h}_t})$. Les résultats des estimations sont présentés dans les tableaux 19 et 20 : il semble que le VAR(1) soit le meilleur modèle. En effet le l test accepte le modèle, la p-value du Jtest est supérieure à 0.05 et les coefficients dont les retards sont supérieurs à 1 ne sont pas significatifs pour les deux séries. Nous voulons

nous assurer qu'ils sont nuls. Considérons p restrictions telles que l'hypothèse nulle soit $H_0 : R(\theta_0) = r$, avec r = 0 dans notre cas. Nous posons θ_T^n l'estimateur non contraint des

Tableau 19 : Estimation de la moyenne bivariée du S&P500 par les GMM1.

GMM1	BB	VAR(1)	VAR(2)	VAR(3)	VAR(4)
$\hat{b}_{10}$	0.536	0.0357	0.0373	0.0383	0.0401
	(0.0097)	(0.0098)	(0.0099)	(0.0099)	(0.0100)
$\hat{b}_{11}$		0.0355	0.0358	0.0319	0.0316
		(0.0171)	(0.0171)	(0.0171)	(0.0171)
$\hat{b}_{21}$		0.1499	0.1541	0.1542	0.1573
		(0.0202)	(0.0209)	(0.0209)	(0.0209)
$\hat{b}_{12}$			-0.0115	-0.0150	-0.0130
			(0.0161)	(0.0160)	(0.0161)
$\hat{b}_{22}$			-0.0085	-0.0061	-0.0076
			(0.0200)	(0.0204)	(0.0204)
$\hat{b}_{13}$				-0.0197	-0.0196
				(0.0158)	(0.0158)
$\hat{b}_{23}$				0.0128	0.0223
				(0.0191)	(0.0195)
$\hat{b}_{14}$					-0.0307
					(0.0161)
$\hat{b}_{24}$					-0.0135
					(0.0203)
J	52.50	1.48	1.37	8.34	2.31
p-value	3.98 e-12	0.4773	0.5043	0.0154	0.3139
Q10	47.98	48.08	42.72	39.68	36.40
Q100	175.45	174.46	169.41	166.09	164.76
l_{10}	14.24	11.83	13.07	12.39	6.80
l_{100}	84.56	93.93	105.85	94.64	91.69

Tableau 20 : Estimation de la moyenne bivariée du TSE 300 par les GMM1.

GMM1	BB	VAR(1)	VAR(2)	VAR(3)	VAR(4)
$\hat{b}_{20}$	0.0578	0.0281	0.0288	0.0275	0.0264
	(0.0074)	(0.0074)	(0.0075)	(0.0076)	(0.0075)
$\hat{b}^*_{11}$		0.0541	0.0488	0.0508	0.0536
		(0.0133)	(0.0132)	(0.0133)	(0.0129)
$\hat{b}^*_{21}$		0.2549	0.2534	0.2533	0.2513
		(0.0170)	(0.0174)	(0.0173)	(0.0174)
$\hat{b}^*_{12}$			-0.0014	-0.0025	0.0004
			(0.0119)	(0.0119)	(0.0119)
$\hat{b}^*_{22}$			0.0039	-0.0050	-0.0060
			(0.0169)	(0.0172)	(0.0172)
$\hat{b}^*_{13}$				0.0106	0.0140
				(0.0110)	(0.0119)
$\hat{b}^*_{23}$				0.0253	0.0243
				(0.0165)	(0.0167)
$\hat{b}^*_{14}$					-0.0282
					(0.0119)
$\hat{b}^*_{24}$					0.0279
					(0.0171)
J	172.11	0.13	7.82	5.96	3.54
p-value	4.23 e-38	0.9378	0.0200	0.0507	0.1696
Q10	291.50	92.24	86.94	75.10	78.17
Q100	431.81	215.29	209.88	201.15	205.08
l_{10}	50.55	18.37	17.55	16.50	16.88
l_{100}	151.70	113.94	113.12	112.20	113.76

GMM et θ_T^c l'estimateur contraint par H_0 des GMM. Nous utilisons deux tests. Tout d'abord, la statistique du test de Wald est calculée en utilisant seulement l'estimation non contrainte, tel que :

$$\text{WT} = T(R(\theta_T^n) - r)'(\Lambda_T (D_T S_T^{-1} D_T)^{-1} \Lambda_T')^{-1} (R(\theta_T^n) - r).$$

S_T, D_T, Λ_T sont estimés avec le modèle non contraint. S_T est la matrice de variance-covariance. Nous avons vu que $\hat{S}_T = (1/T)\sum_{t=1}^{T} (f_t(\hat{\theta}_T))(f_t(\hat{\theta}_T))'$.

$D = E\left(\frac{\partial f_t(\theta)}{\partial \theta'}\right)$, et D_T est un estimateur convergent de D.

Avec $f_t(\theta) = z_t \varepsilon_t$, où ε_t sont les résidus du modèle et z_t les variables instrumentales.

$\Lambda = \frac{\partial R}{\partial \theta'}$, et Λ_T est un estimateur convergent de Λ.

Nous employons, ensuite, la statistique du ratio de vraisemblance qui est T fois la différence entre la valeur minimum de la fonction objective quand les paramètres sont restreints et la valeur minimum de la fonction objective quand les paramètres ne sont pas contraints :

$$\text{LR} = T(Q_T(\theta_T^c) - Q_T(\theta_T^n)).$$

Il est nécessaire d'utiliser la même matrice S pour l'estimateur contraint et non contraint. Comme l'estimateur non contraint est convergent sous l'hypothèse nulle et alternative, un estimateur non contraint, convergent, de θ est employé pour trouver S. Les deux statistiques suivent asymptotiquement une $\chi^2(p)$ où p est le nombre de restrictions. Les résultats des statistiques sont présentés dans les tableaux 21 et 22.

Tableau 21 : Statistiques du test de Wald et du ratio de vraisemblance pour le S&P500.

S&P500	VAR(2)	VAR(3)	VAR(4)
WT	1.3733	3.2945	11.3994
LR	1.3720	3.2918	11.3899

Table 22 : Statistiques du test de Wald et du ratio de vraisemblance pour le TSE.

TSE300	VAR(2)	VAR(3)	VAR(4)
WT	0.0548	6.6299	12.6421
LR	1.3720	6.6238	12.6308

Dans tous les cas l'hypothèse H_0 est acceptée, nous concluons donc que les deux séries suivent un VAR(1) en moyenne. Les coefficients du VAR(1) étant tous significatifs, nous en déduisons que le S&P500 cause en moyenne le TSE et que le TSE cause en moyenne le S&P500. Nous rappellons que Karolyi (1995) sélectionne un VAR(1) lorsque la variance est un modèle GARCH univarié. En ce qui concerne la causalité, nous arrivons à la même conclusion que lui, c'est à dire qu'il y a causalité dans les deux sens entre les deux indices.

10) Conclusion

Dans cette étude, nous nous sommes intéressé à la modélisation et à l'estimation du taux de rendement du S&P500 et de celui du TSE 300 qui représentent les marchés boursiers des États Unis et du Canada respectivement. Ces deux indices ont été modélisés de manière standard, par un modèle autorégressif pour la moyenne et un modèle GARCH pour la variance. Nous avons tout d'abord estimé ces deux séries avec la méthode du quasi maximum de vraisemblance qui est une estimation jointe de la moyenne et de la variance spécifiée. Puis nous avons employé la méthode des moments généralisés pour une estimation de la moyenne sans détermination de la variance. Les estimations nous ont permis de faire une comparaison des deux méthodes. Il apparaît que la méthode des moments généralisés est aussi efficace que la méthode du maximum de vraisemblance, et qu'elle est robuste à la forme de l'hétéroscédasticité. En effet, elle permet une estimation de la moyenne qui ne nécessite aucune spécification de la variance. Ainsi, la moyenne ne peut être contaminée par une variance mal spécifiée. Nous utilisons donc cette méthode pour choisir le meilleur modèle pour chacun des deux indices. Un AR(4) semble être une bonne spécification pour la moyenne du S&P500 et un AR(3) pour le TSE 300. Nous pouvons également retenir de cette étude qu'il existe une causalité dans les deux sens entre l'indice américain et l'indice canadien.

Bibliographie

- 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,* pp177-181.

- Bollerslev, T.;(1986); « Generalized Autoregressive Conditional Heteroskedasticity. »; *Journal of Econometrics,* vol.31, pp307-327.

- Bollerslev, T.;(1987); « A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. »; *Review of Economics and Statistics*, vol.69, pp542-547.

- Bollerslev, T.;(1988); « On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process. »; *Journal of Time Series Analysis*, vol.9, pp121-131.

- Bollerslev, T.; Chou, R.Y.; Jayaraman, N.; Kroner, K.F.;(1991); « Les modèles ARCH en finance : un point sur la théorie et les résultats empiriques. »; *Annales d'Économie et de Statistiques*, n° 24.

- Campbell, J.Y.; Lo, A.W.; Mackinlay, A.C.;(1997); « The Econometrics of Financial Markets. »; *Princeton University Press.*

- Comte, F.; Lieberman, O.;(1997); « Variance Noncausality in Multivariate Garch Processes. »; *manuscript, Université Paris 6.*

- Cumby, R.E.;Huizinga, J.;(1992); « Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions »; *Econometrica,* vol. 60, No.1,pp185-195.

-Engle, R.F.;(1982); « Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. »; *Econometrica,* vol.50, pp987-1007.

- Engle, R.F.; Bollerslev, T.;(1986); « Modelling the Persistence of Conditional Variances. »; *Econometric Reviews,* vol.5(1), pp1-50.

- Engle, R.F.; Kroner, K.F.;(1995); « Multivariate Simultaneous Generalized ARCH »; *Econometric Theory,* vol. 11,pp122-150

- Engle, R.F.; Nelson, D.B.;(1994); « ARCH Models »; *Handbook of Econometrics,* vol. IV.

- Gallant R A;(1987); « Non Linear Statistical Models »; Wiley.

- Gouriéroux, C;(1992); « Modèles ARCH et Applications Financières »; *Economica.*

- Gouriéroux, C; Monfort, A;(1989); « Statistiques et Modèles Économétriques »; *Economica,* tome 2.

- Gouriéroux, C; Monfort, A; Trognon A;(1984); « Pseudo Maximum Likelihood Methods : Theory »; *Econometrica* 52, p681-700.

- Granger, C.W.J.;(1969); « Investigating Causal Relations by Econometric Models and Cross Spectral Methods. »; *Econometrica*, vol.37, pp424-438.

- Granger, C.W.J.;(1980); « The Variation of Certain Speculative Prices. »; *Journal of Economics Dynamics and control*, vol.36, pp394-419.

- Granger, C.W.J.; Robins, R.P.; Engle, R.F.;(1986); « Wholesale and Retail Prices : Bivariate Time-Series Modelling with Forcastable Error Variances. »; *In Besley, D. & Kuh, E. (Eds.), Model Reliability, MIT Press, Cambridge,*MA,1-17.

- Hamilton, J.D.;(1994); « Time Series Analysis. »; *princeton university Press.*

- Karolyi, G.A.;(1995); « A Multivariate GARCH Model of International Transmissions of Stock Returns and Volatility : The Case of the United States and Canada. »; *Journal of Business & Economic Statistics,* (January), vol.13,n° 1.

- Mandelbrot, B.;(1963); « The Variation of Certain Speculative Prices. »; *Journal of Business*, vol.36, pp394-419.

- Nelson, D.B.;(1989); « Modelling stock Market Volatility Changes. »; *in Proceedings from the American Statistical Assocaition, Business and Economic statistics Section .*

- Nelson, D.B.;(1991); « Conditional Heteroskedasticity in Asset Returns : A New Approch. »; *Econometrica*, vol.59, pp347-370.

- Nelson, D.B.; Cao, C.Q.;(1992); « Inequality Constraints in the Univariate GARCH Model »; *Journal of Business and Economic Statistics,* vol 10, n 2.

- Ogaki, M.;(1993); « Generalized Method of Moments : Econometric Applications »; *Hanbooks of Statistics,* vol. 11, pp455-488.

Annexe 1 : Valeurs de la fonction de vraisemblance, du critère d'akaike et de Schwartz pour le S&P500.

Valeurs de la fonction de vraisemblance pour le S&P500.

$y_t = \mu_t + \varepsilon_t$	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constante	-2985.28	-2971.70	-2969.49	-2966.63	-2963.71	-2961.98
Arch(1)	-2608.29	-2590.95	-2586.72	-2581.47	-2569.29	-2563.23
Arch(2)	-2494.64	-2467.66	-2467.51	-2465.53	-2450.85	-2448.53
Arch(3)	-2433.58	-2418.11	-2417.81	-2409.87	-2401.94	-2398.73
Arch(4)	-2363.17	-2333.97	-2333.88	-2330.53	-2324.03	-2322.65
Garch(11)	-2162.69	-2142.49	-2142.16	-2140.64	-2134.29	-2133.97
Garch(2,1)	-2156.64	-2137.42	-2137.41	-2135.99	-2130.57	-2129.91
Garch(1,2)	-2149.93	-2132.01	-2132.00	-2130.34	2125.68	-2125.05
Garch(2,2)	-2156.92	-2137.22	-2137.21	-2136.66	-2130.56	-2121.79

Valeurs du critère d'Akaike pour le S&P500.

AIC	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constante	-5974.56	-5949.40	-5946.98	-5943.26	-5939.42	-5937.96
Arch(1)	-5222.58	-5189.90	-5183.44	-5174.94	-5152.58	-5142.46
Arch(2)	-4997.28	-4945.32	-4847.62	-4945.06	-4917.70	-4915.06
Arch(3)	-4877.16	-4848.22	-4849.62	-4835.74	-4821.88	-4823.88
Arch(4)	-4738.34	-4681.94	-4683.76	-4679.06	-4668.06	-4667.30
Garch(1,1)	-4333.98	-4294.98	-4296.32	-4295.28	-4284.58	-4285.94
Garch(2,1)	-4323.28	-4286.84	-4288.82	-4287.98	-4279.14	-4279.82
Garch(1,2)	-4309.86	-4276.02	-4278.00	-4276.68	-4269.68	-4270.10
Garch(2,2)	-4325.84	-4288.44	-4290.42	-4291.32	-4281.12	-4265.58

Tableau 5 : Valeurs du critère de Schwartz pour le S&P500.

SIC	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constante	-5988.11	-5969.73	-5974.09	-5977.14	-5980.08	-5985.40
Arch(1)	-5242.91	-5217.01	-5102.58	-5215.60	-5200.02	-5196.67
Arch(2)	-5024.39	-4979.20	-4987.68	-4992.50	-4971.91	-4976.05
Arch(3)	-4898.06	-4888.88	-4897.06	-4889.95	-4882.87	-4885.23
Arch(4)	-4779.00	-4729.38	-4737.97	-4740.05	-4735.83	-4741.85
Garch(1,1)	-4360.49	-4328.86	-4336.98	-4342.72	-4338.79	-4346.93
Garch(2,1)	-4357.16	-4327.50	-4336.26	-4323.77	-4340.13	-4347.59
Garch(1,2)	-4343.74	-4316.68	-4325.44	-4330.89	-4328.99	-4337.87
Garch(2,2)	-4366.50	-4335.88	-4344.63	-4352.31	-4348.97	-4340.13

Annexe 2 : Valeurs de la fonction de vraisemblance, du critère d'akaike et de Schwartz pour le TSE 300.

Valeurs de la fonction de vraisemblance pour le TSE.

$y_t = \mu_t + \varepsilon_t$	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constante	-1609.94	-1478.92	-1476.99	-1472.27	-1470.63	-1465.21
Arch(1)	-1093.78	-956.58	-933.27	-900.82	-899.11	-897.18
Arch(2)	-889.99	-722.70	-720.19	-713.03	-711.02	-709.75
Arch(3)	-796.32	-623.89	-621.17	-618.57	-618.11	-618.01
Arch(4)	-693.25	-509.75	-508.88	-507.75	-507.62	-505.99
Garch(1,1)	-566.11	-390.57	-389.86	-386.67	-385.33	-383.11
Garch(2,1)	-556.83	-387.27	-387.14	-383.70	-382.48	-380.17
Garch(1,2)	-546.34	-384.14	-384.05	-380.03	-378.92	-376.56
Garch(2,2)	-537.35	-365.99	-365.98	-363.02	-383.50	-378.30

Valeurs du critère d'Akaike pour le TSE.

AIC	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constante	-3223.88	-2963.84	-2961.98	-2954.54	-2953.26	-2944.42
Arch(1)	-2193.56	-1921.16	-1876.54	-1813.64	-1812.22	-1810.36
Arch(2)	-1787.56	-1455.40	-1452.38	-1440.06	-1438.04	-1437.50
Arch(3)	-1602.64	-1259.78	-1256.34	-1253.14	-1254.22	-1256.02
Arch(4)	-1398.5	-1033.50	-1033.76	-1033.49	-1035.24	-1033.98
Garch(1,1)	-1140.22	-791.14	-791.72	-787.34	-786.66	-784.22
Garch(2,1)	-1123.66	-786.54	-788.28	-783.40	-782.96	-780.34
Garch(1,2)	-1102.68	-780.28	-782.10	-776.06	-785.00	-773.12
Garch(2,2)	-1086.70	-745.98	-747.96	-744.04	-787.00	-778.60

Valeurs du critère de Schwartz pour le TSE.

SIC	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constante	-3237.43	-2984.17	-2989.09	-2988.42	-2993.92	-2991.86
Arch(1)	-2240.22	-1948.27	-1910.42	-1854.30	-1859.66	-1864.57
Arch(2)	-1815.09	-1489.28	-1493.04	-1487.50	-1492.25	-1498.49
Arch(3)	-1636.52	-1300.44	-1303.78	-1307.35	-1315.21	-1323.79
Arch(4)	-1439.16	-1080.94	-1087.97	-1094.49	-1103.01	-1108.53
Garch(1,1)	-1167.33	-825.02	-832.38	-834.78	-840.87	-845.21
Garch(2,1)	-1157.54	-827.20	-835.72	-837.61	-843.95	-822.14
Garch(1,2)	-1136.56	-820.94	-829.94	-830.27	-845.99	-844.37
Garch(2,2)	-1127.36	-793.41	-802.17	-805.03	-854.77	-853.15

Annexe 3 : Estimation par le maximum de vraisemblance du S&P500 et du TSE300.

Estimation par le maximum de vraisemblance du S&P500.

MV	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0602	0.0549	0.0549	0.0557	0.0580	0.0586
	(0.0097)	(0.0097)	(0.0097)	(0.0097)	(0.0097)	(0.0097)
$\hat{b}_1$		0.0842	0.0841	0.0840	0.0832	0.0828
		(0.0137)	(0.0137)	(0.0137)	(0.0137)	(0.0137)
$\hat{b}_2$			0.0003	0.0011	0.0011	0.0009
			(0.0127)	(0.0127)	(0.0128)	(0.0128)
$\hat{b}_3$				-0.0114	-0.0085	-0.0086
				(0.0130)	(0.0131)	(0.0131)
$\hat{b}_4$					-0.0436	-0.0428
					(0.0132)	(0.0133)
$\hat{b}_5$						-0.0105
						(0.0127)
$\hat{\omega}$	0.0101	0.0100	0.0100	0.0100	0.0101	0.0102
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
$\hat{a}_1$	0.0632	0.0633	0.0634	0.0632	0.0636	0.0637
	(0.0013)	(0.0013)	(0.0014)	(0.0015)	(0.0016)	(0.0016)
$\hat{\beta}_1$	0.9272	0.9271	0.9270	0.9271	0.9265	0.9264
	(0.0028)	(0.0028)	(0.0029)	(0.0029)	(0.0030)	(0.0030)

Estimation par le maximum de vraisemblance du TSE 300.

MV	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0547	0.0411	0.0414	0.0399	0.0403	0.0401
	(0.0075)	(0.0073)	(0.0073)	(0.0073)	(0.0074)	(0.0074)
$\hat{b}_1$		0.2537	0.2549	0.2554	0.2555	0.2559
		(0.0130)	(0.0139)	(0.0140)	(0.0140)	(0.0141)
$\hat{b}_2$			-0.0046	-0.0132	-0.0131	-0.0136
			(0.0112)	(0.0126)	(0.0127)	(0.0128)
$\hat{b}_3$				0.0331	0.0359	0.0357
				(0.0130)	(0.0134)	(0.0134)
$\hat{b}_4$					-0.0112	-0.0137
					(0.0127)	(0.0134)
$\hat{b}_5$						0.0101
						(0.0127)
$\hat{\omega}$	0.0184	0.0167	0.0167	0.0165	0.0165	0.0165
	(0.0014)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)
$\hat{a}_1$	0.1308	0.1240	0.1240	0.1231	0.1231	0.1232
	(0.0034)	(0.0035)	(0.0035)	(0.0035)	(0.0036)	(0.0039)
$\hat{\beta}_1$	0.8417	0.8494	0.8495	0.8506	0.8506	0.8506
	(0.0055)	(0.0053)	(0.0054)	(0.0053)	(0.0054)	(0.0055)

Annexe 4 : Estimation de la moyenne du S&P500 et du TSE 300 par les MCO.

Estimation de la moyenne du S&P500 par les MCO.

MCO	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0545	0.0510	0.0524	0.0531	0.0547	0.0535
	(0.0119)	(0.0125)	(0.0130)	(0.0134)	(0.0133)	(0.0137)
$\hat{b}_1$		0.0641	0.0658	0.0655	0.0651	0.0657
		(0.0372)	(0.0368)	(0.0369)	(0.0366)	(0.0365)
$\hat{b}_2$			-0.0268	-0.0259	-0.0267	-0.0265
			(0.0351)	(0.0348)	(0.0345)	(0.0348)
$\hat{b}_3$				-0.0141	-0.0121	-0.0116
				(0.0303)	(0.0302)	(0.0297)
$\hat{b}_4$					-0.0302	-0.0316
					(0.0183)	(0.0190)
$\hat{b}_5$						0.0206
						(0.0352)
Q10	47.98	21.77	17.28	16.05	11.07	8.22
Q100	175.45	148.07	143.98	141.51	137.29	134.95
$l_{q,10}$	14.24	8.34	8.29	8.34	5.71	5.42
$l_{q,100}$	84.58	83.35	84.46	83.35	83.55	83.69

Estimation de la moyenne du TSE 300 par les MCO.

MCO	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0427	0.0342	0.0348	0.0334	0.0332	0.0321
	(0.0097)	(0.0098)	(0.0103)	(0.0103)	(0.0105)	(0.0107)
$\hat{b}_1$		0.1998	0.2033	0.2040	0.2037	0.2034
		(0.0340)	(0.0330)	(0.0324)	(0.0324)	(0.0324)
$\hat{b}_2$			-0.0178	-0.0256	-0.0254	-0.0265
			(0.0376)	(0.0377)	(0.0380)	(0.0376)
$\hat{b}_3$				0.0384	0.0366	0.0375
				(0.0315)	(0.0343)	(0.0334)
$\hat{b}_4$					0.0085	0.0022
					(0.0284)	(0.0265)
$\hat{b}_5$						0.0308
						(0.0284)
Q10	291.50	30.96	29.70	19.78	19.39	13.10
Q100	431.81	153.12	151.13	144.94	144.45	136.36
$l_{q,10}$	50.59	11.13	10.38	6.39	7.04	8.50
$l_{q,100}$	151.50	102.58	101.82	100.66	100.58	101.52

Annexe 5 : Estimation de la moyenne du S&P500 et du TSE 300 par les GMM.

Estimation de la moyenne du S&P500 par les GMM.

GMM	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0570	0.0509	0.0536	0.0520	0.0510	0.0468
	(0.0113)	(0.0119)	(0.0124)	(0.0132)	(0.0130)	(0.0130)
$\hat{b}_1$		0.0585	0.0593	0.0648	0.0822	0.0.831
		(0.0358)	(0.0354)	(0.0363)	(0.0359)	(0.0357)
$\hat{b}_2$			-0.0234	-0.0212	-0.0264	-0.0173
			(0.0333)	(0.0336)	(0.0330)	(0.0347)
$\hat{b}_3$				-0.0048	-0.0041	-0.0068
				(0.0284)	(0.0290)	(0.0294)
$\hat{b}_4$					-0.0318	-0.0392
					(0.0167)	(0.0178)
$\hat{b}_5$						0.0349
						(0.0323)
J	4.79	3.14	2.77	5.64	2.65	2.51
p-value	0.4422	0.3702	0.4288	0.1302	0.4477	0.4728
Q10	47.98	22.03	17.62	16.76	14.17	12.34
Q100	175.45	148.26	144.26	142.92	138.83	137.52
$l_{q,10}$	14.24	8.36	8.32	8.33	5.83	5.54
$l_{q,100}$	84.58	83.32	84.42	84.81	82.24	82.48

Estimation de la moyenne du TSE 300 par les GMM.

GMM	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0515	0.0338	0.0394	0.0334	0.0332	0.0310
	(0.0093)	(0.0094)	(0.0099)	(0.0100)	(0.0102)	(0.0104)
$\hat{b}_1$		0.2178	0.2185	0.2030	0.2045	0.2048
		(0.0318)	(0.0320)	(0.0322)	(0.0322)	(0.0324)
$\hat{b}_2$			-0.0433	-0.0340	-0.0349	-0.0255
			(0.0347)	(0.0346)	(0.0359)	(0.0374)
$\hat{b}_3$				0.0530	0.0532	0.0378
				(0.0265)	(0.0303)	(0.0334)
$\hat{b}_4$					0.0017	0.0063
					(0.0256)	(0.0257)
$\hat{b}_5$						0.0287
						(0.0282)
J	41.06	3.01	4.58	1.96	2.65	1.91
p-value	9.12 e-08	0.3894	0.2051	0.5803	0.4480	0.5893
Q10	291.50	33.68	35.80	21.09	20.98	13.21
Q100	431.81	155.14	156.06	147.71	148.51	137.54
$l_{q,10}$	50.59	11.03	10.23	6.27	6.99	8.57
$l_{q,100}$	151.50	102.43	101.75	100.65	101.53	102.57

Annexe 6 : Estimation de la moyenne du S&P500 et du TSE 300 par les GMM1.

Estimation de la moyenne du S&P500 par les GMM1.

GMM1	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0530	0.0412	0.0409	0.0409	0.0418	0.0430
	(0.0097)	(0.0097)	(0.0099)	(0.0100)	(0.0100)	(0.0100)
$\hat{b}_1$		0.1035	0.1041	0.1033	0.1066	0.1059
		(0.0142)	(0.0142)	(0.0143)	(0.0143)	(0.0143)
$\hat{b}_2$			0.0006	0.0036	0.0050	0.0046
			(0.0136)	(0.0136)	(0.0135)	(0.0135)
$\hat{b}_3$				-0.0102	-0.0068	-0.0068
				(0.0132)	(0.0133)	(0.0133)
$\hat{b}_4$					-0.0296	-0.0288
					(0.0132)	(0.0133)
$\hat{b}_5$						-0.0129
						(0.0134)
J	55.04	6.20	6.98	12.32	7.14	6.34
p-value	1.27 e-10	0.1023	0.0725	0.0064	0.0674	0.0963
Q10	47.98	30.88	31.31	31.22	29.65	34.76
Q100	175.45	157.49	157.93	156.89	153.87	158.94
$l_{q,10}$	14.23	8.08	7.75	7.59	5.63	5.31
$l_{q,100}$	84.54	83.27	84.00	84.19	82.05	81.78

Estimation de la moyenne du TSE 300 par les GMM1.

GMM1	BB	AR(1)	AR(2)	AR(3)	AR(4)	AR(5)
$\hat{b}_0$	0.0540	0.0294	0.0290	0.0265	0.0263	0.0252
	(0.0075)	(0.0074)	(0.0075)	(0.0075)	(0.0075)	(0.0075)
$\hat{b}_1$		0.2939	0.2929	0.2966	0.2961	0.2966
		(0.0145)	(0.0147)	(0.0148)	(0.0148)	(0.0149)
$\hat{b}_2$			0.0021	-0.0046	-0.0048	-0.0051
			(0.0144)	(0.0147)	(0.0147)	(0.0147)
$\hat{b}_3$				0.0342	0.0332	0.0335
				(0.0138)	(0.0141)	(0.0142)
$\hat{b}_4$					0.0041	-0.0004
					(0.0138)	(0.0141)
$\hat{b}_5$						0.0227
						(0.0134)
J	385.39	6.31	9.60	4.14	5.02	2.16
p-value	4.16 e-81	0.0975	0.0222	0.2464	0.4134	0.5401
Q10	291.50	89.83	89.56	82.48	82.01	76.97
Q100	431.81	209.99	209.82	207.38	206.74	200.00
$l_{q,10}$	50.53	11.88	8.79	6.80	7.19	8.54
$l_{q,100}$	151.57	101.05	98.00	98.29	98.32	100.44

Annexe 7 : Le l test.

Supposons que $y_t = X_t\delta + \varepsilon_t$. X_t est un vecteur (1×k) de variables explicatives.

δ est un vecteur (k×1) de paramètres inconnus.

Z_t est un vecteur (h×1) de variables instrumentales.

Soit $d = (X'ZA_T^{-1}Z'X)^{-1}X'ZA_T^{-1}Z'y$, où A_T est proportionnelle à une estimation convergente de la matrice S vue précédemment dans le développement sur les GMM (section 5.1).

La matrice de covariance asymptotique de l'estimateur d est notée V_d :

$V_d = DSD'$, avec $D = p\lim T(X'ZA_T^{-1}Z'X)^{-1}X'ZA_T^{-1}$.

Le théorème central-limite permet d'écrire que :

$\sqrt{T}\hat{r} = \sqrt{T}r + \frac{\partial r}{\partial \delta}\sqrt{T}(d-\delta)$, où r est un vecteur s×1 des autocorrélations supérieur à q,

$r = [r_{q+1}, r_{q+2}, ..., r_{q+s}]'$.

D'après la proposition 1 de l'article de Cumby et Huizinga (1992) :

$\sqrt{T}\hat{r} \xrightarrow{A} N(0, V_{\hat{r}})$, où $V_{\hat{r}} = V_r + BV_dB' + CD'B' + BDC'$.

La matrice B d'éléments i et j est :
$B(i,j) = -[E(\varepsilon_{t-q-i}X_{j,t}) + E(\varepsilon_t X_{t-q-i})]/E(\varepsilon_t^2)$.

Si on pose $\xi_{i,t} = \varepsilon_t\varepsilon_{t-q-i}$, $\omega_{j,t} = \varepsilon_t Z_{j,t}$, i=1...s et j=1...h, le ij ^{ieme} élément de la matrice V_r est donné par :

$$V_r(i,j) = \sigma_\varepsilon^{-4}\sum_{n=-q}^{q} E(\xi_{i,t}\xi_{j,t-n}),$$

et le ij ^{ieme} élément de la matrice C est donné par :

$$C(i,j) = \sigma_\varepsilon^{-2}\sum_{n=-q}^{q} E(\xi_{i,t}\omega_{j,t-n}).$$

Annexe 8 : Valeurs de la fonction de vraisemblance, du critère d'Akaike et de Schwartz pour le S&P500 estimé de manière bivariée.

Valeurs de la fonction de vraisemblance pour le S&P500

$y_{1t} = \mu_{1t} + \varepsilon_{1t}$	(i)	(ii)	(iii)	(iv)
Constante	-2959.41	-2953.24	-2950.01	-2942.73
Arch(1)	-2555.89	-2545.03	-2535.50	-2543.39
Arch(2)	-2427.93	-2425.80	-2424.08	-2409.34
Arch(3)	-2381.08	-2379.49	-2371.59	-2362.12
Arch(4)	-2308.58	-2303.03	-2299.08	-2291.54
Garch(1,1)	-2113.69	-2112.92	-2111.25	-2103.61
Garch(2,1)	-2107.16	-2106.24	-2105.72	-2098.60
Garch(1,2)	-2101.41	-2100.46	-2098.63	-2092.37
Garch(2,2)	-2107.40	-2106.40	-2105.75	-2089.39

Valeurs du critère d'Akaike pour le S&P500.

AIC	(i)	(ii)	(iii)	(iv)
Constante	-5926.82	-5918.48	-5916.02	-5905.46
Arch(1)	-5121.78	-5104.06	-5089.00	-5108.78
Arch(2)	-4867.86	-4867.60	-4868.16	-4842.68
Arch(3)	-4776.16	-4776.98	-4765.18	-4750.24
Arch(4)	-4633.06	-4626.06	-4622.16	-4611.08
Garch(1,1)	-4239.38	-4241.84	-4242.50	-4231.22
Garch(2,1)	-4228.32	-4230.48	-4233.44	-4223.20
Garch(1,2)	-4216.82	-4218.92	-4220.26	-4210.74
Garch(2,2)	-4230.80	-4232.80	-4235.50	-4206.78

Valeurs du critère de Schwartz pour le S&P500.

SIC	(i)	(ii)	(iii)	(iv)
Constante	-5959.93	-5959.14	-5970.23	-5973.23
Arch(1)	-5155.66	-5151.50	-5149.99	-5183.33
Arch(2)	-4908.52	-4921.81	-4944.71	-4924.00
Arch(3)	-4823.44	-4837.97	-4839.73	-4838.34
Arch(4)	-4687.37	-4693.83	-4703.48	-4705.96
Garch(1,1)	-4280.04	-4296.05	-4310.17	-4312.54
Garch(2,1)	-4275.76	-4291.47	-4307.99	-4311.30
Garch(1,2)	-4264.26	-4279.91	-4293.81	-4298.84
Garch(2,2)	-4285.01	-4300.57	-4316.82	-4301.66

Annexe 9 : Valeurs de la fonction de vraisemblance, du critère d'Akaike et de Schwartz pour le TSE 300 estimé de manière bivariée.

Valeurs de la fonction de vraisemblance pour le TSE 300.

$y_{2t} = \mu_{2t} + \varepsilon_{2t}$	(i)	(ii)	(iii)	(iv)
Constante	-1447.89	-1445.13	-1439.89	-1427.73
Arch(1)	-950.21	-914.49	-883.71	-893.80
Arch(2)	-718.60	-713.81	-706.24	-701.56
Arch(3)	-617.46	-612.20	-608.40	-606.28
Arch(4)	-503.26	-502.39	-500.23	-496.21
Garch(1,1)	-382.58	-382.47	-377.39	-370.70
Garch(2,1)	-379.89	-379.84	-378.40	-367.62
Garch(1,2)	-377.21	-376.89	-370.46	-363.95
Garch(2,2)	-377.61	-377.54	-372.56	-346.99

Valeur du critère d'Akaike pour le TSE 300.

AIC	(i)	(ii)	(iii)	(iv)
Constante	-2903.78	-2902.26	-2895.78	-2875.46
Arch(1)	-1910.42	-1842.98	-1785.42	-1809.60
Arch(2)	-1449.20	-1443.62	-1432.48	-1427.12
Arch(3)	-1248.92	-1242.40	-1238.80	-1238.56
Arch(4)	-1022.52	-1024.78	-1024.46	-1020.42
Garch(1,1)	-777.16	-780.94	-774.78	-765.40
Garch(2,1)	-773.78	-777.68	-778.80	761.24
Garch(1,2)	-768.42	-771.78	-762.92	753.90
Garch(2,2)	-771.22	-775.08	-769.12	721.98

Valeur du critère de Schwartz pour le TSE 300.

SIC	(i)	(ii)	(iii)	(iv)
Constante	-2930.89	-2942.92	-2949.99	-2943.23
Arch(1)	-1944.30	-1890.42	-1846.41	-1884.15
Arch(2)	-1489.86	-1497.83	-1500.25	-1508.44
Arch(3)	-1296.36	-1303.39	-1313.35	-1326.66
Arch(4)	-1076.73	-1092.55	-1105.78	-1115.30
Garch(1,1)	-817.82	-835.15	-842.55	-846.72
Garch(2,1)	-821.22	-838.67	-853.35	-849.34
Garch(1,2)	-815.86	-832.77	-837.47	-842.00
Garch(2,2)	-825.43	-842.85	-850.44	-816.86

Annexe 10: Estimation de la moyenne bivariée du S&P500 et du TSE 300 par les MCO.

Estimation de la moyenne bivariée du S&P500 par les MCO.

MCO	BB	VAR(1)	VAR(2)	VAR(3)	VAR(4)
$\hat{b}_{10}$	0.0545	0.0495	0.0510	0.0520	0.0531
	(0.0119)	(0.0126)	(0.0131)	(0.0134)	(0.0134)
$\hat{b}_{11}$		0.0157	0.0117	0.0111	0.0099
		(0.0469)	(0.0445)	(0.0446)	(0.0440)
$\hat{b}_{21}$		0.0957	0.1064	0.1061	0.1077
		(0.0461)	(0.0378)	(0.0378)	(0.0377)
$\hat{b}_{12}$			-0.0560	-0.0579	-0.0601
			(0.0383)	(0.0386)	(0.0384)
$\hat{b}_{22}$			0.0308	0.0379	0.0380
			(0.0333)	(0.0344)	(0.0343)
$\hat{b}_{13}$				-0.0148	-0.0182
				(0.0309)	(0.0312)
$\hat{b}_{23}$				-0.0096	-0.0024
				(0.0312)	(0.0324)
$\hat{b}_{14}$					-0.0600
					(0.0312)
$\hat{b}_{24}$					0.0516
					(0.0438)
Q10	47.98	29.80	21.62	18.54	12.20
Q100	175.45	155.73	148.36	142.94	138.55
l_{10}	14.24	11.84	13.02	14.02	9.76
l_{100}	84.58	90.11	89.47	89.78	88.57

Estimation de la moyenne bivariée du TSE 300 par les MCO.

MCO	BB	VAR(1)	VAR(2)	VAR(3)	VAR(4)
$\hat{b}_{20}$	0.0427	0.0321	0.0331	0.0315	0.0321
	(0.0096)	(0.0097)	(0.0102)	(0.0104)	(0.0105)
$\hat{b}^*_{11}$		0.0981	0.0958	0.0964	0.0957
		(0.0356)	(0.0331)	(0.0329)	(0.0328)
$\hat{b}^*_{21}$		0.1238	0.1309	0.1313	0.1322
		(0.0463)	(0.0391)	(0.0389)	(0.0388)
$\hat{b}^*_{12}$			-0.0277	-0.0250	-0.0269
			(0.0391)	(0.0370)	(0.0369)
$\hat{b}^*_{22}$			0.0077	-0.0035	-0.0033
			(0.0360)	(0.0329)	(0.0328)
$\hat{b}^*_{13}$				0.0158	0.0130
				(0.0269)	(0.0276)
$\hat{b}^*_{23}$				0.0252	0.0301
				(0.0272)	(0.0281)
$\hat{b}^*_{14}$					-0.0555
					(0.0265)
$\hat{b}^*_{24}$					0.0553
					(0.0404)
Q10	291.50	32.20	30.97	19.23	17.47
Q100	431.81	158.98	157.70	149.27	148.83
l_{10}	50.59	11.95	11.03	11.36	8.88
l_{100}	151.50	104.35	104.84	104.14	103.82

Annexe 11 : Estimation de la moyenne bivariée du S&P500 et du TSE 300 par les GMM.

Estimation de la moyenne bivariée du S&P500 par les GMM.

GMM	BB	VAR(1)	VAR(2)	VAR(3)	VAR(4)
$\hat{b}_{10}$	0.0578	0.0463	0.0497	0.0545	0.0529
	(0.0113)	(0.0121)	(0.0127)	(0.0133)	(0.0133)
$\hat{b}_{11}$		0.0040	0.0166	-0.0044	0.0127
		(0.0451)	(0.0439)	(0.0435)	(0.0439)
$\hat{b}_{21}$		0.1352	0.1052	0.1130	0.1112
		(0.0338)	(0.0378)	(0.0378)	(0.0373)
$\hat{b}_{12}$			-0.0558	-0.0564	-0.0683
			(0.0376)	(0.0387)	(0.0375)
$\hat{b}_{22}$			0.0329	0.0317	0.0362
			(0.0324)	(0.0322)	(0.0322)
$\hat{b}_{13}$				-0.0104	-0.0192
				(0.0303)	(0.0309)
$\hat{b}_{23}$				0.0050	0.0080
				(0.0269)	(0.0293)
$\hat{b}_{14}$					-0.0512
					(0.0256)
$\hat{b}_{24}$					0.0301
					(0.0329)
J	4.17	1.93	0.44	3.92	3.38
p-value	0.1237	0.3816	0.8032	0.1407	0.1841
Q10	47.98	34.06	21.76	21.71	13.20
Q100	175.45	159.95	148.48	148.14	140.62
l_{10}	14.24	11.85	12.50	14.45	7.42
l_{100}	84.5	90.84	88.98	89.86	86.97

Estimation de la moyenne bivariée du TSE 300 par les GMM.

GMM	BB	VAR(1)	VAR(2)	VAR(3)	VAR(4)
$\hat{b}_{20}$	0.0489	0.0308	0.0331	0.0325	0.0340
	(0.0093)	(0.0094)	(0.0100)	(0.0104)	(0.0103)
$\hat{b}^*_{11}$		0.0851	0.0952	0.0828	0.1000
		(0.0292)	(0.0329)	(0.0326)	(0.0319)
$\hat{b}^*_{21}$		0.1471	0.1382	0.1380	0.129
		(0.0333)	(0.0390)	(0.0390)	(0.0380)
$\hat{b}^*_{12}$			-0.0192	-0.0186	-0.0309
			(0.0376)	(0.0372)	(0.0357)
$\hat{b}^*_{22}$			-0.0004	-0.0132	-0.0186
			(0.0313)	(0.0299)	(0.0300)
$\hat{b}^*_{13}$				0.0319	0.0169
				(0.0245)	(0.0268)
$\hat{b}^*_{23}$				0.0434	0.0493
				(0.0241)	(0.0234)
$\hat{b}^*_{14}$					-0.0334
					(0.0199)
$\hat{b}^*_{24}$					0.0193
					(0.0283)
J	24.13	0.50	1.40	4.20	1.23
p-value	5.75 e-06	0.7804	0.4956	0.1223	0.5185
Q10	291.50	33.15	31.58	22.89	22.25
Q100	431.81	158.87	157.72	155.52	155.56
l_{10}	50.97	11.80	16.85	9.83	7.22
l_{100}	152.18	104.32	108.89	103.17	101.73