JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

PASCAL LOUVET
OLLIVIER TARAMASCO

Estimation non paramétrique de la dépendance sérielle sur les rentabilités boursières d'indices internationaux

Journal de la société statistique de Paris, tome 138, n° 1 (1997), p. 53-79
<http://www.numdam.org/item?id=JSFS_1997__138_1_53_0>

ESTIMATION NON PARAMÉTRIQUE DE LA DÉPENDANCE SÉRIELLE SUR LES RENTABILITÉS BOURSIÈRES D'INDICES INTERNATIONAUX

Pascal LOUVET
CERAG - Grenoble

Ollivier TARAMASCO
Ecole Supérieure des Affaires - CERAG - Grenoble

RÉSUMÉ

Les cours boursiers ne sont pas à accroissements indépendants, comme le voudrait l'efficience informationnelle, mais semblent conserver une mémoire de leur trajectoire passée. Quelles sont l'importance et la forme de cette dépendance ? L'analyse canonique non linéaire, qui permet de répondre à cette double question, est appliquée sur les rentabilités journalières de neuf indices boursiers internationaux. Les résultats montrent que la dépendance s'explique par un double effet : le marché reproduit la tendance et la variance qu'il a constatées la veille. La mémoire du marché est cependant courte ; la rentabilité la plus récente fournit l'essentiel de l'information nécessaire pour déterminer la distribution conditionnelle, une rentabilité supplémentaire permet toutefois d'affiner l'estimation de la variance conditionnelle. Ces résultats suggèrent l'utilisation d'un modèle AR(1)-ARCH(2) pour rendre compte de la dépendance temporelle. Celle-ci tend à disparaître sur la période récente.

Mots clés : Rentabilité, indice boursier, dépendance, processus ARCH, analyse canonique, non-linéarité, non paramétrique.

ABSTRACT

In this article, we use a new technique to estimate dependence between consecutive stock markets returns : *the non linear canonical analysis*. It is applied on nine international indexes returns. Results show that dependence can be explained by two effects : one affects the trend and the other the variance of returns. These results lead to a non parametrical AR-ARCH representation of the returns process.

Keywords : Return, stock index, dependence, ARCH process, canonical analysis, non-linearity, non parametric.

I. Introduction

L'objectif de cet article est de décrire la dépendance entre les rentabilités boursières successives à l'aide d'une technique non paramétrique : l'analyse canonique non linéaire. Cette méthode est une généralisation à des variables continues de l'analyse factorielle des correspondances (A.F.C.) utilisée par LOUVET et TARAMASCO (1993). Elle permet non seulement d'identifier les caractéristiques du lien entre les rentabilités successives, mais aussi de déterminer les moments conditionnels de manière non paramétrique. Depuis ENGLE (1982) et BOLLERSLEV (1986), la plupart des modèles utilisés pour décrire le comportement des cours boursiers, appartiennent à la classe des modèles ARCH. Ils apparaissent dans tous les domaines de la finance de marché, que ce soit la gestion de portefeuille, les études d'événements, l'évaluation d'options, comme en témoigne la revue de la littérature sur ces modèles effectuée par BOLLERSLEV *et alii* (1992).

Lorsque ENGLE (1982) propose de modéliser les taux d'inflation par un modèle ARCH(p), il explique deux phénomènes repérés depuis longtemps sur les données économiques et financières : les distributions empiriques sont leptokurtiques, et les taux (de rentabilité, d'inflation,...) d'amplitude élevée se succèdent fréquemment. En effet, les modèles ARCH postulent que la variance conditionnelle est une fonction croissante de l'amplitude des variations passées. Lorsque cette amplitude est élevée, la variance augmente et la probabilité d'observer à nouveau une forte variation augmente elle aussi. Pour résultat, les queues de la distribution marginale sont alourdies[1]. Tous les modèles de type ARCH sont construits sur le même principe : formuler la relation entre les moments conditionnels et le passé du processus. Seule change la forme des fonctions caractérisant ces relations. Par exemple, le modèle GARCH(p, q) de BOLLERSLEV (1986) permet de limiter le nombre de retards dans l'expression de la variance conditionnelle, le modèle TARCH(p) de ZAKOIAN (1990) ou le modèle EGARCH de NELSON (1990) admettent une asymétrie dans les distributions conditionnelles,... Empiriquement, la supériorité des modèles ARCH par rapport à tout autre type de modèles ne semble faire aucun doute : pour HSIEH (1991), le rejet d'une distribution identique des rentabilités successives par le test B.D.S. est due à l'hétéroscédasticité conditionnelle ; LONGIN (1993) observe que la distribution empirique des cours annuels extrêmes (plus haut et plus bas) est compatible avec une distribution ARCH des rentabilités journalières ; pour DING *et alii* (1993), les très forts coefficients d'autocorrélation entre les puissances (non entières) des rentabilités s'expliquent grâce au modèle A-PARCH ; enfin, LOUVET et TARAMASCO (1993) montrent par une approche descriptive que les rentabilités sont générées par ce type de modèle.

1. Il est facile de montrer, comme le fait GOURIÉROUX (1993), que tous les modèles de type ARCH proposent la même propriété de générer des distributions à queues épaisses.

L'utilisation des modèles ARCH n'est cependant pas sans inconvénient. En particulier, dans les études empiriques, il n'y a aucun moyen simple de choisir entre les différents types de modèles (qui se comptent par dizaines) ; une fois choisi le modèle, le nombre de retards à retenir est très difficile à déterminer (contrairement aux modèles ARIMA, pour lesquels les fonctions d'autocorrélation, d'autocorrélation partielle et d'autocorrélation inverse permettent ce choix) ; enfin, l'estimation des paramètres par le maximum de vraisemblance pose de graves problèmes numériques[2]. Deux de ces difficultés sont levées par la technique utilisée dans cet article : le choix du modèle, et les problèmes numériques liés au maximum de vraisemblance.

L'analyse canonique est décrite dans la partie suivante. La troisième partie reprend les résultats de son application aux rentabilités de divers indices boursiers internationaux.

II. Méthodologie

L'analyse canonique non linéaire (par la suite, A.C.), dont l'origine remonte à BENZÉCRI (1973), permet de mesurer et de caractériser la dépendance entre deux vecteurs aléatoires. Elle s'appuie sur une définition précise de la notion de dépendance. Deux vecteurs aléatoires X et Y sont dépendants s'il existe une transformation déterministe[3] de X et une transformation déterministe de Y qui soient de corrélation (linéaire) non nulle. Dans l'analyse canonique, la dépendance est mesurée par le coefficient de corrélation maximal obtenu sur toutes les transformations possibles de X et de Y. Ce coefficient est appelé *coefficient de corrélation canonique.*

Un exemple permet de comprendre la puissance de l'A.C. Soit X une variable aléatoire de loi $N(0, 1)$. Construisons la variable $Y = X^2$. La dépendance entre les deux variables X et Y peut être qualifiée de "totale" puisque X détermine entièrement Y. Or la corrélation linéaire entre ces deux variables est nulle. Mesurer la dépendance par le seul calcul du coefficient de corrélation linéaire, apparaît clairement insuffisant. En revanche, lorsque la variable X est transformée par la fonction carrée et Y par la fonction identité, le coefficient de corrélation linéaire entre les deux transformées est égal à 1, démontrant ainsi la dépendance "totale" entre X et Y.

L'A.C. permet, elle, de rendre compte de la dépendance réelle en fournissant la transformation $\phi(X) = X^2$ et $\theta(Y) = Y$, et en donnant un coefficient de corrélation canonique égal à 1.

---

2. Les algorithmes classiques utilisés pour maximiser la vraisemblance convergent à vitesse très lente, lorsqu'ils convergent !

3. Si X est un vecteur de dimension p, une transformation déterministe du vecteur X consiste à lui appliquer une fonction déterministe de $\mathbb{R}^p \longrightarrow \mathbb{R}$.

Une fois trouvée la transformation de corrélation maximale, il peut être utile de rechercher d'autres transformations susceptibles de compléter la description de la dépendance. Rien n'oblige en effet à ce que la dépendance se réduise à une seule dimension. Parmi ces autres transformations, on sera tenté d'éliminer toutes celles qui sont en partie redondantes avec la transformation initiale et qui n'apporteraient que peu d'informations complémentaires. Aussi peut-on continuer l'analyse en recherchant une deuxième transformation de X et de Y, orthogonale (c'est-à-dire indépendante de la première) et de corrélation maximale, puis une troisième orthogonale aux deux premières, et ainsi de suite.

Pour aller plus loin dans la technique, il est nécessaire de décrire les propriétés probabilistes de l'A.C.

2.1 La définition de l'analyse canonique non linéaire

Soient X et Y deux vecteurs aléatoires de dimensions respectives p et q. L'analyse canonique de ces deux vecteurs consiste à résoudre dans un premier temps le problème de maximisation suivant :

$$\rho_1^* = \max_{\{\phi \in F(X), \theta \in G(Y)\}} \left[\rho\big(\phi(X), \theta(Y) \big) \right] \tag{1}$$

où

- $\rho(.,.)$ est le coefficient de corrélation (linéaire) ;
- $F(X)$ est l'ensemble des fonctions ϕ de $\mathbb{R}^p$ vers $\mathbb{R}$ qui vérifient :

$$E\big(\phi(X)\big) = 0 \text{ et } E\Big(\phi(X)^2\Big) < \infty$$

- $G(Y)$ est l'ensemble des fonctions θ de $\mathbb{R}^q \mapsto \mathbb{R}$ qui vérifient [4] :

$$E\big(\theta(Y)\big) = 0 \text{ et} E\Big(\theta(Y)^2\Big) < \infty$$

Une fois déterminé le premier coefficient de corrélation canonique ρ_1^* ainsi que le premier couple de fonctions canoniques (ϕ_1^*, θ_1^*) conduisant à la solution du problème (1), les autres coefficients et couples de fonctions canoniques sont déterminés par récurrence :

$$\rho_{i+1}^* = \max_{\{\phi \in F_i(X), \theta \in G_i(Y)\}} \left[\rho\big(\phi(X), \theta(Y) \big) \right] \tag{i+1}$$

4. Supposer $E\big(\phi(X)\big) = E\big(\theta(Y)\big) = 0$ permet d'éliminer la solution triviale $\phi(X) = constante, \theta(Y) = constante$, qui donne un coefficient de corrélation canonique égal à 1. La deuxième condition : $E\big(\phi(X)^2\big) < \infty$ et $E\big(\theta(Y)^2\big) < \infty$ est une condition d'existence du coefficient de corrélation.

où

- $F_i(X)$ est l'ensemble des fonctions ϕ de $F(X)$, telles que la variable aléatoire $\phi(X)$ et les variables aléatoires $\phi_1^*(X), \phi_2^*(X) \dots \phi_i^*(X)$ soient deux à deux non corrélées ;
- $G_i(Y)$ est l'ensemble des fonctions θ de $G(Y)$, telles que la variable aléatoire $\theta(Y)$ et les variables aléatoires $\theta_1^*(Y), \theta_2^*(Y), \dots, \theta_i^*(Y)$ soient deux à deux non corrélées.

En d'autres termes, l'A.C. recherche les transformées de X et de Y les mieux corrélées, puis itère sous contrainte d'indépendance (d'orthogonalité). Les coefficients de corrélation canonique sont donc naturellement classés dans l'ordre décroissant.

Remarques :

1. Les coefficients de corrélation canonique sont compris entre 0 et 1 ;
2. Lorsque X et Y sont des vecteurs continus, le nombre de triplets canoniques $(\rho_i^*, \phi_i^*, \theta_i^*)$ est *a priori* infini ;
3. Lorsque X et Y sont deux variables aléatoires discrètes, l'A.C. de (X, Y) est équivalente à l'analyse factorielle des correspondances de X et de Y (*cf.* BENZÉCRI (1973)).

2.1.1 Les propriétés des triplets canoniques

Le théorème suivant[5] permet de calculer l'ensemble des triplets canoniques sans qu'il soit nécessaire de résoudre séquentiellement les problèmes de maximisation sous contraintes décrits dans le paragraphe précédent.

Théorème 1. *Soit $(\rho_i^*, \phi_i^*, \theta_i^*)$ l'analyse canonique du couple (X, Y). Lorsque θ_i^* et ϕ_i^* ont été choisies telles que $Var(\theta_i^*(Y)) = Var(\phi_i^*(X)) = 1$[6], il vient :*

$$E[\phi_i^*(X) \,|\, Y] = \rho_i^* \theta_i^*(Y)$$
$$E[\theta_i^*(Y) \,|\, X] = \rho_i^* \phi_i^*(X)$$

Cette relation de réciprocité est appelée Alternative Conditional Expectation *ou* ACE *par Breiman et Freidman (1985). En combinant les deux équations, on obtient :*

$$E\big[E[\phi_i^*(X) \,|\, Y] \,|\, X\big] = {\rho_i^*}^2 \phi_i^*(X)$$

5. *cf.* TARAMASCO (1993) pour la démonstration.
6. θ_i^* et ϕ_i^* peuvent être choisies à une constante multiplicative près. En particulier, les deux fonctions $\theta_i^{**} = \dfrac{\theta_i^*}{\sigma(\theta_i^*(Y))}$ et $\phi_i^{**} = \dfrac{\phi_i^*}{\sigma(\phi_i^*(X))}$, sont aussi solutions du problème (i) et telles que $Var(\theta_i^{**}(Y)) = Var(\phi_i^{**}(X)) = 1$.

et de même :

$$E\big[E[\theta_i^*(Y) \,|\, X] \,|\, Y\big] = {\rho_i^*}^2 \theta_i^*(Y)$$

Par conséquent, les fonctions canoniques sont les vecteurs propres d'un opérateur linéaire (une "double espérance conditionnelle") et les carrés des coefficients de corrélation canonique sont les valeurs propres de ce même opérateur. Cette propriété, très importante, permettra d'estimer les triplets canoniques.

2.1.2 Les deux vecteurs X et Y sont continus

Lorsque X et Y sont deux vecteurs continus, qu'ils possèdent une densité par rapport à la mesure de Lebesgue, et sous des conditions non contraignantes [7], il est possible d'obtenir le résultat suivant [8] :

Théorème 2. *Soit $f_{(X,Y)}$ la densité du couple (X, Y), f_X la densité du vecteur X et f_Y la densité du vecteur Y. Soit $(\rho_i^*, \phi_i^*, \theta_i^*)$ l'analyse canonique du couple (X, Y), alors :*

$$\forall (x,y) \in \mathbb{R}^{p+q},\ f_{(X,Y)}(x,y) = f_X(x)\, f_Y(y)\Big(1 + \sum_i \rho_i^*\, \phi_i^*(x)\, \theta_i^*(y)\Big)$$

La densité conditionnelle de Y sachant $X{=}x$ est immédiate :

$$\forall y \in \mathbb{R}^q,\ f_{Y|X=x}(y) = f_Y(y)\Big(1 + \sum_i \rho_i^*\, \phi_i^*(x)\, \theta_i^*(y)\Big)$$

La densité du couple (X, Y) s'écrit alors comme "somme de produits". Il faut remarquer que X et Y sont indépendants si et seulement si tous les coefficients de corrélation canonique sont nuls, puisqu'alors :

$$f_{(X,Y)} = f_X \times f_Y$$

Le théorème 3 permet de calculer les moments conditionnels.

Théorème 3. *Soit $(\rho_i^*, \phi_i^*, \theta_i^*)$ l'analyse canonique du couple (X, Y). Y est de dimension 1 ($q = 1$). Pour $k \in I\!N^*$, si pour tout i, $E\,|\,Y^k\theta_i^*(Y)\,| < \infty$, alors le k^e moment conditionnel est donné par :*

$$E\big[Y^k \,|\, X\big] = E\big(Y^k\big) + \sum_i \rho_i^*\, E\big(Y^k \theta_i^*(Y)\big)\, \phi_i(X)$$

Corollaire. *En particulier, si $\mu_{i,1} = E\big(Y\theta_i(Y)\big)$ et $\mu_{i,2} = E\big(Y^2\theta_i(Y)\big)$, alors l'espérance et la variance conditionnelles de Y sachant X sont données par :*

$$m_Y(X) = E\,[Y \mid X] = E(Y) + \sum\nolimits_i \rho_i^* \mu_{i,1} \phi_i(X)$$
$$\sigma_Y^2(X) = Var(Y \mid X) = h_Y(X) - m_Y(X)^2$$

---

7. *cf.* BESSE *in* DROESBEKE *et alii* (1992) pour une description plus précise de ces conditions.

8. *cf.* TARAMASCO (1993) pour la démonstration

avec $h_Y(X) = E[Y^2 \mid X] = E(Y^2) + \sum_i \rho_i^* \mu_{i,2} \phi_i(X)$, *le moment conditionnel d'ordre 2.*

Ce théorème permettra d'obtenir une représentation conditionnellement hétéroscédastique du processus des rentabilités.

2.2 L'estimation des triplets canoniques

Jusqu'à présent, seules les propriétés probabilistes de l'A.C. ont été présentées. Intéressons-nous maintenant à l'estimation statistique des fonctions canoniques à partir de la donnée d'une trajectoire $(R_t, t = 0, \dots, T)$ des rentabilités.

L'estimation fonctionnelle peut emprunter deux voies. La première, de type paramétrique, choisit de limiter la recherche dans une famille de fonctions prédéterminée et estime les paramètres identifiant ces fonctions. La seconde est de type non paramétrique (aucune hypothèse n'est faite sur la nature de la fonction à estimer).Elle consiste à estimer la valeur de cette fonction en chacun des k points d'un ensemble $\{x_1, x_2, \dots, x_k\}$ préalablement choisi. Le résultat obtenu est alors le vecteur $\left\{\widehat{f}(x_1), \widehat{f}(x_2), \dots, \widehat{f}(x_k)\right\}$ des valeurs estimées de f aux k points. Le problème fonctionnel est donc transformé en un problème vectoriel et permet l'utilisation du calcul matriciel. Dans cet article, c'est cette optique qui est utilisée : les fonctions canoniques duales ϕ_i^* et θ_i^* sont estimées sur les deux ensembles de points respectifs $\{x_1, x_2, \dots, x_k\}$ et $\{y_1, y_2, \dots, y_l\}$. Ces fonctions sont caractérisées par les deux vecteurs[9] :

$$\Phi_i^* = \left\{\widehat{\varphi_i^*}(x_1),\ \widehat{\varphi_i^*}(x_2), \dots,\ \widehat{\varphi_i^*}(x_k)\right\}$$

$$\Theta_i^* = \left\{\widehat{\theta_i^*}(y_1),\ \widehat{\theta_i^*}(y_2), \dots,\ \widehat{\theta_i^*}(x_l)\right\}$$

2.2.1 La méthode de calcul

Pour extraire l'ensemble de vecteurs Φ_i^* et Θ_i^*, il faut procéder en deux étapes. Dans la première, il faut obtenir les $k \times p \times l \times q$ estimateurs $\widehat{p}_{n,m}$ des probabilités $P(X = x_n, Y = y_m)$. Ces calculs sont menés par la *méthode du noyau.* Dans la seconde étape, le calcul des espérances conditionnelles du théorème 1 est remplacé par un calcul de moyenne (empirique) conditionnelle.

9. Si la dimension de X ou de Y est supérieure à 1, il faudra "aplatir" les vecteurs. Par exemple, si X est de dimension 2, le vecteur Φ_i^* est de dimension $2 \times k$ et est égal à

$\Phi_i^* = \{\widehat{\varphi_i^*}(x_{11}, x_{12}), \dots, \widehat{\varphi_i^*}(x_{11}, x_{k2}), \widehat{\varphi_i^*}(x_{21}, x_{12}), \dots, \widehat{\varphi_i^*}(x_{k1}, x_{k2})\}$

2.2.2 Estimation des probabilités empiriques

La méthode la plus simple pour obtenir les probabilités $\widehat{p}_{n,m}$ consiste à partager l'étendue de X et Y en classes, et à calculer la fréquence dans chacune de ces classes. Lorsque cette méthode est appliquée aux deux rentabilités successives (R_t, R_{t-1}), le résultat obtenu est celui de LOUVET et TARAMASCO (1993). Dans cet article, une méthode plus générale est employée : *la méthode du noyau*[10]. La densité du vecteur V au point v, à partir d'un échantillon v_t, $t = 1 \ldots T$ est estimé par :

$$\widehat{f}_h(v) = \frac{1}{Th_T} \sum_{t=1}^{T} K\left(\frac{v_t - v}{h_T}\right)$$

avec

- h_T : paramètre de lissage positif,
- v_t : t^e observation.

Dans cet article, le noyau K est un produit de noyaux d'Epanechnikov. Le paramètre de lissage h est fixé à $\dfrac{\widehat{\sigma}}{T^{1/(4+r)}}$ où r est la dimension du vecteur V (2 ou 3 dans les applications qui suivront) et $\widehat{\sigma}$ est l'écart type empirique des rentabilités[11].

Cette méthode présente deux avantages par rapport au calcul simple de la fréquence : premièrement, le calcul peut être fait sur un plus grand nombre de points, et, deuxièmement, les fonctions obtenues sont plus "lisses".

2.2.3 Calcul des estimateurs

Une fois obtenues les probabilités empiriques $\widehat{p}_{n,m}$ du couple (X, Y), il est facile de calculer les probabilités empiriques $\widehat{q}_n$ du vecteur X et $\widehat{r}_m$ du vecteur Y par la formule classique :

$$\forall n \in 1, \ldots, k, \; \widehat{q}_n = \sum_{j=1}^{l} \widehat{p}_{n,j}$$

$$\forall m \in 1, \ldots, l, \; \widehat{r}_m = \sum_{j=1}^{k} \widehat{p}_{j,m}$$

Soient alors $\widehat{P}_{X,Y}$ la matrice des $\widehat{p}_{n,m}$, $\widehat{P}_X$ la matrice diagonale dont les éléments diagonaux sont les $\widehat{q}_n$ et $\widehat{P}_Y$ est la matrice diagonale dont les

10. cf. BOSQ et LECOUTRE (1987) pour une description plus complète de cette méthode.

11. Cette valeur est proportionnelle à la valeur de h_T (inconnue) qui minimise l'erreur quadratique intégrée.

éléments diagonaux sont les $\widehat{r}_m$. En remplaçant dans le théorème 1 le calcul des espérances conditionnelles par leur estimation, il vient :

$$\widehat{P}_Y^{-1}\ {}^t\widehat{P}_{X,Y}\ \widehat{P}_X^{-1}\ \widehat{P}_{X,Y}\ \Theta_i^* = \widehat{\rho}_i^{*\,2}\ \Theta_i^*$$

et

$$\widehat{P}_X^{-1}\ \widehat{P}_{X,Y}\ \widehat{P}_Y^{-1}\ {}^t\widehat{P}_{X,Y}\ \Phi_i^* = \widehat{\rho}_i^{*2}\ \Phi_i^*$$

Ces deux dernières formules sont celles de l'analyse factorielle des correspondances (A.F.C.) du tableau $\widehat{P}_{X,Y}$: les valeurs propres de l'A.F.C. sont alors les carrés des coefficients de corrélation canonique, et les vecteurs propres de l'A.F.C. les fonctions canoniques.

2.3 Les données

L'analyse canonique est appliquée entre la rentabilité à la date t : $Y_t = R_t$ et le vecteur des n rentabilités du passé : $X_t = \underline{R_{t-1}} = \begin{pmatrix} R_{t-1} \\ R_{t-2} \\ \dots \\ R_{t-n} \end{pmatrix}$

Le paramètre n sera choisi par la suite égal à 1 ou 2.

L'étude porte sur neuf indices internationaux représentatifs des marchés américain, asiatique et européen. Les données ont été collectées sur la base de données Datastream[12]. La période d'étude s'étend de 1969 à 1995 sauf exception. Les taux de rentabilité sont calculés de la manière suivante :

$$R_t = ln\left(\frac{I_t}{I_{t-1}}\right)$$

où I_t est la valeur de marché de l'indice à la date t. Tous les indices sont calculés avec des cours de clôture à l'exception de l'indice SBF (France) qui retient des cours d'ouverture.

Pour pouvoir appliquer l'analyse canonique ou calculer les moments conditionnels de la distribution de R_t, une seule hypothèse est nécessaire : la stationnarité de cette distribution. Certes, de nombreuses raisons portent à croire qu'il s'agit d'une hypothèse forte ; on peut néanmoins espérer qu'une grande part de la non-stationnarité apparente est en fait imputable à la dépendance mal prise en compte. Pour tous les indices, une étude générale est menée sur la dépendance entre deux rentabilités journalières successives. Une étude plus particulière est consacrée aux indices SP (Etats-Unis) et SBF (France). Pour ces indices, la période globale a été découpée en trois sous-périodes : 69-76, 77-84 et 85-fin, afin d'étudier la stabilité de la dépendance.

12. Par exception, les indices du Japon nous ont été fournis par Yamachi Securities (Genève).

III. Résultats

L'analyse canonique peut être conduite en retenant un vecteur de rentabilités passées de dimension quelconque. Il en est de même pour l'estimation des moments conditionnels. Plus la dimension de ce vecteur est grande, mieux il rend compte du passé et plus les chances sont grandes de capter la dépendance réelle. Toutefois, l'interprétation des facteurs de dépendance en devient plus difficile. Pour faciliter l'analyse des résultats, il est préférable d'adopter une démarche progressive en limitant d'abord l'information passée à la rentabilité la plus récente puis en enrichissant l'étude par l'ajout d'autres rentabilités passées. Dans cet esprit, les résultats qui sont présentés dans la section 3.1 portent exclusivement sur la dépendance entre la rentabilité d'un jour R_t et celle de la veille R_{t-1}. La valeur des moments conditionnels qui en découle sera présentée dans la section 3.2. Les sections suivantes proposeront diverses extensions de l'analyse susceptibles d'affiner les résultats, parmi lesquelles la prise en compte de deux rentabilités passées.

3.1 La dépendance entre deux rentabilités journalières consécutives

Le tableau 1 reprend l'ensemble des résultats de l'analyse canonique menée sur chacun des indices internationaux. La première constatation qui s'impose est qu'il suffit de deux facteurs pour expliquer la quasi-totalité de la dépendance (plus de 95 % de l'inertie quel que soit l'indice étudié[13]). La dépendance ne présente donc pas un profil très complexe d'autant que le premier facteur explique la plus grande part de cette inertie (entre 54 % et 84 % selon l'indice). La simplicité de la dépendance, et ce quel que soit l'indice, laisse à penser que l'analyse détecte un phénomène bien réel et non quelques anomalies ou coïncidences comme en recèle toute série statistique.

---

13. C'est pour cette raison que nous ne reproduisons pas les résultats concernant les facteurs suivants qui n'expliquent, chacun, qu'une part infime de l'inertie.

Tableau 1

Coefficients de corrélation canonique sur $(R_{t-1} - R_t)$ journaliers

Indice	1er facteur (a)	Effet	2e facteur	Effet
Standard & Poor's	0,098 62,2 %	variance 62,2 %	0,073 34,7 %	variance 97,0 %
DJI	0,084 73,1 %	linéaire 73,1 %	0,050 25,7 %	variance 98,8 %
Toronto SE	0,184 77,9 %	linéaire 77,9 %	0,129 21,2 %	variance 99,1 %
Tokyo SE	0,177 62,4 %	variance 62,4 %	0,096 32,9 %	linéaire 95,3 %
Nikkei	0,168 83,4 %	variance 83,4 %	0,067 13,1 %	linéaire 96,5 %
Sydney	0,156 78,7 %	linéaire 78,7 %	0,077 18,9 %	variance 97,6 %
FTSE	0,149 72,6 %	linéaire 72,6 %	0,089 26,2 %	variance 98,8 %
SBF	0,152 65,9 %	linéaire 65,9 %	0,106 32,1 %	variance 98,- %
FAZ	0,096 54,1 %	linéaire 54,1 %	0,086 43,5 %	variance 97,6 %

(a) Pour chaque indice, et pour chaque facteur, la première ligne reprend la valeur du coefficient de corrélation canonique et le type d'effet induit par le facteur ; la seconde ligne reprend l'inertie expliquée par le facteur ainsi que l'inertie cumulée.

Les coefficients de corrélation canonique, dont l'interprétation est analogue à celle de coefficients de corrélation linéaire, oscillent entre 0,08 et 0,19 pour le premier facteur, et entre 0,05 et 0,13 pour le second. Pour les marchés américain et allemand, le degré de dépendance est relativement faible puisque le plus fort coefficient n'atteint pas 0,10. Ces deux marchés démontrent une certaine efficience dans le sens où l'information contenue dans la rentabilité la plus récente affecte peu le niveau de la rentabilité future. En revanche, le degré de dépendance est plus important sur les autres marchés, qui sont donc à cet égard moins efficients, le premier coefficient étant compris entre 0,15 et 0,19, et le second pouvant dépasser 0,10, comme c'est le cas pour les marchés canadien et français. L'importance du phénomène n'est pas la même pour tous les marchés.

L'interprétation de chacun de ces deux facteurs en nous renseignant sur la nature précise du phénomène peut nous en livrer les causes éventuelles. L'analyse non paramétrique ne permet pas naturellement de connaître la forme analytique des fonctions de dépendance. En revanche, elle permet de les représenter graphiquement (à l'aide d'une estimation point par point). C'est grâce à cet examen graphique que nous pouvons caractériser l'impact de ces fonctions de dépendance sur la distribution de probabilité conditionnelle de la rentabilité.

Les fonctions de dépendance des divers indices présentent de nombreuses similitudes. En premier lieu, pour chaque facteur, quel que soit l'indice étudié, les deux duaux sont identiques. Cela revient à dire que la corrélation est maximale lorsqu'elle s'applique à une même transformation de R_{t-1} et de R_t ou encore que la distribution du couple (R_{t-1}, R_t) est symétrique :

$$\forall(x,y) \in \mathbb{R}^2,\ Prob(R_t < x \text{ et } R_{t-1} < y) = Prob(R_t < y \text{ et } R_{t-1} < x)$$

Et donc, sachant que, par hypothèse, $Prob(R_t < x) = Prob(R_{t-1} < x)$,

$$\forall(x,y) \in \mathbb{R}^2,\ Prob(R_t < x \,|\, R_{t-1} = y) = Prob(R_{t-1} < x \,|\, R_t = y)$$

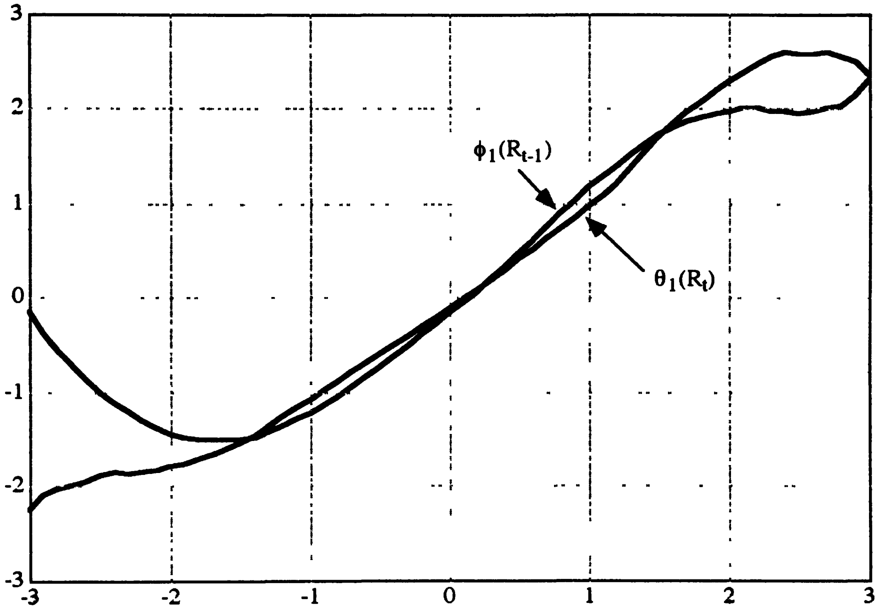
Cette symétrie tout à fait remarquable suggère une forme de réversabilité du temps boursier. La distribution de probabilité de la rentabilité d'un jour est conditionnée par la rentabilité de la veille de la même manière qu'elle l'est par celle du lendemain. Par conséquent, pour analyser la dépendance entre les rentabilités successives, il n'est pas nécessaire de savoir si l'on regarde vers le futur ou vers le passé. Le sens financier de cette symétrie reste énigmatique.

En second lieu, certaines similitudes peuvent être observées dans le profil même des fonctions. Elles sont, à quelques exceptions près, d'une lecture facile. En guise d'illustration, nous reprenons deux cas extrêmes : celui du SBF (*cf.* graphiques 1 et 2) et celui du SP (*cf.* graphiques 3 et 4). Dans l'interprétation de ces graphiques, deux précautions doivent être prises :

- Le graphique peut être lu indistinctement en opérant une symétrie par rapport à l'axe des abcisses. En effet, il faut se rappeler qu'il représente la valeur des fonctions $\phi(R_{t-1})$ et $\theta(R_t)$ de corrélation maximale. Il est facile de se rendre compte que les fonctions $-\phi(R_{t-1})$ et $-\theta(R_t)$ admettent la même corrélation ;
- Toute l'étude repose sur une estimation correcte de la distribution de probabilité du couple (R_{t-1}, R_t) dont le domaine de définition est $\mathbb{R}^2$. Tout échantillon empirique est nécessairement borné et ne peut rendre compte fidèlement de la distribution de variables définies sur un espace infini. Si l'estimation de la distribution n'en est guère affectée en son centre, elle l'est beaucoup plus fortement dans les queues. C'est pourquoi la forme des fonctions de dépendance[14] doit être interprétée en se concentrant sur le centre et en occultant la périphérie.

14. Il en sera de même pour l'interprétation des moments conditionnels.

Graphique 1 : ***Premier dual de l'indice*** SBF

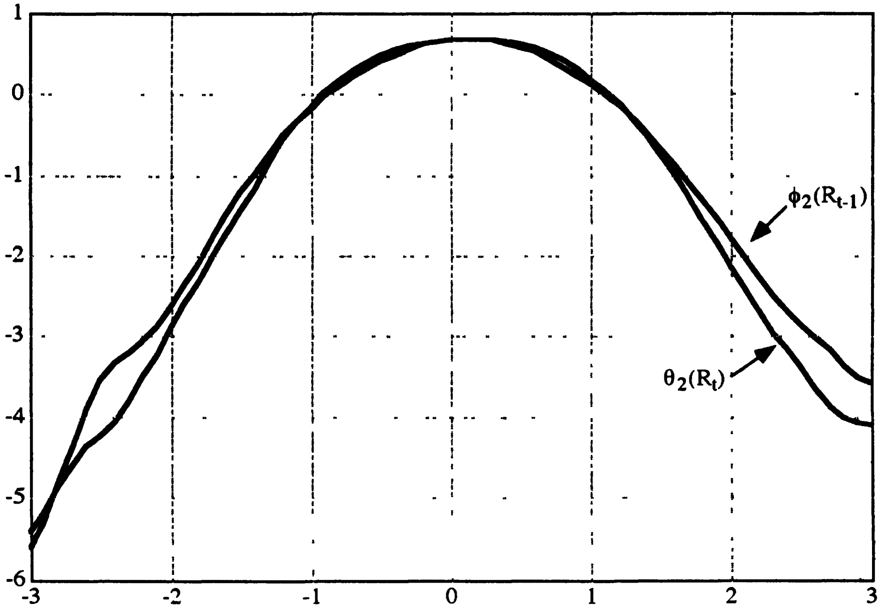


Comme on peut le lire sur le tableau 1, pour l'indice SBF, les deux premiers facteurs ont un impact important sur la dépendance (98 % de l'inertie). On ne peut donc en négliger aucun, même si l'un prend une part deux fois plus grande que l'autre. Le premier facteur (*cf.* graphique 1) admet une forme linéaire assez évidente dans la zone de rentabilité médiane. Celle-ci est par ailleurs centrée sur une valeur proche de 0. Approximativement, le profil des deux fonctions est de type :

$$\phi_1(R_{t-1}) = R_{t-1} \text{ et } \theta_1(R_t) = R_t$$

Par conséquent, l'indice SBF est en premier lieu affecté par une dépendance linéaire qui amène la rentabilité d'un jour à reproduire le niveau de la veille par une forme d'effet d'entraînement (nous parlerons par la suite d'un effet linéaire). Ce phénomène peut s'expliquer par une forme de paresse du marché qui tend à poursuivre la tendance de la veille : la morosité entraîne la morosité, l'euphorie entraîne l'euphorie. La dépendance linéaire permet de penser que l'espérance conditionnelle peut être appréhendée par une fonction linéaire de R_{t-1} ; en d'autres termes, que le processus est autorégressif.

Graphique 2 : ***Second dual de l'indice*** SBF



Le second facteur est de forme parabolique et centré sur 0. En tenant compte de l'échelle, une bonne approximation de ce facteur peut être obtenue par :

$$\phi_2(R_{t-1}) = R_{t-1}^2 \text{ et } \theta_1(R_t) = R_t^2$$

Cette dépendance de type quadratique exprime le lien entre l'amplitude des variations successives. Elle agit particulièrement lorsque les rentabilités atteignent des niveaux extrêmes. Si l'indice subit un jour une forte variation, il faut s'attendre le lendemain à ce qu'il répète une variation aussi forte dans le même sens ou en sens inverse. Cette variation est probablement due à la difficulté que le marché a à assimiler des informations importantes. Procédant par tâtonnements successifs, il peut soit sur-réagir en première analyse puis corriger le lendemain, ou au contraire sous-réagir et prolonger le mouvement ensuite. Ce phénomène induit une persistance entre les niveaux de variance (nous parlerons par la suite d'un effet variance). La variance de la rentabilité étant conditionnée par le carré de la rentabilité de la veille, le processus peut être décrit par un processus de type ARCH. Au total, sachant que les autres facteurs ont une influence négligeable, l'analyse précédente permet de penser que la dépendance peut être captée par les deux premiers moments de la distribution conditionnelle de la rentabilité.

L'indice SBF est exemplaire pour la simplicité des fonctions de dépendance. Les autres indices offrent des configurations souvent aussi limpides, quelquefois moins[15]. Globalement, il est possible de repérer l'effet variance sur tous les

15. Ces graphiques peuvent être obtenus auprès des auteurs.

indices et l'effet linéaire sur la plupart (le tableau 1 indique indice par indice quel est l'effet capté par chaque facteur). Tantôt le facteur principal est l'effet variance, tantôt c'est l'effet linéaire.

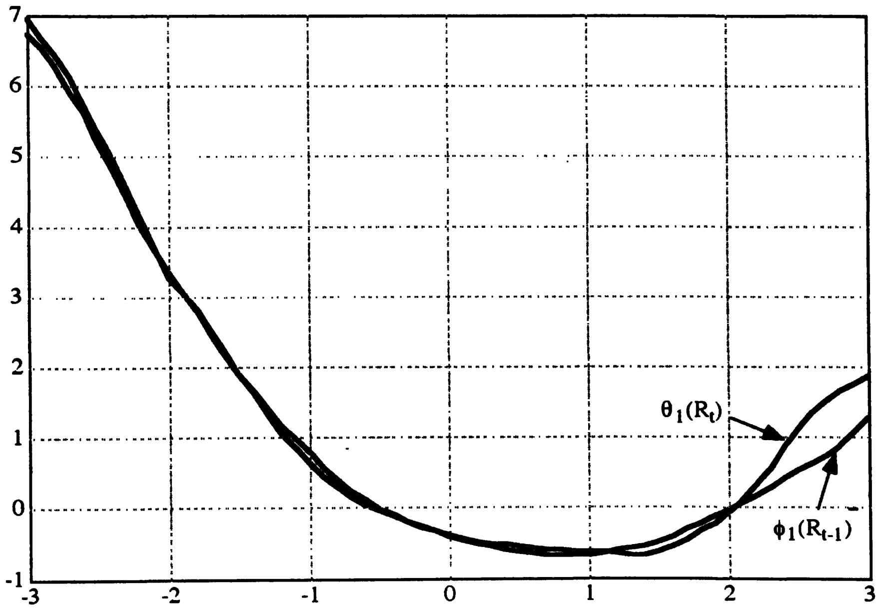
Par rapport aux autres indices, le SP offre une configuration originale. Les deux couples de fonctions canoniques (*cf.* graphiques 3 et 4) admettent une allure parabolique décentrée. Le premier couple est centré sur une rentabilité voisine de 1 % ; le second, de manière symétrique, sur une rentabilité voisine de -1 %. S'il fallait proposer une expression analytique de ces fonctions, nous pourrions choisir :

$$\phi_1(R_{t-1}) = (R_{t-1} - 1)^2 \text{ et } \theta_1(R_t) = (R_t - 1)^2$$

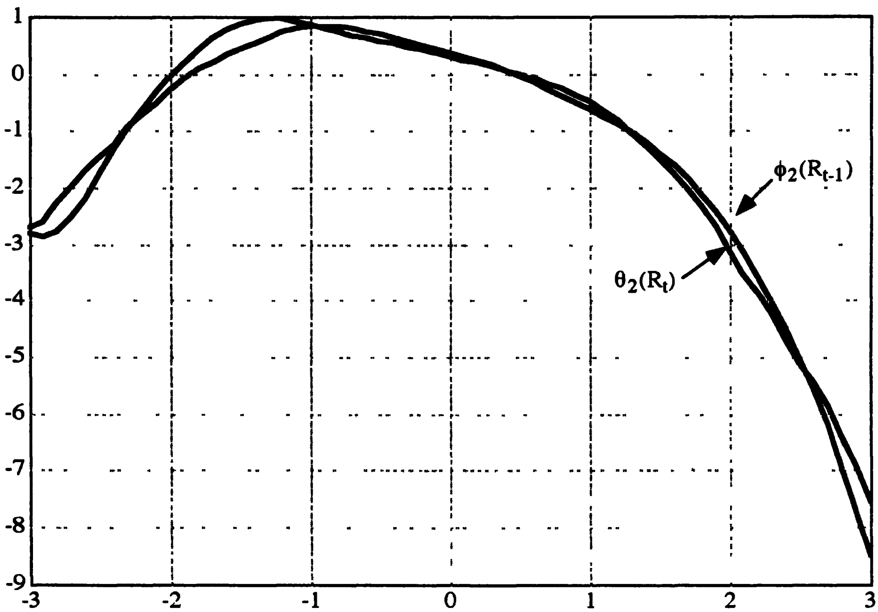
$$\phi_2(R_{t-1}) = (R_{t-1} + 1)^2 \text{ et } \theta_2(R_t) = (R_t + 1)^2$$

En quelque sorte, le premier facteur pousse la rentabilité d'un jour sur l'autre à reproduire l'amplitude de son écart par rapport à un taux cible de +1 % ; le second par rapport à un taux cible de -1 %. Chercher les ressorts de ce comportement est hasardeux, d'autant qu'il ne faut pas oublier que, pour cet indice, la dépendance est faible et donc que la nature des fonctions de dépendance présente moins d'intérêt. En tout état de cause, il paraît beaucoup plus difficile de rendre compte de cette dépendance à l'aide d'un modèle ARCH classique.

Graphique 3 : ***Premier dual de l'indice*** SP



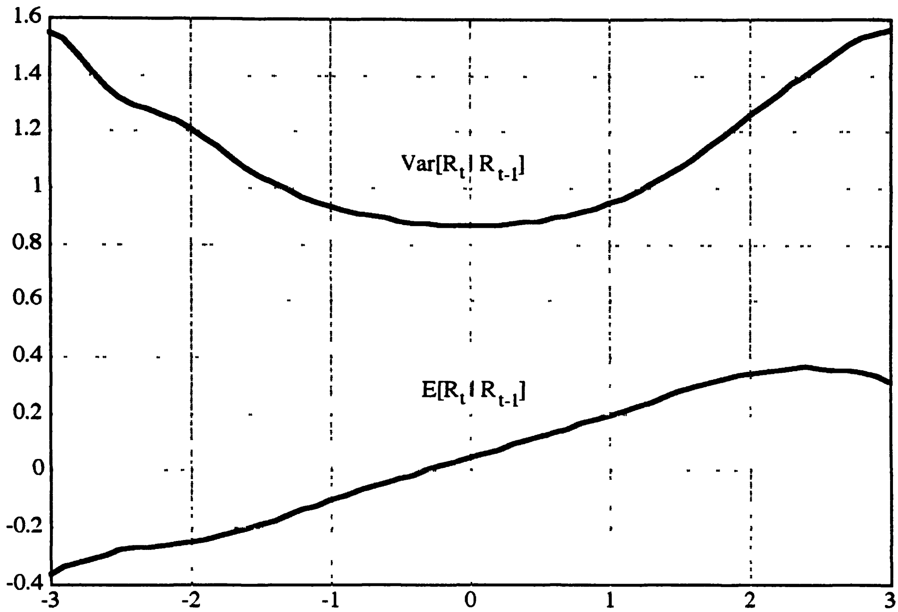
Graphique 4 : ***Second dual de l'indice*** SP



3.2 L'estimation des moments conditionnels de la rentabilité journalière

La nature des fonctions canoniques pour la plupart des indices suggèrent que les deux premiers moments conditionnels de la distribution de R_t dépendent de R_{t-1}. La valeur de ces moments peut être calculée à l'aide des fonctions canoniques, comme nous l'avons indiqué dans la partie précédente (*cf.* 2.1.2). Il est préférable toutefois d'employer une démarche directe : puisque nous disposons d'une estimation de la distribution du couple (R_t, R_{t-1}), il est facile de calculer toutes les caractéristiques de cette distribution et en particulier les moments de la distribution conditionnelle de R_t, sachant R_{t-1}. Les graphiques 5 et 6 représentent l'espérance et la variance conditionnelles en fonction de R_{t-1} pour les indices SBF et SP. Les moments conditionnels pour les autres indices donnent lieu à une interprétation identique (ils figurent en annexe).

Graphique 5 : $E[R_t \mid R_{t-1}$ **et** $Var\, R_t \mid [R_{t-1}]$ ***pour l'indice*** SBF



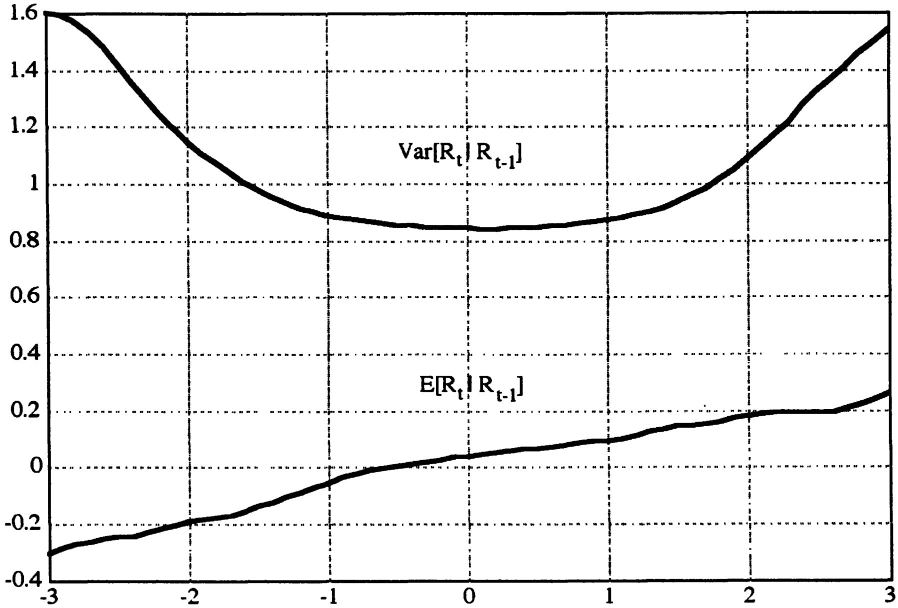
Comme attendu, l'espérance conditionnelle de l'indice SBF est une fonction linéaire et sa variance conditionnelle une fonction quadratique de la rentabilité précédente. A titre indicatif, on pourrait approcher la valeur de ces moments par les fonctions suivantes :

$$E[R_t \mid R_{t-1}] = 0,04 + 0,15R_{t-1}$$

$$Var[R_t \mid R_{t-1}] = 0,85 + 0,09R_{t-1}^2$$

Pour ce qui concerne la forme de ces moments, l'interprétation nous renvoie à ce qui a été dit lors de l'analyse des fonctions de dépendance : l'espérance conditionnelle est d'autant plus forte que le marché est déjà orienté à la hausse la veille et la variance conditionnelle est plus grande lorsque le marché vient de subir une forte variation à la hausse ou à la baisse. Quant à la taille du phénomène, on peut noter que les moments sont substantiellement affectés par le niveau de rentabilité de la veille, puisque l'espérance varie dans une plage de [-0,4 ; +0,4] et la variance entre 0,8 et 1,6.

Graphique 6 : $E[R_t \mid R_{t-1}$ ***et** $Var\, R_t \mid [R_{t-1}]$ **pour l'indice*** SP



Il est plus surprenant de trouver une même configuration pour les moments conditionnels de l'indice SP, compte tenu de la différence observée entre les fonctions canoniques des deux indices. Certes, on peut noter que l'espérance conditionnelle est légèrement moins sensible à la valeur de R_{t-1}, ce qui peut s'expliquer par l'absence d'un facteur de dépendance linéaire clairement identifié. Ce résultat surprenant tend à prouver que la dépendance ne peut être résumée par l'écriture des deux premiers moments conditionnels.

3.3 La sensibilité au pas de calcul des rentabilités

Comme pour toute autre analyse de données, les résultats de l'analyse canonique sont d'autant plus significatifs qu'ils portent sur un échantillon de grande taille. C'est pourquoi nous avons retenu jusqu'ici des rentabilités journalières. Il convient toutefois de s'interroger sur les implications de ce choix. L'objectif fondamental de l'étude est de comprendre en quoi le comportement futur d'un cours boursier dépend de sa trajectoire passée et quelle est la logique financière qui sous-tend ce phénomène. Pour en comprendre les raisons, il est essentiel de mesurer la dépendance à sa source.

L'analyse sur des rentabilités journalières nous a permis de donner une interprétation simple de leur dépendance sérielle. Ceci suggère que notre choix était pertinent. Pour le vérifier, il suffit de reprendre l'analyse en modifiant le pas de calcul et d'observer si la nature de la dépendance se simplifie encore ou au contraire se complexifie. L'analyse a été appliquée à des rentabilités cumulées sur deux jours, puis à des rentabilités hebdomadaires.

Tableau 2

Coefficients de corrélation canonique
sur $R_{t-1} - R_t$ journaliers, bi-journaliers et hebdomadaires

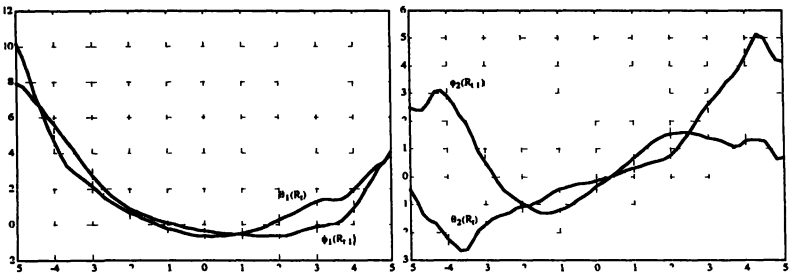
Indice	Périodicité	1er facteur (a)	Effet	2e facteur	Effet
SP	journalière	0,098 62,2 %	variance 62,2 %	0,073 34,7 %	variance 97,- %
	bi-journalière	0,069 80,- %	variance 80,- %	0,030 15 %	non caractérisé 95,- %
	hebdomadaire	0,099 89,7 %	variance 89,7 %	0,026 6,- %	non caractérisé 95,7 %
SBF	journalière	0,152 65,9 %	linéaire 65,9 %	0,106 32,1 %	variance 98,- %
	bi-journalière	0,097 66,1 %	variance 66,1 %	0,066 29,9 %	non caractérisé 96,- %
	hebdomadaire	0,117 59,3 %	variance 59,3 %	0,093 37,2 %	non caractérisé 96,5 %

(a) Pour chaque indice, et pour chaque facteur, la première ligne reprend la valeur du coefficient de corrélation canonique et le type d'effet induit par le facteur ; la seconde ligne reprend l'inertie expliquée par le facteur ainsi que l'inertie cumulée.

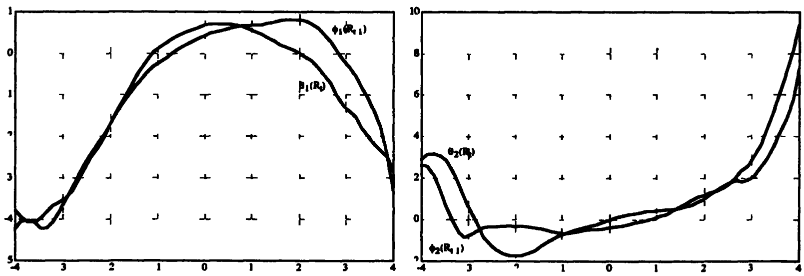
Lorsque l'on compare les résultats obtenus sur des rentabilités bijournalières (*cf* graphiques 7 et 8) avec ceux obtenus sur des données journalières, on note une baisse des coefficients de corrélation et une complexification de la forme des fonctions canoniques. De même, les moments conditionnels sont beaucoup moins sensibles à la rentabilité précédente. Cette dégradation suggère que les rentabilités bijournalières récupèrent, en l'altérant, la dépendance temporelle qui affecte les rentabilités journalières.

En revanche, les résultats retrouvent une certaine lisibilité pour les rentabilités hebdomadaires (*cf.* graphiques 9 et 10). Pour l'indice SP, seul l'effet sur la variance conditionnelle persiste. Pour l'indice SBF, l'effet variance devient dominant. On peut donc penser que les opérateurs tirent des informations de la dernière rentabilité hebdomadaire tout autant que de la dernière rentabilité journalière.

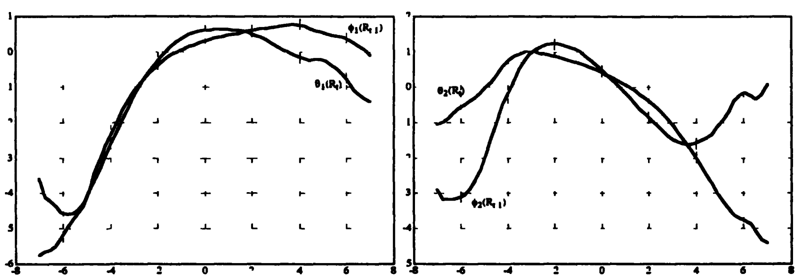
Graphique 7 : ***Résultats pour les rentabilités bi-journalières du*** SBF



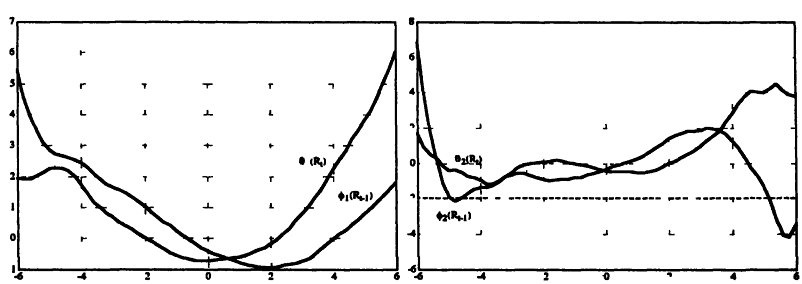
Graphique 8 : ***Résultats pour les rentabilités bi-journalières du*** SP



Graphique 9 : ***Résultats pour les rentabilités hebdomadaires du*** SBF



Graphique 10 : ***Résultats pour les rentabilités hebdomadaires du*** SP



3.4 La stabilité des fonctions de dépendance au cours du temps

Sur la période 1969-1995, les marchés financiers ont subi de nombreuses évolutions. La bourse française, jusqu'alors peu active, a pris son véritable essor dans les années 1980. Sur le plan international, le fonctionnement des marchés a gagné en efficience grâce à la systématisation de la gestion de portefeuille, en particulier de la gestion indicielle, et à l'augmentation du volume des échanges. Le développement des marchés dérivés a favorisé la spéculation sur la volatilité et a contraint les opérateurs à mieux anticiper l'évolution de la variance. Enfin, l'internationalisation et la globalisation financière tend à standardiser le comportement des marchés.

La modification de la dépendance temporelle rend compte de ces évolutions, comme on peut le constater à la lecture des résultats obtenus sur trois sous-périodes (*cf.* tableau 3 et graphiques 11 et 12). En premier lieu, le degré de dépendance décroît de manière significative sur les périodes les plus récentes, témoignant ainsi d'un gain d'efficience. On ne détecte presque plus de dépendance sur l'indice SP (les coefficients de corrélation tombent à 0,066 et 0,036) et les moments conditionnels sont quasiment plats. Il faut des niveaux de rentabilité extrêmes pour que la distribution conditionnelle soit remise en cause. En second lieu, la dépendance tend à s'unifier sur tous les marchés : l'effet variance devient dominant sur le SBF et l'effet linéaire tend à disparaître.

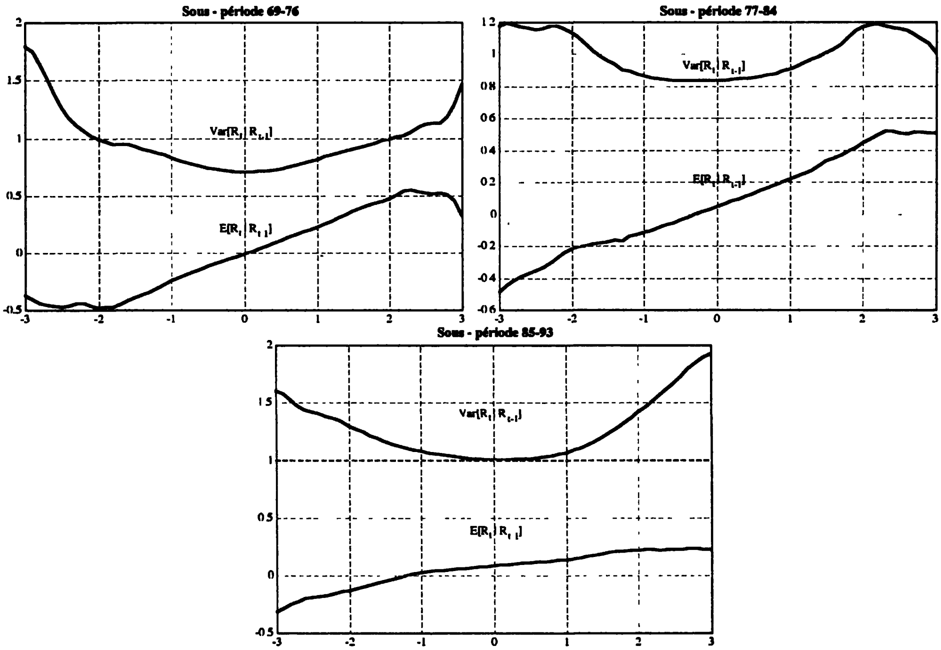
Tableau 3

Coefficients de corrélation canonique entre les rentabilités journalières sur les sous-périodes

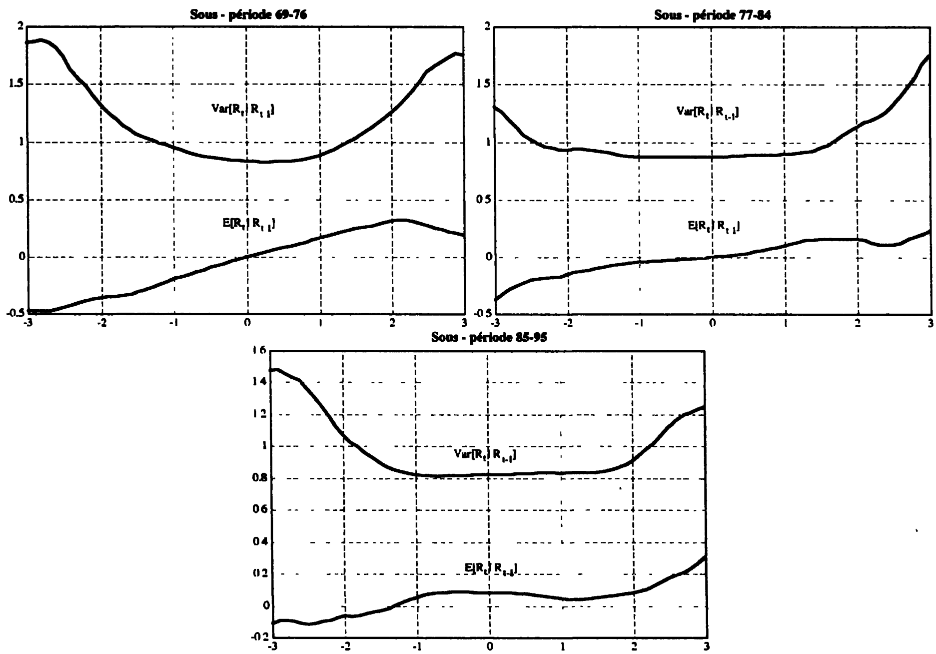
Indice	Période	1er facteur (a)	Effet	2e facteur	Effet
SP	1969-76	0,182 62,6 %	variance 62,6 %	0,134 33,9 %	variance 96,5 %
	1977-84	0,088 55,1 %	variance 55,1 %	0,061 26,5 %	variance 81,6 %
	1985-95	0,066 73,6 %	variance 73,6 %	0,036 21,7 %	non caractérisé 95,3 %
SBF	1969-76	0,237 73,4 %	linéaire 73,4 %	0,134 23,3 %	effet variance 96,7 %
	1977-84	0,177 78,3 %	linéaire 78,3 %	0,088 19,6 %	effet variance 97,9 %
	1985-93	0,100 58,7 %	variance 58,7 %	0,079 36,5 %	linéaire 95,2 %

(a) Pour chaque indice, et pour chaque facteur, la première ligne reprend la valeur du coefficient de corrélation canonique et le type d'effet induit par le facteur ; la seconde ligne reprend l'inertie expliquée par le facteur ainsi que l'inertie cumulée.

Graphique 11 : ***Moments conditionnels pour l'indice*** SBF



Graphique 12 : ***Moments conditionnels pour l'indice*** SP



3.5 La dépendance entre R_t et (R_{t-1}, R_{t-2})

Quelle information faut-il retenir de sa trajectoire passée pour prévoir le comportement futur d'un indice? Cette information est-elle toute entière contenue dans la rentabilité la plus récente ou requiert-elle au contraire une mémoire plus ancienne ? Pour le savoir, nous proposons maintenant de retenir du passé les deux dernières rentabilités (R_{t-1}, R_{t-2}) et d'étudier en quoi cet enrichissement permet d'affiner notre analyse. Les résultats comparés figurent dans le tableau 4 et les graphiques 13 et 14[16].

Tableau 4

Coefficients de corrélation canonique entre R_t et (R_{t-1}, R_{t-2})

Indice	Variables	1er facteur (a)	Effet	2e facteur	Effet
SP	R_t et R_{t-1}	0,098 62,2 %	variance 62,2 %	0,073 34,7 %	variance 97,- %
	R_t et (R_{t-1}, R_{t-2})	0.100 53,8 %	variance 53,8 %	0,086 39,9 %	linéaire 93,7 %
SBF	R_t et R_{t-1}	0,152 65,9 %	linéaire 65,9 %	0,106 32,1 %	variance 98,- %
	R_t et (R_{t-1}, R_{t-2})	0,146 62,5 %	linéaire 62,5 %	0,106 33,3 %	variance 95,8 %

(a) Pour chaque indice, et pour chaque facteur, la première ligne reprend la valeur du coefficient de corrélation canonique et le type d'effet induit par le facteur ; la seconde ligne reprend l'inertie expliquée par le facteur ainsi que l'inertie cumulée.

Les résultats ne sont guère modifiés par la prise en compte de R_{t-2}. Les coefficients de corrélation sont quasiment stables. La nature de l'effet reste la même : l'examen des fonctions canoniques $\theta_1(R_t)$ et $\theta_2(R_t)$ permet de constater que l'indice SBF subit principalement un effet linéaire mais aussi un effet sur la variance et que, pour l'indice SP, l'effet dominant porte toujours sur la variance, bien qu'un effet linéaire apparaisse comme deuxième facteur.

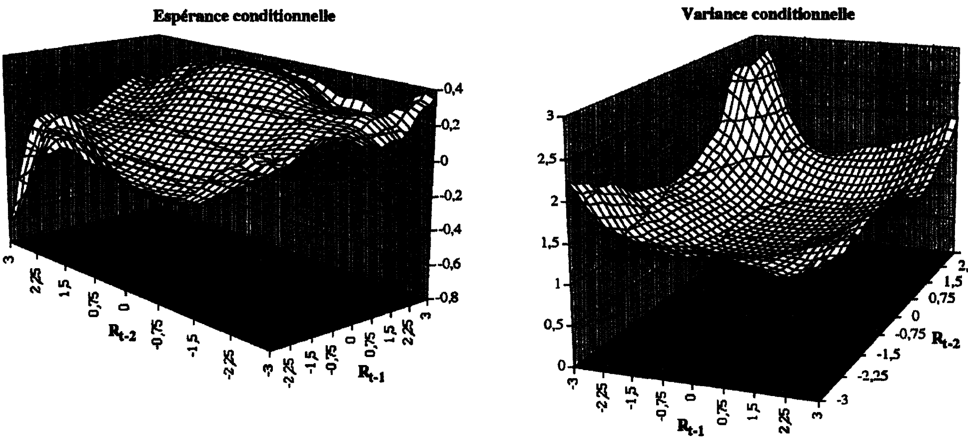
Pour l'indice SBF, la fonction canonique $\phi_1(R_{t-2}, R_{t-1})$ reste de forme linéaire et $\phi_2(R_{t-2}, R_{t-1})$ de forme parabolique ; R_{t-2} joue un rôle modeste sur ϕ_1 mais plus manifeste sur ϕ_2. En d'autres termes, l'ajout de R_{t-2} n'apporte pas de modifications substantielles dans la caractérisation de la dépendance linéaire mais permet d'affiner la prévision de la variance. Les conclusions sont les mêmes pour l'indice SP si l'on inverse l'ordre des facteurs.

Le profil des moments conditionnels confirme cette analyse. L'espérance conditionnelle est quasiment déterminée par R_{t-1} seul (ϕ_2 est un plan parallèle à R_{t-2}), alors que la variance conditionnelle, si elle est affectée principalement

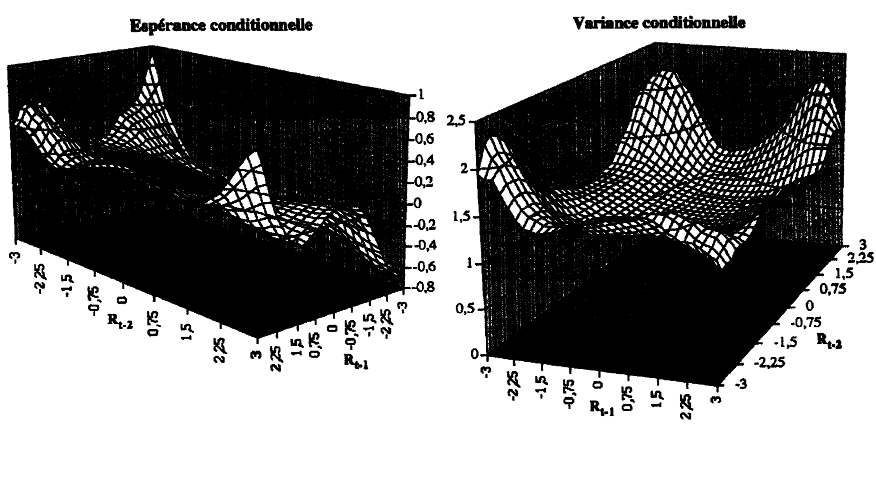
16. Il convient de préciser qu'en deux dimensions la qualité d'estimation diminue, ce qui explique par exemple que le coefficient de corrélation empirique de la première fonction soit plus faible lorsqu'on ajoute R_{t-2} dans le cas du SBF.

par R_{t-1}, dépend aussi de R_{t-2}. Au total, on vérifie que R_{t-1} joue un rôle prépondérant comme mémoire du passé et que le gain informationnel tiré R_{t-2}, et sans doute de rentabilités plus anciennes, est pauvre. L'utilisation d'un modèle AR(1)-ARCH(2) semble suffisant pour rendre compte de la dépendance temporelle.

Graphique 13 : *Moments conditionnels pour l'indice* SBF



Graphique 14 : *Moments conditionnels pour l'indice* SP



IV. Conclusion

Au cours de cette étude, nous nous sommes proposé de décrire la dépendance temporelle qui affecte les rentabilités boursières à l'aide de l'analyse canonique. Grâce à cette analyse, nous avons pu comprendre comment la distribution de probabilité de la rentabilité pouvait être conditionnée par les rentabilités passées. Appliquée à un échantillon de neuf indices internationaux sur la période 1969-1985, elle nous a permis de constater que la dépendance présente une structure simple et relativement similaire sur tous les marchés. Pour la résumer, on peut dire que le marché tend à reproduire la tendance et la variance qu'il a observées la veille. Sur le plan probabiliste, cette dépendance peut être exprimée en conditionnant les deux premiers moments de la distribution de R_t par R_{t-1} : l'espérance par une fonction linéaire et la variance par une fonction quadratique. Les séries de rentabilités journalières pourraient donc s'écrire sous la forme d'un processus AR(1)-ARCH(1). L'estimation directe des moments conditionnels nous a permis de vérifier la pertinence d'une telle écriture.

La dépendance a tendance à décroître : sur les années 1985-1995, le niveau de corrélation devient très faible et il ne subsiste quasiment plus qu'une dépendance sur les variances. Les marchés auraient-ils gagné en efficience sur cette période ?

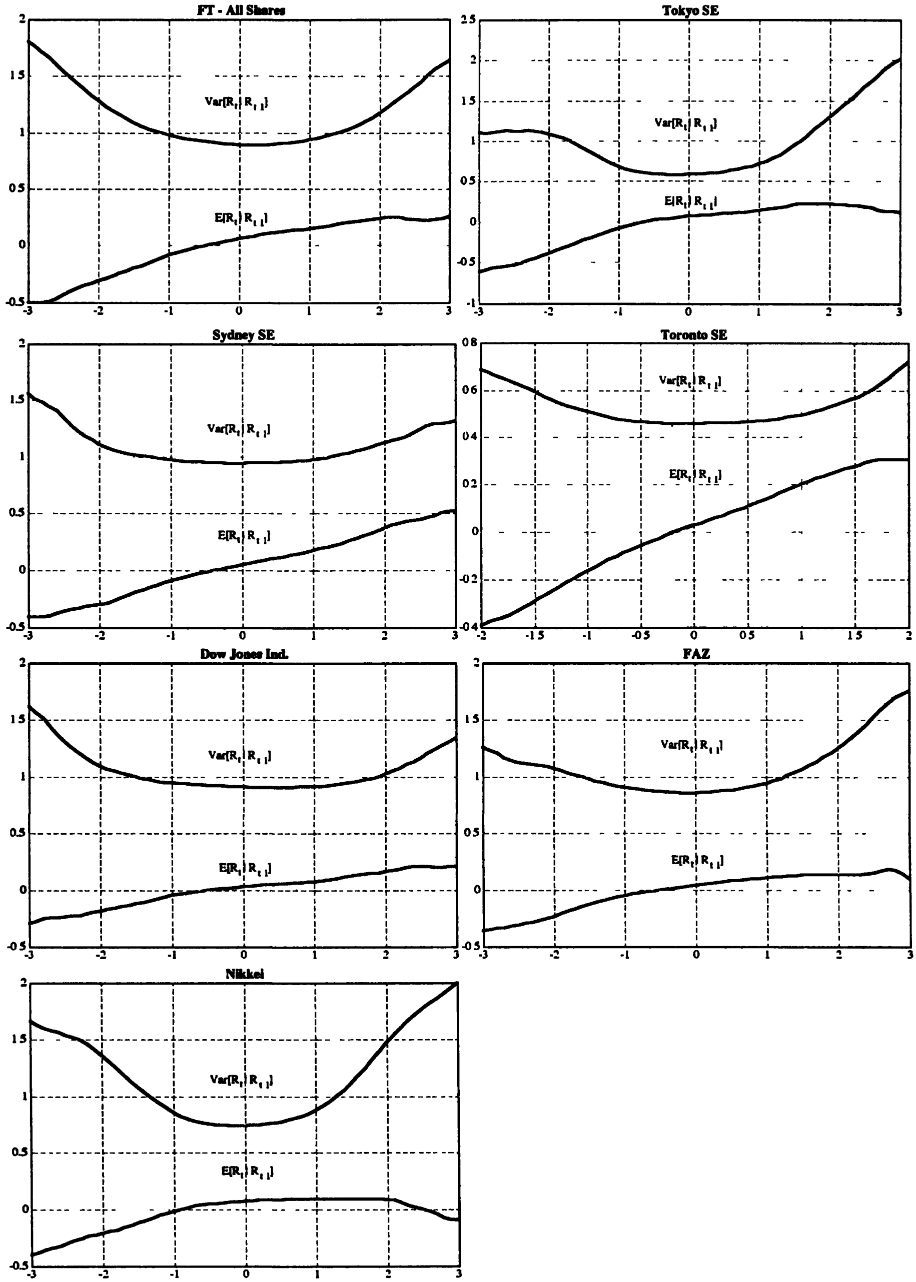
A l'image des fractales, lorsque l'on agrandit l'angle d'observation, on repère curieusement une même configuration des fonctions de dépendance alors que l'on aurait pu s'attendre à une complexification : les rentabilités hebdomadaires, en effet, offrent, elles aussi, une structure de dépendance simple et de type voisin à celle qui régit les rentabilités journalières. A notre sens, l'explication de ce phénomène tiendrait au comportement des opérateurs qui utiliseraient explicitement les niveaux de rentabilités passées pour réviser leurs anticipations, les uns se référençant à la rentabilité journalière, les autres à la rentabilité hebdomadaire. Mais il ne s'agit que d'une hypothèse et ce phénomène reste énigmatique.

Une autre énigme à laquelle nous n'avons pas apporté de réponse est celle de la réversibilité de la structure de dépendance. Si l'on se proposait d'expliquer la rentabilité d'aujourd'hui, non plus par le niveau des rentabilités passées, mais par le niveau des rentabilités futures, les mêmes fonctions pourraient être utilisées.

Lorsque l'on tente d'expliquer la rentabilité journalière par la conjugaison des deux rentabilités passées, on parvient à affiner l'estimation de la variance conditionnelle, mais l'on n'enrichit pas fondamentalement l'analyse. Un modèle AR(1)-ARCH(2), s'il est plus adéquat, ne doit pas accorder un poids significatif à la seconde rentabilité passée. La dernière rentabilité observée semble contenir à elle seule l'essentiel de l'information utile.

Annexe

Moments conditionnels pour les autres indices



Références

[1] BENZÉCRI J.P. (1973) *L'analyse des données, tome 2; l'analyse factorielle des correspondances*, Dunod.

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

[3] BOLLERSLEV T., CHOU R. Y. et KRONER K. F. (1992) "ARCH Modeling in Finance", *Journal of Econometrics* vol. 1, Annals, pp. 5-59.

[4] BOSQ D. et LECOUTRE J.P. (1987) *Théorie de l'estimation fonctionnelle*, Economica, coll. ESA, 342 pages.

[5] BREIMAN L. et FREIDMAN J. (1985) "Estimating Optimal Transformation for Multiple Regression and Correlation", *Journal of the American Statistical Association* vol. 80, n° 391, Septembre, pp. 580-598.

[6] DING Z., GRANGER C. et ENGLE R. (1993) "A Long Memory Property of Stock Market Returns and a New Model", *Journal of Empirical Finance,* vol. 1, n° 1, pp. 83.106.

[7] DROESBEKE J.-J., FICHET B. et TASSI P. (1992) *Modèles pour l'analyse des données multidimensionnelles*, Economica, coll. ASU, 363 pages.

[8] ENGLE R. (1982) "Autoregressive Conditionnal Heteroskedasticity", *Economica,* vol. 50, pp. 987-1008.

[9] GOURIÉROUX C. (1993) *Modèle* ARCH *et applications financières*, Economica, coll. ASU, 288 pages.

[10] HSIEH D. (1991) "Chaos and Nonlinear Dynamics; Application to Financial Markets", *Journal of Finance,* vol. 46, no 5, Décembre, pp. 1389-1877.

[11] LONGIN F. (1993) *Booms and Crashes. Application of Extreme Value Theoty to the US Stock Market*, Journée de l'Association Française de Finance, Décembre.

[12] LOUVET P. et TARAMASCO O. (1993) "Analyse descriptive de la dépendance sérielle sur les rentabilités boursières", *Finance,* vol. 14, n° 1, pp. 67-94

[13] NELSON D. (1990) "Conditionnal Heteroskedasticity in Asset Returns : a New Approach", *Economica,* vol. 59, pp. 347-370.

[14] TARAMASCO O. (1993) *Modélisation non-paramétrique du comportement des cours boursiers*, Thèse de l'Université, Université Joseph Fourier - Grenoble I, 154 pages.

[15] ZAKPIAN J-M. (1990) *Modèles hétéroscédastiques à seuil*, Document de travail n° 9012, CREST, 50 pages.