

2000s-44

**Étude du Modèle d'Évaluation
par Arbitrage sur le marché
des actions suisses**

Thierry Vessereau

Série Scientifique
Scientific Series



CIRANO
Centre interuniversitaire de recherche
en analyse des organisations

Montréal
Octobre 2000

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- Université Concordia
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- Université McGill
- MEQ
- MRST
- Alcan Aluminium Ltée
- AXA Canada
- Banque Nationale du Canada
- Banque Royale du Canada
- Bell Québec
- Bombardier
- Bourse de Montréal
- Développement des ressources humaines Canada (DRHC)
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Imasco
- Industrie Canada
- Pratt & Whitney Canada Inc.
- Raymond Chabot Grant Thornton
- Ville de Montréal

© 2000 Thierry Vessereau. Tous droits réservés. All rights reserved.

Reproduction partielle permise avec citation du document source, incluant la notice ©.

Short sections may be quoted without explicit permission, provided that full credit, including © notice, is given to the source.

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

This paper presents preliminary research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

Étude du Modèle d'Évaluation par Arbitrage sur le marché des actions suisses*

Thierry Vessereau[†]

Résumé / Abstract

Cet article étudie le Modèle d'Évaluation par Arbitrage (MEA) sur le marché suisse des actions. Pour déterminer les facteurs communs dans le modèle, la méthode de l'analyse en composantes indépendantes est employée. Cette méthode, reposant sur les réseaux neuronaux, permet d'inclure les informations des statistiques d'ordre élevé dans la décomposition. Son emploi sur le marché suisse des actions pour une période de vingt ans commençant le 1er janvier 1975 permet des améliorations sensibles dans l'étude du modèle multifactoriel, en comparaison avec la méthode classiquement employée d'analyse en composantes principales.

In this paper the application of Arbitrage Pricing Theory (APT) and multifactorial pricing is studied on the Swiss stock market. In order to estimate the factors used in the multifactorial model, it is proposed to use the new method of Independent Component Analysis. This method laying on neural networks allows including information on high-order statistics and dependences between distributions of stock returns. Its use on a twenty years period of swiss stock returns beginning January 1, 1975 allows significative improvement in the multifactorial model tests when comparing with the commonly used Principal Component Analysis.

Mots Clés : Modèle d'Évaluation par Arbitrage, Analyse en Composantes Indépendantes, réseaux neuronaux

Keywords: Arbitrage Pricing Theory, Independent Component Analysis, neural networks

* Adresse de correspondance : Thierry Vessereau, CIRANO, 2020 rue University, 25^e étage, Montréal, Qc, Canada H3A 2A5 Tél. : (514) 985-4000 poste 3035 Fax : (514) 985-4039 courriel : vesseret@cirano.umontreal.ca
Cet article est tiré du quatrième chapitre de ma thèse, soutenue à l'Université de Genève : *Application en finance des réseaux neuronaux artificiels : L'approche multifactorielle des marchés de l'évaluation non paramétrique des options*. Je remercie particulièrement mon directeur de thèse Jean Lefoll, ainsi que les membres du jury : Marc Chesney, Alain Haurie, Christian Pellegrini et Gilbert Ritschard.

[†] CIRANO

1 Le Modèle d'Evaluation par Arbitrage

Le Modèle d'Evaluation par Arbitrage (MEA), développé initialement par Ross (1976) [31], représente à la fois une alternative et par certains aspects une extension du Modèle d'Evaluation des Actifs Financiers (MEDAF). Contrairement à ce dernier modèle, selon lequel le rendement d'un actif est une fonction linéaire d'un facteur commun unique, le marché, le MEA repose sur l'hypothèse que les rentabilités \tilde{r}_i des N actifs dépendent linéairement d'un nombre restreint de K facteurs communs indépendants \tilde{f}_j :

$$\tilde{r}_i = E[\tilde{r}_i] + b_{i1}\tilde{f}_1 + b_{i2}\tilde{f}_2 + \dots + b_{iK}\tilde{f}_K + \tilde{\epsilon}_i, \quad (1)$$

pour $i = 1..N$ actifs, où b_{ij} est le coefficient de sensibilité de l'actif i au j -ème facteur commun \tilde{f}_j ; les erreurs $\tilde{\epsilon}_i$ sont d'espérance nulle et de variance finie et ne sont corrélées ni avec un facteur commun, ni avec l'erreur liée à un autre actif.

En prenant appui sur le principe de l'absence d'arbitrage systématique, Ross (1976) [31] montre qu'il existe donc un ensemble de $(K + 1)$ coefficients $\lambda_0, \lambda_1, \dots, \lambda_K$ tel que pour chacun des actifs i :

$$E[\tilde{r}_i] = \lambda_0 + \lambda_1 b_{i1} + \dots + \lambda_K b_{iK}, \quad (2)$$

où λ_0 représente la prime de risque d'un actif qui n'est sensible à aucun des facteurs du marché (actif hors-risque ou portefeuille zéro-beta), les λ_k les primes de risque liés à chaque facteur k , et les b_{ik} les sensibilités de l'actif i à ces facteurs. La prime de risque de l'actif i , $E[\tilde{r}_i] - \lambda_0$, est donc une combinaison des primes de risque sur les différents facteurs. Huberman (1982) [19], puis Ingersoll (1984) [20], proposent une approche complémentaire à l'approche de Ross en considérant une séquence d'économie avec un nombre croissant d'actifs risqués. Chamberlain et Rothschild (1983) [7] y apportent une contribution significative en considérant que les rendements des actifs suivent une structure factorielle approximative, c'est-à-dire que la matrice de covariance des erreurs du modèle n'est pas nécessairement strictement diagonale.

Les Modèles d'Evaluation par Arbitrage reposant uniquement sur le principe de non arbitrage ne proposent cependant pas dans leur dérivation un modèle d'équilibre de marché, puisque l'absence d'arbitrage n'implique pas nécessairement l'existence d'un équilibre. Les premières dérivations du MEA d'équilibre sont effectuées par Connor (1984) [11], Chen

et Ingersoll (1983) [8], et Grinblatt et Titman (1983) [16], en introduisant principalement l'hypothèse que le portefeuille de marché n'a pas de risque idiosyncratique dans le modèle factoriel considéré, et qu'il est possible pour un investisseur de construire un portefeuille optimal en éliminant les risques diversifiables sans restreindre ses choix. Avec cette hypothèse supplémentaire, la rentabilité anticipée d'un actif est alors une combinaison linéaire exacte des sensibilités de l'actif aux différents facteurs. Chen et Ingersoll (1983) [8] et Grinblatt et Titman (1987) [17] montrent qu'une condition nécessaire et suffisante à la dérivation du MEA en tant que modèle d'équilibre est qu'au moins l'un des portefeuilles de référence parfaitement corrélés aux facteurs soit une combinaison linéaire de portefeuilles efficients dans le cadre moyenne-variance. Enfin Wei (1987) [36] étudie le cas où le portefeuille de marché possède un risque idiosyncratique $\tilde{\epsilon}_{im}$ pour un modèle factoriel donné, c'est-à-dire, selon le sens que donne Wei à ce terme, qu'il est mal diversifié, et l'impact de cette existence de risque sur le modèle d'évaluation. Wei montre alors qu'une relation exacte peut être obtenue en introduisant dans la décomposition factorielle ce portefeuille de marché résiduel.

La principale difficulté de l'étude empirique du Modèle d'Evaluation par Arbitrage est de traiter l'hypothèse selon laquelle les rendements des actifs sont générés par une structure factorielle. En effet, ni les facteurs et les sensibilités des actifs aux différents facteurs, ni même le nombre de facteurs nécessaires, ne sont a priori déterminées. Au pire, il est toujours possible choisir le nombre de facteurs égal au nombre d'actifs pour que les rendements des actifs soient parfaitement expliqués historiquement. Le principal problème méthodologique, contrairement au MEDAF dont l'unique facteur est le marché, dont au moins un proxy peut être observé, est celui de la méthodologie à employer pour déterminer empiriquement la structure factorielle de générations des rendements.

La quasi-totalité des études du Modèle d'Evaluation par Arbitrage emploient pour l'estimation des portefeuilles de facteurs des techniques d'analyse factorielle, en particulier l'analyse en facteurs communs, comme dans l'étude initiale de Roll et Ross (1980) [30] ou Lehmann et Modest (1988) [22], et l'analyse en composantes principales, par exemple par Connor et Korajczyk (1988) [12]. La plupart de ces études concluent, si le modèle d'évaluation par arbitrage est accepté, à un nombre de facteurs communs déterminés empiriquement compris entre trois et cinq, avec une prédominance du premier facteur. Elles concluent que le MEA ne peut pas être rejetée, mais ne peuvent lever les incertitudes du modèle. Le principal

problème mis en évidence empiriquement est que le nombre des facteurs communs n'est pas stable selon les observations et les actifs employés. Dhrymes et al. (1984) [14] démontrent empiriquement que le nombre de facteurs augmente lorsque le nombre d'actifs sur le marché augmente, ce qui est contraire aux hypothèses même de la théorie. Trzcinka (1986) [34] montre d'une manière similaire que les facteurs ne sont pas stables si on augmente le nombre des actifs, et conclut que seul le premier facteur peut effectivement être retenu. Les résultats de Trzcinka ont cependant été très controversés, en particulier par Brown (1989) [5] et Connor et Korajczyk (1993) [13], qui expliquent les résultats de Trzcinka par le fait que son approche empirique serait incohérente. Enfin, l'approche autorégressive est appliquée au problème des facteurs communs initialement sous une forme semi-autorégressive puis sous une forme autorégressive par Mei (1993) [24], qui trouve par cette méthode qu'au moins sept facteurs sont nécessaires pour évaluer correctement l'évolution des rendements des actifs sur le marché des Etats-Unis.

L'absence de stabilité de la structure factorielle obtenue lorsque le nombre d'actifs augmente, ou suivant les différents sous-groupes d'actifs, comme cela est mis en évidence par Dhrymes et al. (1984) [14], est une critique fondamentale du Modèle d'Evaluation par Arbitrage puisque la décomposition factorielle correcte en est l'hypothèse principale. Le principal problème rencontré par ce modèle est en effet l'absence d'une méthode permettant d'obtenir effectivement la décomposition factorielle correcte.

Dans cet article nous proposons d'appliquer au problème de l'estimation des séries de facteurs communs dans le cadre du Modèle d'Evaluation par Arbitrage (MEA), une méthode récente d'analyse factorielle: l'Analyse en Composantes Indépendantes. Cette méthode, proposée initialement par Hérault et Juten, repose sur une estimation récurrente des facteurs indépendants à des ordres élevés en employant des techniques de réseaux neuronaux.

Dans la seconde section, nous introduirons et décrirons l'analyse en composantes indépendantes. Cette méthode de décomposition sera appliquée à l'étude du Modèle d'Evaluation par Arbitrage sur le marché suisse réduit au marché des actions; la troisième section décrira les données utilisées ainsi que les méthodes de détermination du nombre de facteurs. Dans la quatrième section, nous essaierons de déterminer le nombre de facteurs nécessaires, en employant pour cela une batterie de tests, dont nous discuterons la pertinence.

2 L'Analyse en Composantes Indépendantes

Soit le vecteur de variables aléatoires $\mathbf{x}(t)$ de dimension n observé aux temps t . On suppose que ce vecteur est une combinaison linéaire de signaux d'entrée $\mathbf{s}(t)$:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \varepsilon_t, \quad (3)$$

où \mathbf{A} est une matrice $n \times n$ inversible, et ε_t un terme de bruit dû à l'observation du signal. Dans la pratique, le terme ε_t est généralement omis car ce terme d'erreur est impossible à distinguer des signaux d'entrée. Dans ce qui suit, on supposera que les \mathbf{x} observés sont d'espérance nulle. Le principe de l'analyse factorielle est de transformer les séries de vecteurs observés $\mathbf{x}(t)$ en des séries de vecteurs $\mathbf{s}(t)$, dont tous les éléments sont statistiquement indépendants entre eux, et représentant les signaux recherchés $\mathbf{y}(t)$. Le but de l'analyse factorielle est donc de déterminer une transformation linéaire optimale:

$$\mathbf{y}(\mathbf{t}) = \mathbf{W}\mathbf{x}(\mathbf{t}) \quad (4)$$

pour un ensemble de dimension n de vecteurs observés $\mathbf{x}(\mathbf{t})$ et une dimension particulière m du vecteur $\mathbf{s}(\mathbf{t})$, avec $m < n$. Elle consiste ainsi à projeter l'espace d'origine à n dimensions, dans un sous-espace à m dimensions, en cherchant à retenir la plus grande partie de l'information (dispersion) contenue dans les données. Deux indéterminations sont cependant inhérentes au problème: il est d'une part impossible de déterminer l'ordre des sources, et d'autre part de déterminer leur amplitude. Les signaux séparés $\mathbf{s}(t)$ sont ainsi une transformation des signaux sources $\mathbf{y}(t)$:

$$\mathbf{s}(t) = \mathbf{D}\mathbf{P}\mathbf{y}(t), \quad (5)$$

où \mathbf{D} est une matrice diagonale et \mathbf{P} une matrice de permutation.

L'analyse en composantes principales retient les m premières composantes principales du processus stochastique réalisé en \mathbf{x} pour déterminer la transformation linéaire $\mathbf{y} = \mathbf{W}\mathbf{x}$, avec $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]' \in R^{m \times n}$, et où $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ sont les m les vecteurs propres orthonormés de la matrice de variance-covariance estimée par $\hat{\Omega} = \frac{1}{T}\mathbf{X}\mathbf{X}'$. La recherche de la transformation en composante principales est adaptée aux réseaux neuronaux par Oja (1982) [27], qui propose un réseau constitué d'une seule unité de traitement linéaire, permettant de déterminer la première composante. Oja (1982) [27], et Oja et Karhunen (1985) [28], proposent

un réseau à une couche comprenant m unités de traitement linéaires permettant d'estimer les m premières composantes principales (figure 1). Dans ce modèle, la minimisation de la fonction de coûts définie par :

$$E(\mathbf{W}) = \frac{1}{2} \|\mathbf{e}\|^2, \quad (6)$$

avec $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$, où $\hat{\mathbf{x}} = \mathbf{W}'\mathbf{y}$ et $\mathbf{y} = \mathbf{W}\mathbf{x}$, conduit à la règle d'apprentissage du réseau:

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \eta(k)[\mathbf{x}'(k) - \mathbf{y}'(k)\mathbf{W}(k)]. \quad (7)$$

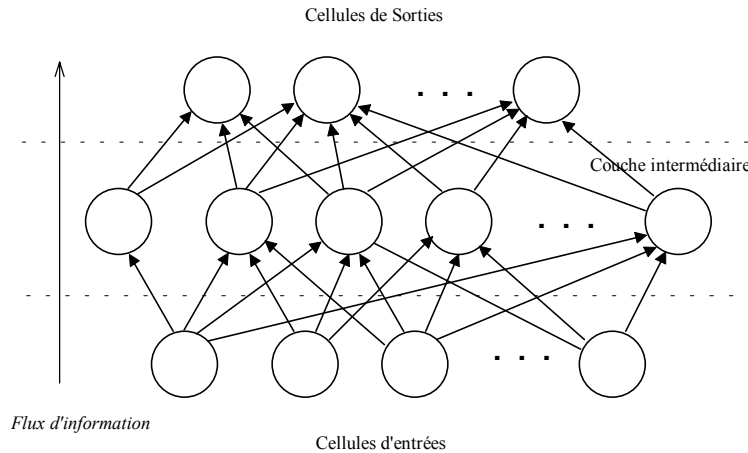


FIG. 1 – Réseau à rétropropagation

Karhunen et Oja démontrent que l'algorithme converge effectivement, mais vers une combinaison linéaire des m principaux vecteurs propres de Ω , indétermination liée à la technique même d'estimation des composantes principales.

L'algorithme d'analyse en composantes principales par les réseaux neuronaux n'apporte cependant pas d'amélioration tangible aux algorithmes classiques, puisque les deux types d'algorithmes mènent aux mêmes résultats. Herault et Jutten (1991) [18] proposent une approche alternative de l'analyse en composantes principales, qu'ils nomment analyse en composantes indépendantes (InCA, *INdependent Component Analysis*). Contrairement à l'analyse en composantes principales, le principe de l'analyse en composantes indépendantes n'est pas de déterminer les sources permettant de représenter la plus grande partie de la dispersion des signaux de sortie, mais de construire les sources les plus indépendantes possibles. Pour cela, il est nécessaire d'inclure dans la méthode de décomposition les statistiques d'ordre élevé, et la mesure d'indépendance statistique:

$$E[s_i(t)s_j(t)] = 0, \quad \forall i, j. \quad (8)$$

est remplacée par la condition que pour toutes les paires de signaux (i,j) (en l'occurrence $i = 1$ et $j = 2$), et tous les ordres k et l :

$$E[s_i^{2k+1}(t) s_j^{2l+1}(t)] = 0, \quad \forall i,j, \text{ et } \forall k,l. \quad (9)$$

L'algorithme initial de décomposition de Herault-Jutten ne traite que le cas de deux sources, et ne peut être étendu trivialement à plus de sources. De nombreuses extensions de l'algorithme initial de Herault-Jutten ont donc été récemment développés, sous le nom d'analyse en composantes principales non-linéaires, par Oja et Karhunen (1995) [29], d'analyse par maximisation de l'entropie, par Bell et Sejnowski (1996) [4], d'algorithme d'adaptation robuste, par Cichocki, Unbehauen, Moszczyński et Rummert (1994) [9], d'algorithme d'adaptation équivariante, par Cardoso et Laheld (1996) [6], ou d'algorithme par le gradient usuel, par Amari, Cichocki, et Yang (1995) [2]. Nous emploierons ici la dérivation de l'algorithme par Amari, Cichocki et Yang, qui repose sur la mesure de divergence de *Kullback-Leiber* et les développements de Gram-Charlier pour les variables stochastiques, et dont nous ne donnerons ici que les grandes lignes.

La dépendance entre les signaux reconstitués $\mathbf{y} = \mathbf{W}\mathbf{x}$, y_i , $i = 1..K$, sera ici mesurée par la mesure de divergence de *Kullback-Leiber* entre les distributions jointes et le produit des distributions marginales, qui s'écrit:

$$D(\mathbf{W}) = \int \frac{p(\mathbf{y}) \ln p(\mathbf{y})}{\prod_{i=1}^K p_i(y_i) \ln p_i(y_i)} d\mathbf{y} = H(\mathbf{y}) - \sum_{i=1}^K H(y_i) \quad (10)$$

avec:

$$\begin{aligned} H(\mathbf{y}) &= \int p(\mathbf{y}) \ln p(\mathbf{y}) d\mathbf{y}, \\ H(y_i) &= \int p_i(y_i) \ln p_i(y_i) dy_i. \end{aligned} \quad (11)$$

Les distributions marginales $p_i(y_i)$ peuvent être approximées par le développement de Gram-Charlier à l'ordre 4 :

$$p_i(y_i) \simeq \alpha(y_i) \left[1 + \frac{\kappa_3^i}{3!} H_3(y_i) + \frac{\kappa_4^i}{4!} H_4(y_i) \right] \quad (12)$$

où:

- $\kappa_3^i = m_3^i$, $\kappa_4^i = m_4^i - 3$, m_k^i sont les moments à l'ordre k de y_i ;
- le moment d'ordre 1 (espérance) est nul, et le moment d'ordre 2 (variance) est fixé à un: $m_1^i \equiv 0$, $m_2^i \equiv 1 \forall i$;

- $\alpha(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ est la densité en y de la loi normale standard;
- $H_k(y)$ sont les polynômes de Hermite-Chebyshev à l'ordre k , définis par: $\alpha(y)H_k(y) = (-1)^k \frac{d^k \alpha(y)}{dy^k}$.

En utilisant ce développement de Gram-Charlier et le développement de Lagrange du logarithme, et tenant compte des propriétés orthogonales des polynômes de Hermite-Chebyshev, la mesure de divergence suivant Kullback-Leiber s'écrit en fonction de \mathbf{W} et de \mathbf{x} comme:

$$-D(\mathbf{W}) \simeq -H(\mathbf{x}) - \ln |\det(\mathbf{W})| + \frac{K}{2} \ln(2\pi e) - \sum_{i=1}^K \left[\frac{(\kappa_3^i)^2}{2 \cdot 3!} + \frac{(\kappa_4^i)^2}{2 \cdot 4!} - \frac{5}{8} (\kappa_3^i)^2 \kappa_4^i - \frac{1}{16} (\kappa_4^i)^3 \right]. \quad (13)$$

Le but de l'algorithme est alors de déterminer les sources indépendantes en minimisant la valeur de $D(\mathbf{W})$, en calculant $\frac{\partial w_{ij}}{\partial t} = -\eta(t) \frac{\partial D}{\partial w_{ij}}$. Le développement de cette minimisation mène à la règle d'apprentissage:

$$\frac{d\mathbf{W}}{dt} = \eta(t) [\mathbf{I} - f(\mathbf{y})\mathbf{y}'] \mathbf{W}'^{-1}. \quad (14)$$

où la fonction d'activation $f(y)$ est définie par:

$$f(y) = \frac{29}{4}y^3 - \frac{47}{4}y^5 - \frac{14}{3}y^7 + \frac{25}{4}y^9 + \frac{3}{4}y^{11} \quad (15)$$

D'autre part, si on suppose que la matrice \mathbf{A} n'est pas singulière comme cela a été fait jusqu'ici, l'algorithme de descente du gradient $\frac{d\mathbf{W}}{dt} = -\eta(t) \frac{\partial D}{\partial \mathbf{W}}$ peut être remplacé par: $\frac{d\mathbf{W}}{dt} = -\eta(t) \frac{\partial D}{\partial \mathbf{W}} \mathbf{W}' \mathbf{W}$, et l'algorithme de descente est alors donné par:

$$\frac{d\mathbf{W}}{dt} = \eta(t) [\mathbf{I} - f(\mathbf{y})\mathbf{y}'] \mathbf{W}. \quad (16)$$

Common (1994) [10] discute en détail le contexte mathématique de la décomposition. L'étude de la convergence et de la stabilité de l'algorithme dans le cas de $m > 2$ sources est étudié par Cardoso et Laheld (1996) [6] qui considèrent des signaux avec de forts composants asymétriques, et par Amari, Chen et Cichocki (1997) [1]. En notant

$$\begin{cases} \sigma_i^2 = E[y_i^2], \\ k_i^2 = E[\dot{f}(y_i)], \\ m_i = E[y_i^2 \dot{f}(y_i)], \end{cases} \quad (17)$$

pour les différents signaux observés y_i , avec $\dot{f}(y) = \frac{\partial f(y)}{\partial y}$, ceux-ci montrent que l'algorithme est stable, et converge vers une solution unique, si les conditions suivantes sont satisfaites:

$$\begin{cases} m_i + 1 > 0, \\ k_i > 0, \\ \sigma_i \sigma_j k_i k_j > 1, \end{cases} \quad (18)$$

pour tous les signaux y_i, y_j .

Nous proposons d'employer l'analyse en composantes indépendantes pour le problème d'estimation des facteurs dans le cadre du Modèle d'Evaluation par Arbitrage, ce qui se justifie pour plusieurs raisons fondamentales.

Tout d'abord, l'un des défauts de l'analyse en composantes principales est, comme son nom l'indique, de rechercher non pas tant les composantes indépendantes que les composantes principales. De ce fait, la méthode va rechercher les composantes ayant les plus fortes influences sur le marché en se limitant aux moments d'ordre deux pour ce calcul. Comme le soulignent Conway et Reinganum (1988) [12], la tendance des méthodes d'analyse factorielle est d'extraire un facteur principal unique. C'est effectivement ce qui est trouvé par un grand nombre d'études du MEA, qui concluent à l'existence d'un facteur prépondérant sur le marché, par exemple par Trzcinka (1986) [34] qui trouve qu'un seul facteur présente une stabilité. Il n'existe pas d'alternative réelle à ces méthodes, à l'exception notable des méthodes semi-autorégressive et autorégressive, qui restent peu employées du fait de leur lourdeur et des capacités de calcul qu'elles nécessitent. Contrairement à l'analyse factorielle, l'analyse en composantes indépendantes ne cherche pas à déterminer une succession de composantes représentant la plus grande dispersion du marché, mais les composantes statistiquement les plus indépendantes entre elles. Cette décomposition semble donc beaucoup mieux adaptée au Modèle d'Evaluation par Arbitrage.

L'étude de l'analyse en composantes indépendantes en comparaison avec l'analyse en composantes principales, menée dans Vessereau (1999) [35] permet d'établir empiriquement que la structure factorielle obtenue est une structure factorielle stable. En divisant l'échantillon en sous-groupes d'actifs, la structure trouvée reste inchangée. De même, elle reste inchangée si les observations sont divisées en deux sous-périodes, pour autant que des moments d'ordre supérieurs à deux existent. Cette stabilité de la structure factorielle obtenue est fondamentale en ce qui concerne l'étude du Modèle d'Evaluation par Arbitrage. En particulier, une instabilité de la structure factorielle interdit d'envisager le test de ce modèle

dans des sous-groupes d'actifs en tant que modèle d'équilibre, alors que c'est le principal avantage du MEA, comme le soulignent Grinblatt et Titman (1983) [16]. L'instabilité de la structure factorielle obtenue par des méthodes d'analyse factorielle est ainsi une des critiques fondamentales qui peut être apportée à l'emploi de ces méthodes, ainsi que le montre en particulier l'étude de Chen (1983) [8].

L'avantage apporté par l'analyse en composantes indépendantes est fortement lié à l'existence de moments d'ordre supérieur à deux dans les distributions des facteurs et par conséquent des actifs, et en particulier à la non normalité et à l'asymétrie des rendements observés. Les tests des rendements sur les marchés rejettent clairement l'hypothèse de normalité. En particulier l'examen de l'asymétrie des lois montre que celles-ci sont significatives. Ce seul fait justifierait à lui seul l'emploi de la méthode proposée et le rejet des méthodes d'analyse factorielle qui ne prennent en compte que les covariances entre les actifs, à l'exclusion des statistiques d'ordre plus élevées. Comme le montre Sinclair (1981) [33], cette non normalité des rendements explique la faiblesse des résultats généralement obtenus par le MEA. Le premier choix généralement fait pour pallier ce problème est de regrouper les actifs en portefeuilles, mais avec comme conséquence de perdre les informations sur les covariances des actifs à l'intérieur de ces portefeuilles, et donc de risquer de sous-estimer le nombre de facteurs nécessaires. Le deuxième choix est d'utiliser des rendements mensuels ou hebdomadaires, qui présentent des distributions plus proches de la normalité, ce qui est fait par exemple par Trzcinka (1986) [34] ou Mei (1993) [24]; de la même manière, ceci conduit à perdre des informations sur les covariances des actifs sur des intervalles de temps plus court, et ne permet pas de cerner le comportement du modèle sous une fréquence quotidienne.

Le MEA ne fait théoriquement aucune hypothèse quant aux distributions des rendements des actifs. Cependant, les études empiriques du modèle utilisent généralement des méthodes d'analyse factorielle (analyse en facteurs communs ou analyse en composantes principales); comme ces méthodes ne tiennent généralement compte que des deuxièmes moments (variances et covariances), elles supposent implicitement que le risque peut être représenté dans un cadre moyenne-variance. Les méthodes d'analyse factorielle supposent d'autre part que la structure factorielle est stricte, ou s'en approche, puisque l'on impose que la matrice des termes résiduels soit diagonale. Ceci signifie qu'on recherche des covariances des erreurs entre elles et avec les facteurs exactement nulles, ce qui dans la pratique est irréaliste. La con-

vergence de la méthode n'est ainsi pas effectivement accomplie. L'analyse en composantes indépendantes par contre, par l'emploi de statistiques d'ordre plus élevé, impose l'existence de covariances de moment d'ordre deux faibles mais pas obligatoirement nulles. Son utilisation dans le cadre du Modèle d'Evaluation par Arbitrage permet ainsi d'introduire explicitement la possibilité d'une structure factorielle approximative, étudiée par Chamberlain et Rothschild (1983) [7]. Elles relâchent ainsi les restrictions sur la possibilité d'annulation complète du risque non systématique par les investisseurs, et autorisent le fait que des facteurs communs à un sous-ensemble d'actifs peuvent exister, en particulier sous forme de facteurs sectoriels.

3 Présentation des données

En termes de capitalisation boursière, le marché des actions suisses occupe le septième rang mondial. Il se caractérise par une grande étroitesse, puisque le nombre de compagnies émettrices est d'un peu plus d'une centaine, et le nombre d'actions cotées dans l'ensemble des bourses suisses, qu'il s'agisse d'actions au porteur, de nominatives, de bons de participation ou de jouissance, va de 200 à 450 selon les périodes. Il est également très concentré puisque les dix plus grandes sociétés cotées constituent plus de 70% de la capitalisation boursière. A peu près la moitié des actions suisses sont émises par des sociétés ayant des activités de production (alimentaire, machines, chimie), et l'autre moitié à des institutions bancaires, d'assurance ou d'autres services (transports, grands magasins). Un aspect important du marché suisse est que chaque société a la possibilité d'émettre des actions de diverses catégories: porteurs, nominatives, bons de participation ou de jouissance, donnant droit à des droits sociaux et patrimoniaux différents, les transactions d'actions nominatives pouvant également restreints selon les sociétés. L'existence de ces différents titres portant sur les mêmes société indique la possibilité de covariances importantes entre certains titres. Au cours de la période d'étude, des modifications sur les titres ont naturellement été effectuées: à partir de 1985, un certain nombre de bons de participation ont été convertis en actions en porteur ou en nominatives, et de nombreuses divisions de titres et simplifications de structure de capital ont été menées plus particulièrement en 1992 et 1993.

La base de données a été constituée en utilisant la connexion DATASTREAM de l'Université de Genève. Toutes les données disponibles du 1er janvier 1975 au 31 décembre 1997 ont été

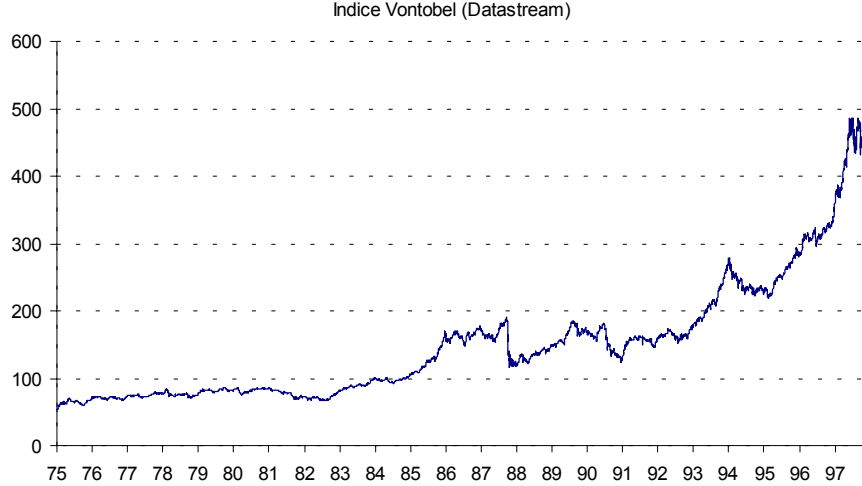


FIG. 2 – Cours de l'indice du 01.01.1975 au 31.12.1997

retenues, excluant les certificats d'investissement dans des sociétés étrangères, constituant une période d'étude de vingt-trois années. Pour éviter le biais de survie, les séries arrêtées au cours de cette période (référéncées par DATASTREAM comme des *dead stocks*) ont été également retenues, en tenant compte cependant du fait que certaines de ces séries avaient été purement et simplement dupliquées pour construire d'autres séries. Cette première étape a permis d'obtenir un premier ensemble de 465 titres, dont 177 titres *dead stocks*.

D'autre part, le taux journalier Euroswisse à trois mois entre le 1er janvier 1975 et le 31 décembre 1997 a été utilisé comme proxy du taux hors-risque sur le marché suisse. Les indices principaux du marché suisse, le SMI et le SPI n'ayant été créés que dans les années 80, et ne concernant qu'une partie des actifs disponibles, l'indice Vontobel de DATASTREAM, qui prend en compte l'ensemble des données disponibles sur le marché et est ajusté aux dividendes et aux changements de capital a été retenu comme proxy du portefeuille de marché. La figure 2 illustre le comportement du marché sur la période retenue: durant l'ensemble de la période, le rendement moyen annuel de l'indice a été de 13% et la valeur de l'indice a été multipliée par six; ce rendement élevé est surtout dû à l'évolution du marché à partir de l'année 1993.

Les données de DATASTREAM sont des prix journaliers de clôture, ajustés aux dividendes et aux opérations en capital. A partir de ces prix, les rendements logarithmiques d'un jour $t - 1$ au jour t des titre ont été calculés par:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right), \quad (19)$$

où P_t et P_{t-1} sont les prix du titre aux temps t et $t - 1$ respectivement.

La période totale allant du 1er janvier 1975 au 31 décembre 1997 a été divisée en 19 sous-périodes de cinq ans chacune. La première période est constituée des données allant du 1er janvier 1975 au 31 décembre 1979, puis chaque période successive est générée en enlevant la première année de la période précédente et en rajoutant l'année suivante à la fin de la période. Pour chaque période, les quatre premières années serviront à l'estimation des paramètres du modèle, et la cinquième année au test de la modélisation. Les rendements ont été calculées pour chacune de ces périodes suivant trois fréquences: fréquences quotidiennes, hebdomadaires et mensuelles. Les fréquences hebdomadaires ont été calculées en prenant les prix de cloture du mercredi, afin d'éviter les anomalies du vendredi au lundi généralement constatées sur les marchés. Les fréquences mensuelles ont été calculées en prenant les prix de cloture du dernier jour de chaque mois.

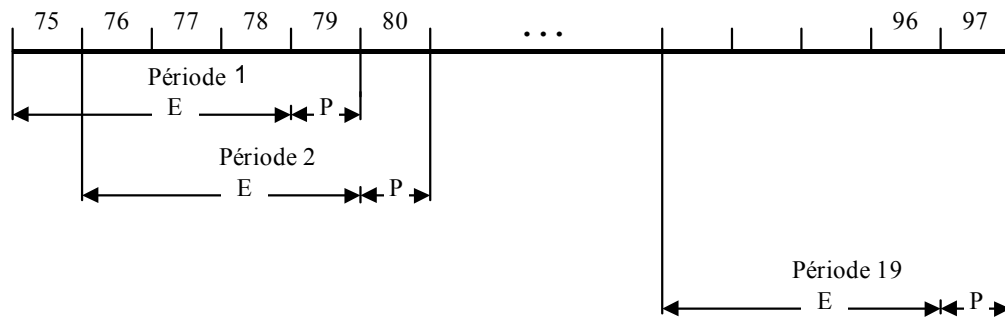


FIG. 3 – *Constitution des périodes d'études*

La première période est constituée des données allant du 1er janvier 1975 au 31 décembre 1979; les quatre premières années servent à l'estimation des paramètres du modèle (E), et la cinquième année à la projection (P). Chaque période successive est générée en enlevant la première année de la période précédente et en rajoutant l'année suivante à la fin de la période.

Le marché suisse souffrant de problèmes de liquidité sur beaucoup de titres, le nombre d'actions effectivement traitées est parfois inférieur à la centaine, et varie suivant les périodes d'étude. Pour déterminer les actions retenues sur chaque période, la procédure suivante a été suivie: d'une part, DATASTREAM fournissant des prix tous les jours de l'année, sans distinction des jours fériés propres à la Suisse, les jours pour lesquels les rendements de l'ensemble des actions sont strictement nuls ont été supprimés; ensuite, certains titres pouvant être émis après le début d'une période ou supprimés avant la fin de celle-ci, les titres dont plus de 10% des cotations n'existent pas dans la période ont été supprimés; enfin, l'on a restreint les titres d'une période aux titres qui étaient cotés en moyenne au moins

deux fois par semaine, c'est-à-dire dont au moins au moins 30% des rendements étaient non nuls. Inclure un titre qui ne serait coté qu'une fois par mois dans une base de données quotidiennes n'aurait en effet pas de sens; d'autre part, il faut noter que les titres ainsi enlevés ne représentent qu'une fraction faible de la capitalisation totale, le manque de liquidité étant dû à un nombre trop faible de titres disponibles et concerne les plus petites sociétés.

		Jour			Semaine		Mois	
Début	Fin	# titres	# sociétés	% marché	# titres	% marché	# titres	% marché
01.01.75	31.12.79	46	34	94.5	71	98.3	71	98.3
01.01.76	31.12.80	51	37	93.5	79	98.0	90	98.7
01.01.77	31.12.81	50	36	92.0	81	97.1	93	98.0
01.01.78	31.12.82	50	36	94.1	86	96.7	93	97.0
01.01.79	31.12.83	53	37	93.6	95	97.3	102	97.6
01.01.80	31.12.84	55	38	94.5	97	96.9	101	97.0
01.01.81	31.12.85	67	47	95.7	107	97.1	109	97.1
01.01.82	31.12.86	73	51	96.8	118	98.2	117	98.2
01.01.83	31.12.87	93	63	94.2	134	98.3	126	98.3
01.01.84	31.12.88	105	71	95.3	139	98.4	134	98.4
01.01.85	31.12.89	121	76	95.3	151	97.5	146	97.5
01.01.86	31.12.90	131	80	92.9	171	95.3	162	94.7
01.01.87	31.12.91	140	89	92.1	193	94.3	190	94.8
01.01.88	31.12.92	125	82	94.3	186	98.9	196	98.9
01.01.89	31.12.93	116	81	95.4	171	99.1	173	99.1
01.01.90	31.12.94	102	76	93.7	158	98.2	165	98.2
01.01.91	31.12.95	85	65	91.7	144	98.0	159	98.0
01.01.92	31.12.96	96	77	90.7	137	97.0	146	97.1
01.01.93	31.12.97	106	86	92.2	147	98.1	146	98.1
01.01.75	31.12.97	88	61	93.8	130	97.5	133	97.6

TAB. 1 – *Compositions des échantillons par sous-périodes*

Ce tableau décrit le type des titres retenus pour chaque sous-période de cinq ans. Sont reportés le nombre de titres retenus, le nombre de sociétés différentes représentés par ces titres, ainsi que la fraction de capitalisation boursière représentée. Le nombre de titres retenu par période est relativement faible: de 46 à 140 pour les données quotidiennes, de 71 à 190 pour les données hebdomadaires et mensuelles. Malgré ce faible nombre, les titres retenus représentent la plus grande partie de la capitalisation du marché, puisque quelle que soit la fréquence de données, plus de 90% de la capitalisation totale du marché sur chaque période est retenue.

Le tableau 1 présente, pour chaque période constituée, le nombre de titres retenus pour les données quotidiennes, hebdomadaires et mensuelles, et la fraction de la capitalisation du marché représenté par ces titres. Sur chaque période de cinq ans, les titres du marché retenus représentent près de 95 % en moyenne de la capitalisation totale du marché pour les données quotidiennes, et 97.6 % en moyenne pour les données hebdomadaires ou mensuelles. L'on peut ainsi estimer que la plus grande part du marché des actions est retenu à chaque période, et surtout la part du marché sur laquelle il y a effectivement des échanges. D'autre part, le

nombre de titres se rapportant à une même société est relativement important pour chaque période, puisqu'en moyenne un quart des titres se rapportent à une société pour laquelle il existe un autre titre dans la période. L'existence de différents titres sur la même société devrait mener à des covariances importantes entre les séries observées correspondantes, et donc entre les séries d'erreurs du modèle correspondantes; dans la spécification du modèle, ceci pourrait être interprété comme l'existence de facteurs sectoriels propres à ces titres. Il est à noter cependant qu'il est nécessaire que le modèle soit robuste à l'existence de tels facteurs sectoriels, et que l'existence de ce point de vue de titres sur la même société n'est pas rédhibitoire; il paraît donc que le seul inconvénient de cette existence est le fait que la taille réelle de l'échantillon est peut-être surestimé.

Au tableau 2 sont reportées les études statistiques menées sur les dix actions les plus importantes du marché suisse, ainsi que l'indice Vontobel, sur l'ensemble de la période de 20 ans. La plupart de ces actions ont un rendement annuel moyen élevé, entre 5% et 10%, mais avec des écarts-type importants (20 à 30%). L'ensemble des actions présente une asymétrie importante, et une kurtosis forte. Les tests de normalité des rendements, par le test de Jarque-Bera, et par celui de Kolomogorov-Smirnov montrent que l'hypothèse de normalité des rendements logarithmiques est rejetée pour tous les actifs, ce dont le modèle d'évaluation de ces actifs devrait donc tenir compte.

	BD	CN	FR	JP	SW	UK	US
Mean	0.106	0.091	0.097	-0.066	0.153	0.114	0.154
StDev	0.159	0.104	0.157	0.185	0.145	0.125	0.129
Skewness	-1.341	-0.990	-0.347	0.347	-1.070	0.004	-0.642
Kurtosis	15.233	10.443	4.304	5.915	10.805	3.108	7.666

TAB. 2 – *Etudes statistiques des données*

4 Le Modèle d'Evaluation par Arbitrage sur le marché suisse

Très peu d'études ont été menées sur le marché suisse des actions, sur lequel les tests seront menés. La seule étude de référence pour le Modèle d'Evaluation par Arbitrage est menée par Beckers, Cummins et Woods (1993) [3]. Encore, cette étude, comme la grande majorité des études traitant du MEA, ne procède pas vraiment à l'estimation des séries de facteurs, mais les détermine a priori comme étant des séries macro-économiques observées ou des caractéristiques associées aux actifs (taille, volatilité). Dans une étude non publiée,

Marañon (1994) [23] développe, adapte et applique la méthode autorégressive au marché suisse sur des rendements hebdomadaires en considérant un ensemble des actions les plus traitées sur ce marché. Il obtient qu’au moins quatre, et probablement cinq facteurs, sont nécessaires sur ce marché.

Nous emploierons ici l’analyse en composantes indépendantes pour déterminer les séries de facteurs. Les résultats de cette méthode seront comparés à ceux de l’analyse en composantes principales; le choix de cette méthode plutôt que l’analyse en facteurs communs, outre qu’elle est la plus proche du fonctionnement de l’analyse en composantes indépendantes, est essentiellement due au fait qu’elle ne nécessite pas l’hypothèse que les facteurs et les résidus suivent des lois localement normales. D’autre part, Chamberlain et Rothschild (1983) [7] montrent qu’elle est asymptotiquement équivalente à l’analyse factorielle. Contrairement aux études faites sur le marché américain, les actifs seront ici traités individuellement, sans qu’ils soient rassemblés en portefeuilles. La constitution de ces portefeuilles est en effet essentiellement dû au nombre très important d’actifs existants sur le marché américain. Cette constitution de groupes permet également de réduire les covariances entre les actifs, comme le montre aussi Lee (1988) [21] qui propose d’adopter une approche par constitution de portefeuille afin d’obtenir une structure factorielle plus stable. Cependant, l’inconvénient de la méthode est de supprimer des covariances entre des actifs, avec le risque d’introduction de biais, comme cela est mis en évidence par Dhrymes et al. (1984) [14].

4.1 Détermination a priori du nombre de facteurs

Le nombre de facteurs nécessaires à retenir pour la structure factorielle de génération des rendements n’a pas encore reçu de solution totalement satisfaisante et applicable quelle que soit la méthode d’analyse utilisé. Si l’on suppose que les facteurs et les termes résiduels sont localement normaux, un test de maximum de vraisemblance peut être construit à partir des matrices de covariance. La statistique de test est:

$$J = -(T - 1 - \frac{1}{6}(2N + 5) - \frac{2}{3}K) [\log |\hat{\Omega}| - \log |\hat{\mathbf{B}}\hat{\mathbf{B}}' + \hat{\mathbf{D}}|], \quad (20)$$

où N , T , et K sont respectivement le nombre d’actifs, le nombre de vecteurs d’observations et le nombre de facteurs du modèle contraint, et où $\hat{\Omega}$, $\hat{\mathbf{B}}$ et $\hat{\mathbf{D}}$ sont des estimateurs des matrices de covariance des facteurs Ω , de sensibilités des actifs aux facteurs \mathbf{B} et de covariance des erreurs \mathbf{D} respectivement. Sous l’hypothèse nulle que K facteurs sont adéquats, J est

asymptotiquement distribuée suivant une chi-carré à $\frac{1}{2}[(N - K)^2 - N - K]$ degrés de liberté.

Dans leur études initiale du Modèle d'Evaluation par Arbitrage, Roll et Ross (1980) [30] utilisent cette approche statistique et concluent qu'un nombre de trois ou quatre facteurs est suffisant, avec cependant un seuil de signification pour le test de χ^2 fixé à 50%, et moins de la moitié des groupes d'actifs constitué pour leur étude satisfaisant le test. Les nombreuses études qui utilisent ce test pour déterminer a priori le nombre de facteurs, trouvent généralement soit un nombre restreint de groupes d'actifs pour lesquels les tests de χ^2 sont effectivement significatifs, soit un nombre de facteurs nécessaire très élevé qui peut aller jusqu'au quart du nombre d'actifs étudié.

Plusieurs critiques peuvent être faites à ce test couramment employé. D'une part, il nécessite l'hypothèse que les rendements conditionnels aux facteurs suivent une distribution jointe normale et sont identiquement distribués à travers le temps. D'autre part, l'hypothèse testée est que les rendements suivent une structure factorielle *stricte*, ce qui n'est pas une hypothèse nécessaire du point de vue théorique, comme nous l'avons vu dans le chapitre quatrième de ce travail. Un défaut connu de ce test discuté par Morrison (1990) [26] est ainsi de tendre à surestimer le nombre de facteurs nécessaires, en particulier pour les petits échantillons. Comme cela est montré sur des marchés simulés par Vessereau (1999), le test est d'autre part très sensible à l'existence de facteurs sectoriels. La présence de tels facteurs, et donc de covariances, biaise nettement le test du χ^2 et tend donc à surestimer le nombre de facteurs nécessaires.

Les résultats du test de χ^2 sur les données du marché suisse sont présentés au tableau 3. Les tests trouvent un nombre de facteurs moyens de 9 pour les données quotidiennes et hebdomadaires sur l'ensemble de la période et un nombre de 7 pour les données mensuelles. Sur des périodes de cinq années, ce nombre est très variable; pour les périodes entre 1975 et 1986, le nombre obtenu est de 5 à 7, et de 12 à 15 pour les périodes allant de 1983 à 1992. Cette augmentation du nombre de facteurs peut être dû à un changement de structure entre ces deux périodes, puisque le nombre de facteurs augmente sensiblement pour les périodes incluant les chutes du marché de 1987 et 1989. Mais cette augmentation peut également être plus simplement dû à une augmentation du nombre d'actifs traités durant ces périodes, comme le montre le graphique de la table 3. Cette augmentation du nombre de facteurs suivant le nombre d'actifs est cohérent avec les résultats par exemple de l'étude de Dhrymes

et al. (1984) [14] sur le marché américain. Dans ce sens, ce test, souvent utilisé dans les études du MEA, semble inadaptée à la détermination du nombre de facteurs nécessaires, car il tend à surestimer ce nombre de facteurs; cependant, le nombre de facteurs nécessaires trouvé, même surestimé, est suffisamment faible pour que le modèle multifactoriel ait un sens.

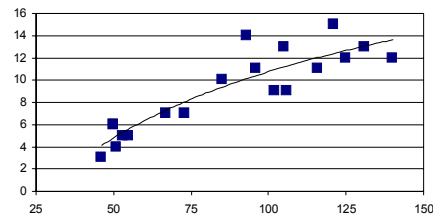
Nombre de facteurs obtenus par périodes d'étude (fréquence quotidienne)

Périodes	# Titres	# Facteurs	Périodes	# Titres	# Facteurs
75/79	46	3	85/89	121	15
76/80	51	4	86/90	131	13
77/81	50	6	87/91	140	12
78/82	50	6	88/92	125	12
79/83	53	5	89/93	116	11
80/84	55	5	90/94	102	9
81/85	67	7	91/95	85	10
82/86	73	7	92/96	96	11
83/87	93	14	93/97	106	9
84/88	105	13	Moyenne	88	9

Moyennes selon les fréquences des données

Fréquences	# Titres	# Facteurs
Quotidiennes	88	9
Hebdomadaires	130	9
Mensuelles	133	7

Nombre de facteurs par nombre d'actifs



TAB. 3 – Détermination a priori du nombre de facteurs sur le marché suisse

Les nombres de facteurs donnés pour les données quotidiennes par le test de χ^2 sur les matrices de covariance sont reportés pour chaque période d'étude de quatre ans, la période commençant le 1er janvier de l'année spécifiée. Le nombre de facteurs obtenus est en moyenne de neuf pour les données quotidiennes et hebdomadaires, et seulement sept pour les données mensuelles, mais ce nombre est très variable selon les périodes d'étude. Ceci semble dû en particulier à la variation des nombres d'actifs retenus suivant les sous-périodes, ce qui est illustré par le graphique, où le nombre de facteurs obtenu est exprimé en fonction du nombre d'actifs dans la période.

4.2 Tests de significativité des primes de risque

Les méthodes employées pour le calcul des primes liées aux facteurs reprennent l'approche utilisée par Fama et McBeth (1973) [15] dans le cadre du MEDAF, formalisée dans le cadre du MEA par Roll et Ross (1980) [30] et Shanken (1992) [32]. Suivant cette approche, les rendements des actions sont régressés transversalement sur leurs sensibilités aux différents

facteurs betas, pour chaque période t :

$$\tilde{\mathbf{R}}_t = \iota\lambda_{0t} + \hat{\mathbf{B}}\lambda_t + \tilde{\eta}_t, \quad (21)$$

où $\tilde{\mathbf{R}}_t$ est le vecteur des rendements des actifs au temps t , $\hat{\mathbf{B}}$ la matrice estimée des sensibilités des actifs aux facteurs, et $\tilde{\eta}_t$ le terme d'erreur de la régression. Les séries temporelles des $\hat{\lambda}_{kt}$ permettent alors d'obtenir une estimation des valeurs moyennes et des écarts-type des primes $\hat{\lambda}_k$: $\hat{\lambda}_k = \frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{kt}$. L'hypothèse que cette prime est non nulle peut alors être testé par un test de Student classique sur la moyenne. Le plus souvent, l'estimation de Λ des primes temporelles ($K \times T$) des facteurs est effectuée par la méthode des moindres carrés généralisés, comme dans l'étude initiale du MEA menée par Roll et Ross (1980) [30]:

$$\Lambda = (\hat{\mathbf{B}}'\hat{\Omega}^{-1}\hat{\mathbf{B}})^{-1}\hat{\mathbf{B}}'\hat{\Omega}\mathbf{R}, \quad (22)$$

où la matrice \mathbf{R} est la matrice ($N \times T$) des réalisations transversales des rendements, et $\hat{\Omega}$ la matrice de variance-covariance des rendements dans le modèle contraint. Avec les hypothèses du MEA, $\hat{\Omega} = \hat{\mathbf{B}}\hat{\mathbf{B}}' + \hat{\mathbf{D}}$, où $\hat{\mathbf{D}}$ est une matrice diagonale dont les éléments sur la diagonale sont les variances des termes résiduels obtenus, les covariances entre les termes résiduels étant supposé nuls.

La méthode la plus employée pour déterminer le nombre de facteurs nécessaires pour la détermination du Modèle d'Evaluation par Arbitrage, consiste donc à tester la significativité des primes estimées à l'étape précédente. Si l'estimation d'une prime de risque d'un facteur ne donne pas une valeur significativement différente de zéro, l'on peut supposer qu'elle correspond à un facteur non nécessaire au modèle. Cependant, comme le remarque déjà Roll et Ross (1980) [30], cette méthode nous protège contre le fait d'obtenir trop de facteurs estimés à la première étape, mais pas contre trop peu. En fait, cette méthode ne protège ni contre le fait de prendre trop de facteurs, ni contre le fait d'en prendre trop peu, comme le montre Vessereau (1999) sur des marchés simulés. En effet, le fait qu'une prime ne soit pas significative n'implique pas nécessairement que ce facteur ne soit pas utile à la décomposition, puisqu'il peut-être rémunéré ou non suivant les périodes. Il est donc nécessaire de tester la significativité des primes sur un ensemble de sous-périodes successives. D'autre part, la rotation des facteurs observé pour l'analyse factorielle classique empêche de donner une quelconque importance à la significativité des primes, puisque cela dépend également de la rotation qui est appliquée aux facteurs réels.

Analyse en Composantes Principales

	CAPM	1	2	3	4	5	6
75/79	2.93 <i>0.19</i>	0.19 <i>0.01</i>	4.51 <i>0.32</i>	-4.64 <i>-0.33</i>	-3.69 <i>-0.26</i>	-4.32 <i>-0.34</i>	-1.44 <i>-0.10</i>
80/84	21.93 <i>1.17</i>	6.44 <i>0.42</i>	-12.36 <i>-1.82</i>	-9.44 <i>-0.53</i>	2.63 <i>0.18</i>	2.59 <i>0.18</i>	-1.83 <i>-0.17</i>
85/89	29.37 <i>2.15</i>	21.53 <i>1.86</i>	-10.79 <i>-1.34</i>	-9.61 <i>-1.19</i>	5.26 <i>0.60</i>	5.25 <i>0.60</i>	-4.51 <i>-0.45</i>
90/94	-0.86 <i>-0.06</i>	-5.19 <i>-0.35</i>	0.55 <i>0.04</i>	0.92 <i>0.06</i>	0.48 <i>0.03</i>	3.55 <i>0.20</i>	3.38 <i>0.24</i>
93/97	24.91 <i>1.52</i>	24.87 <i>1.58</i>	-6.84 <i>-0.44</i>	-6.91 <i>-0.44</i>	2.97 <i>0.30</i>	2.99 <i>0.30</i>	2.89 <i>0.29</i>
75/97	13.68 <i>0.79</i>	11.62 <i>0.71</i>	7.64 <i>0.91</i>	6.72 <i>0.55</i>	2.84 <i>0.28</i>	3.59 <i>0.32</i>	3.15 <i>0.28</i>

Analyse en Composantes Indépendantes

	CAPM	1	2	3	4	5	6
75/79	2.93 <i>0.19</i>	—	-1.89 <i>-0.16</i>	3.75 <i>0.25</i>	2.23 <i>0.14</i>	1.97 <i>0.15</i>	1.10 <i>0.08</i>
80/84	21.93 <i>1.17</i>	—	9.02 <i>0.69</i>	5.52 <i>0.41</i>	-1.12 <i>-0.07</i>	2.64 <i>0.17</i>	0.03 <i>0.00</i>
85/89	29.37 <i>2.15</i>	—	-22.54 <i>-2.31</i>	-8.89 <i>-0.88</i>	-0.16 <i>-0.01</i>	-0.59 <i>-0.06</i>	0.16 <i>0.01</i>
90/94	-0.86 <i>-0.06</i>	—	13.95 <i>0.90</i>	5.32 <i>0.36</i>	5.17 <i>0.37</i>	-3.12 <i>-0.20</i>	3.53 <i>0.23</i>
93/97	24.91 <i>1.52</i>	—	8.25 <i>0.88</i>	9.96 <i>0.54</i>	9.74 <i>0.60</i>	0.79 <i>0.05</i>	6.82 <i>0.43</i>
75/97	13.68 <i>0.79</i>	—	14.21 <i>1.20</i>	6.11 <i>0.46</i>	1.84 <i>0.12</i>	3.21 <i>0.17</i>	1.85 <i>0.10</i>

TAB. 4 – *Primes de risque les plus faibles obtenues sur le marché suisse*

Le tableau rapporte les primes les plus faibles, annualisées, et les valeurs du test de Student, obtenues en utilisant respectivement l'analyse en composantes principales et l'analyse en composantes indépendantes pour les données quotidiennes. Pour chaque période, la structure factorielle est estimée sur les quatre premières années, et les primes de risque calculées sur la cinquième année. Sur la totalité de la période de 1975 à 1997, aucun facteur ne serait significatif; la prime du MEDAF elle-même n'est pas significativement différente de zéro. L'analyse par composantes principales ne permet de trouver une prime significative sur aucune des périodes. Par contre, l'emploi de l'analyse par composantes indépendantes fait apparaître un deuxième facteur pour la période allant de 1985 à 1990. La présence de ce deuxième facteur, même sur une sous-période uniquement, justifie a priori l'étude d'un modèle multifactoriel d'évaluation sur le marché suisse.

Le tableau 4 reporte les primes plus faibles obtenues sur le marché suisse en utilisant l'analyse en composantes principales et l'analyse en composantes indépendantes, pour un nombre de facteurs de un à six. La structure factorielle est estimée sur une période de quatre ans, et les primes de risque correspondantes sont calculées sur la cinquième année. Les commencements des périodes d'étude ont été prises par intervalles de cinq ans. Une période totale d'étude allant de 1975 à 1997 a également été utilisée, et la prime de risque du marché calculée en utilisant le Modèle d'Evaluation des Actifs Financiers (MEDAF) a été ajoutée aux tableaux pour comparaison. Sur ces tableaux, l'on constate que le nombre de facteurs donné par les deux types de décomposition serait inférieur à deux si l'on prend l'ensemble de la période. Cependant, il existe une sous-période, allant de 1985 à 1989 pour laquelle un deuxième facteur est rémunéré, lorsque la méthode de l'analyse en composantes indépendantes est utilisée. L'existence de cette rémunération d'un deuxième facteur, même limité dans le temps, est une forte indication que le modèle multifactoriel a un sens sur le marché suisse.

4.3 Test du nombre de facteurs par le modèle de Wei

Les deux tests que nous avons fait jusqu'ici ont surtout indiqué qu'ils n'étaient guère adaptés au Modèle d'Evaluation par Arbitrage. Le seul test pertinent pour le MEA semble être de tester les primes résiduelles du marché, non pris en compte par le modèle; ce test est le seul test possible si une succession de sous-périodes est utilisée, puisque les facteurs estimés ne peuvent pas être identifiés d'une période à l'autre du fait de l'existence d'une probable rotation ou simple permutation de ces facteurs. Nous suivrons ici la même approche que Lehmann et Modest (1988) [22] ou Connor et Korajczyk (1988) [12], en testant le modèle pour des valeurs différentes de K et en observant si les résultats des tests sont sensibles à l'augmentation du nombre de facteurs.

Le premier test utilisé pour déterminer le nombre de facteurs est le test de la significativité de la prime résiduelle du marché. Comme nous l'avons vu, Wei (1988) [36] écrit l'équation fondamentale du MEA en incluant dans sa formulation la sensibilité des actifs à un portefeuille constitué des termes résiduels du portefeuille de marché:

$$E[\tilde{r}_i] = \lambda_0 + b_{im}E[\tilde{r}_m] + \sum_{k=1}^K b_{ik}f_k, \quad (23)$$

où $E[\tilde{r}_m]$ est le rendement du portefeuille de marché résiduel. Avec cette formulation, le test des primes d'évaluation des facteurs s'écrit:

$$E[\tilde{R}_i] = \lambda_0 + \lambda_m b_{im} + \sum_{k=1}^K \lambda_k b_{ik}. \quad (24)$$

La prime de risque associée au portefeuille de marché résiduel rémunère le risque systématique des actifs qui ne sont pas pris en compte par les K facteurs de la base factorielle estimée. Elle est donc positive et une fonction croissante du nombre de facteurs omis par le modèle multifactoriel. Lorsque des facteurs sont omis, et qu'un nombre $J < K$ de facteurs est retenu, K étant le nombre correct de facteurs nécessaires, la prime de risque associée au portefeuille de marché résiduel est positive non nulle. Il suffit donc a priori de tester cette prime de risque pour déterminer le nombre de facteurs nécessaire au modèle.

La procédure suivie pour tester la significativité des primes suit la procédure développée par Fama et McBeth (1973) [15]: les vecteurs des rendements \mathbf{r}_t des actifs $i = 1..N$ au temps t sont régressés transversalement sur les coefficients de sensibilité des actifs aux facteurs \mathbf{b} et au portefeuille résiduel de marché b_{mt} par:

$$\mathbf{r}_t = \lambda_{0t} + \lambda_t \mathbf{b} + \lambda_{mt} b_{mt} + \epsilon_{it}. \quad (25)$$

Pour chacune des 19 sous-périodes, les quatre premières années sont utilisées pour estimer la structure factorielle, c'est-à-dire les coefficients de sensibilité \mathbf{b} des actifs à ces facteurs et b_{mt} au portefeuille résiduel de marché, et les primes λ_{mt} sont calculées sur la cinquième année de la période. Une fois estimées les séries des primes λ_{mt} associées au portefeuille résiduel de marché, la valeur de cette prime est calculée par

$$\hat{\lambda}_m = \frac{1}{T} \sum_{t=1}^T \hat{\lambda}_{mt}, \quad (26)$$

et sa significativité testée en utilisant le test de Student sur la moyenne. Ce test suppose qu'il n'existe pas d'autocorrélation dans la série λ_{mt} . Comme le discutent Fama et MacBeth (1973) [15], la présence d'autocorrélation tendrait à surévaluer la valeur de t-stat.

Le tableau 5 présente les résultats globaux obtenus par l'analyse en composantes indépendantes et par l'analyse en composantes principales pour un nombre de facteurs compris entre un et sept. Les données ont été utilisées selon une fréquence quotidienne, hebdomadaire et mensuelle. Pour l'analyse en composantes indépendantes, les résultats obtenus avec cette

Analyse en Composantes Principales

# facteurs	1	2	3	4	5	6	7
Quotidiennes	13.12	12.28	10.34	8.33	8.34	8.22	5.94
t-stat	2.36	2.49	2.05	1.62	1.58	1.57	1.11
<i>p-value</i>	0.02	0.01	0.04	<i>0.11</i>	<i>0.11</i>	<i>0.12</i>	<i>0.27</i>
Hebdomadaires	18.49	18.02	14.39	18.59	24.35	21.14	14.99
t-stat	2.26	2.12	1.71	2.07	2.61	2.27	1.60
<i>p-value</i>	0.03	0.03	<i>0.09</i>	<i>0.04</i>	<i>0.01</i>	<i>0.02</i>	<i>0.11</i>
Mensuelles	40.21	15.86	3.18	5.82	2.28	5.81	7.05
t-stat	3.02	1.21	0.21	0.34	0.13	0.37	0.44
<i>p-value</i>	0.00	<i>0.23</i>	<i>0.84</i>	<i>0.74</i>	<i>0.89</i>	<i>0.71</i>	<i>0.66</i>

Analyse en Composantes Indépendantes

# facteurs	1	2	3	4	5	6	7
Quotidiennes	—	9.98	8.16	6.12	3.25	1.03	2.20
t-stat	—	1.84	1.35	0.99	0.52	0.16	0.33
<i>p-value</i>	—	<i>0.07</i>	<i>0.18</i>	<i>0.32</i>	<i>0.61</i>	<i>0.87</i>	<i>0.74</i>
Hebdomadaires	—	14.58	15.17	7.90	7.36	4.03	1.43
t-stat	—	1.52	1.41	0.73	0.69	0.37	0.14
<i>p-value</i>	—	<i>0.13</i>	<i>0.16</i>	<i>0.47</i>	<i>0.49</i>	<i>0.71</i>	<i>0.89</i>
Mensuelles	—	29.86	23.69	37.17	27.93	37.03	30.76
t-stat	—	2.20	1.66	2.64	1.97	2.67	2.11
<i>p-value</i>	—	0.03	0.10	0.01	0.05	0.01	0.04

TAB. 5 – Valeur et significativité des primes résiduelles du modèle de Wei

Le tableau reporte le calcul de la prime λ_m du portefeuille résiduel de marché dans le modèle de Wei: $E[\tilde{R}_i] = \lambda_0 + \lambda_m b_{im} + \sum_{k=1}^K \lambda_k b_{ik}$, en utilisant respectivement l'analyse en composantes indépendantes et l'analyse en composantes principales, pour les données quotidiennes, hebdomadaires et mensuelles. Le tableau reporte les valeurs du test de Student sur ces primes, ainsi que les p-value correspondantes. Le nombre de facteurs pour lequel cette prime ne devient plus significative est de deux pour les données quotidiennes et hebdomadaires en utilisant l'analyse en composantes indépendantes, la méthode échouant totalement pour les données mensuelles qui sont sous-échantillonnées. Le nombre de facteurs nécessaires obtenus par l'analyse en composantes principales est supérieur, puisqu'il est de quatre pour les données quotidiennes et de trois pour les données hebdomadaires.

méthode de décomposition montrent clairement que deux facteurs sont nécessaires pour les données quotidiennes ou hebdomadaires. La p-value d'acceptation du test est cependant limite pour ces décompositions. La méthode échoue totalement pour les données mensuelles. Il faut noter ici que le nombre d'observations pour ces données est très insuffisant, puisque sur quatre ans, 48 observations sont utilisées, alors que le nombre d'actifs est en moyenne de 130. Les résultats obtenus par l'analyse en composantes principales sont inférieurs aux résultats obtenus par l'analyse en composantes indépendantes. Pour les données quotidiennes, le nombre de facteurs nécessaires pour que la prime résiduelle soit non significative est au minimum de quatre facteurs; de plus la p-value d'acceptation du test pour ces données reste faible, autour de 0.01, jusqu'à une décomposition à sept facteurs. Pour les données hebdomadaires, le nombre de facteurs minimal nécessaire est de trois.

Le tableau 6 présente les détails des résultats obtenus pour un nombre de deux à sept facteurs pour différentes sous-périodes de cinq ans, le commencement des périodes étant séparées par des intervalles de trois années, en employant respectivement l'analyse en composantes principales et l'analyse en composantes indépendantes. Pour toutes les sous-périodes, la prime résiduelle n'est pas significative avec deux facteurs, sauf pour une sous-période allant de 1987 à 1991, incluant donc les années 1987 et 1989, pendant lesquelles se sont produites deux baisses brutales de marché. Pour cette période, l'analyse en composantes principales échoue à obtenir une structure factorielle correcte pour un nombre de facteur inférieur à huit, la prime résiduelle restant toujours significative quel que soit le nombre de facteurs. L'analyse en composantes indépendantes permet par contre d'obtenir une structure factorielle correcte par l'ajout d'un seul facteur pour cette période, facteur qui n'est pas nécessaire à la structure sur plus long terme.

La comparaison des deux méthodes d'analyse montrent la supériorité manifeste de l'analyse en composantes indépendantes, puisque pour les données quotidiennes, seuls deux facteurs au lieu de quatre sont nécessaires pour obtenir une prime de Wei nulle (dans le sens qu'elle est non significative), alors que l'analyse en composantes principales ne permet pas de distinguer un nombre de facteurs adéquat inférieur à quatre. Ce nombre est d'autre part le même que la fréquence de données soit quotidienne ou hebdomadaire.

Analyse en Composantes Principales

# facteurs	2	3	4	5	6	7
75/79	0.04	6.17	3.49	3.21	5.55	5.10
	0.00	0.41	0.21	0.20	0.37	0.34
78/82	16.52	17.29	15.83	20.26	2.74	0.82
	0.66	0.71	0.66	0.86	0.12	0.04
81/85	41.33	24.84	13.38	12.27	1.00	21.53
	1.69	0.99	0.55	0.50	0.04	0.80
84/88	21.03	21.34	7.07	19.98	2.43	19.11
	1.07	1.07	0.36	0.92	0.11	0.79
87/91	59.04	64.72	76.01	44.69	54.56	53.58
	2.75	2.92	3.32	1.80	2.20	2.19
90/94	22.22	26.90	24.96	15.42	12.89	14.34
	1.03	1.36	1.26	0.60	0.51	0.57
93/97	17.42	18.50	19.20	14.10	13.65	6.55
	1.04	1.11	1.15	0.86	0.83	0.40
Total	12.28	10.34	8.33	8.34	8.22	5.94
	2.49	2.05	1.62	1.58	1.57	1.11

Analyse en Composantes Indépendantes

# facteurs	2	3	4	5	6	7
75/79	8.80	17.12	14.61	14.22	12.00	3.00
	0.56	1.21	1.07	1.02	0.86	0.23
78/82	13.17	8.32	6.29	3.08	29.38	13.34
	0.60	0.40	0.30	0.16	1.23	0.57
81/85	35.81	27.47	34.46	5.90	9.19	6.16
	1.20	0.93	1.13	0.18	0.27	0.18
84/88	13.90	17.71	13.15	9.79	12.02	8.58
	0.85	0.64	0.47	0.35	0.47	0.34
87/91	57.50	30.76	33.72	33.42	35.72	36.51
	2.56	1.16	1.26	1.25	1.35	1.36
90/94	24.87	47.38	62.46	55.57	58.74	70.29
	1.06	1.50	1.69	1.42	1.53	1.84
93/97	24.94	18.27	18.22	18.10	18.79	17.72
	2.29	1.71	1.67	1.67	1.69	1.61
Total	9.98	8.16	6.12	3.25	1.03	2.20
	1.84	1.35	0.99	0.52	0.16	0.33

TAB. 6 – *Détail des évaluations des primes de Wei par période*

Le tableau présente le détail des valeurs et des t-stat des primes des portefeuilles résiduels de marché, pour des périodes de cinq ans, séparées par intervalles de trois années. Pour chaque période, la structure factorielle est estimée sur les quatre premières années, et les primes calculées sur la cinquième année. Pour chaque période, deux facteurs sont suffisants en employant l'analyse en composantes indépendantes, et quatre pour l'analyse en composantes principales. Un troisième facteur supplémentaire n'est nécessaire pour l'analyse en composantes indépendantes que dans la période allant de 1987 à 1991, période incluant deux baisses brutales du marché en 1987 et 1989. Pour cette période, la prime résiduelle reste toujours significative pour l'analyse en composantes principales, cette méthode échouant donc à déterminer une structure correcte avec moins de huit facteurs.

4.4 Test de Connor-Korajczyk

Pour confirmer le résultat obtenu par le test de la prime résiduelle du modèle de Wei, en particulier pour les données quotidiennes où l'acceptation de l'hypothèse est proche du seuil critique, nous emploierons également le test développé par Connor et Korajczyk (1993) [13]. Ce test inclut explicitement le fait que la structure factorielle de génération des rendements peut être une structure factorielle approximative, en supposant que la matrice de variance des erreurs de la décomposition peut être diagonale, selon la généralisation du MEA faite par Chamberlain et Rothschild (1983) [7] et Ingersoll (1984) [20].

Connor et Korajczyk (1993) [13] proposent ainsi un test basé sur le fait que si K est le nombre adéquat de facteurs, il ne doit pas y avoir de diminution significative des rendements idiosyncratiques lorsque l'on passe de K à $K + 1$ facteurs: dans ce cas, la répartition du risque ne nécessite pas l'adjonction d'un facteur supplémentaire à la structure factorielle commune. Ce test est donc proche du test du MEA avec le modèle de Wei, avec l'ajout d'un facteur qui est le portefeuille résiduel de marché. Il en diffère cependant par le fait que les structures factorielles sont entièrement réestimées avec K et $K + 1$ facteurs, et que le test porte sur la comparaison des portefeuilles résiduels de ces deux décompositions.

Avec une structure factorielle à K facteurs, le modèle de MEA s'écrit au temps t pour un nombre N d'actifs:

$$\tilde{\mathbf{r}}_t = c + \mathbf{B}\tilde{\mathbf{f}}_t + \tilde{\epsilon}_t, \quad (27)$$

où \mathbf{B} est dimension $N \times K$, et $\tilde{\mathbf{f}}_t$ de dimension $K \times 1$. En ajoutant un $(K + 1)$ -ème facteur dans la décomposition, le modèle de MEA à $K + 1$ facteurs s'écrit:

$$\tilde{\mathbf{r}}_t = c + \mathbf{B}\tilde{\mathbf{f}}_t + \beta\tilde{f}_t^* + \tilde{\epsilon}_t^*, \quad (28)$$

où β est un vecteur de dimension $N \times 1$.

En considérant que si les rendements des actifs sont décrits correctement par une structure factorielle approximative à K facteurs, les sensibilités au $(K + 1)$ -ème facteur peuvent avoir une valeur non nulle pour certains actifs, mais pas pour une proportion significative de ceux-ci, Connor et Korajczyk proposent alors d'utiliser la statistique suivante pour tester la détermination du nombre correct de facteurs K :

$$\nu_t - \nu_t^* = \epsilon_t' \epsilon_t - \epsilon_t^* \epsilon_t^*, \quad t = 1..T \quad (29)$$

La statistique $\nu_t - \nu_t^*$ ne peut cependant pas être étudiée directement, car elle converge vers une distribution dégénérée quand $N \rightarrow \infty$ ($\frac{\nu_t - \nu_t^*}{\sqrt{N}} = \frac{f_t^* \beta' \beta}{\sqrt{N}}$ tend vers 0 à la limite). C'est pourquoi l'on sépare les données en deux parties, l'une pour les t pairs et l'autre pour les t impairs. La statistique étudiée Δ_s est alors définie par:

$$\Delta_s = \nu_{2s-1} - \nu_{2s}, \quad s = 1..T/2. \quad (30)$$

Par les hypothèses faites sur la structure des rendements, les ϵ_{it} et $\epsilon_{i\tau}^*$ sont indépendants pour $t \neq \tau$ et identiquement distribués. Pour tenir compte du fait que ν_s et ν_s^* sont des variables estimées, celles-ci sont corrigées respectivement par les facteurs $1/(1 - (K+1)/T - K/N)$ et $1/(1 - (K+2)/T - (K+1)/N)$. Sous les hypothèses adéquates, Connor et Korajczyk montrent que Δ converge vers une distribution de variables indépendantes normalement distribuées:

$$d \lim_{n \rightarrow \infty} \sqrt{N} \hat{\Delta} \simeq N(0, \Gamma), \quad (31)$$

et sous l'hypothèse nulle de K facteurs adéquats, $\hat{\Delta} \Gamma^{-1/2}$ est asymptotiquement distribué comme une normale standardisée.

Le tableau 7 reporte les résultats globaux du test de Connor-Korajczyk sur le marché suisse en employant l'analyse en composantes indépendantes et l'analyse en composantes principales. Les résultats confirment ceux obtenus en testant la prime de marché résiduel du modèle de Wei, avec des niveaux de signification plus grands. Quelle que soit la fréquence des données, quotidienne, hebdomadaire ou mensuelle, il est en effet inutile de rajouter un facteur à une décomposition à deux facteurs par l'analyse en composantes indépendantes. En particulier le test confirme le nombre de deux facteurs nécessaires pour les données quotidiennes, alors que le test de significativité de la prime résiduelle pour ces données était trop proche des valeurs critiques pour être assuré. Pour l'analyse en composantes principales, si un seul facteur semble nécessaire pour les données mensuelles, au moins quatre facteurs sont nécessaires en ce qui concerne les données quotidiennes pour que l'ajout d'un facteur supplémentaire à la décomposition soit inutile. Pour les données hebdomadaires, ce nombre minimal est de trois.

Le tableau 8 présente les détails des résultats du test de Connor-Korajczyk sur les mêmes sous-périodes que celles qui ont été utilisées dans l'étude du modèle de Wei, savoir des sous-périodes de cinq ans, le commencement des périodes étant séparées par des intervalles de trois années. L'examen de ces sous-périodes semble montrer que la décomposition à un facteur

Analyse en Composantes Principales

# facteurs	1 à 2	2 à 3	3 à 4	4 à 5	5 à 6	6 à 7	7 à 8
Quotidiennes	40.78	29.06	22.20	19.03	18.05	17.05	16.26
t-stat	4.09	2.63	1.99	1.71	1.62	1.55	1.48
<i>p-value</i>	0.00	0.01	0.05	<i>0.09</i>	<i>0.11</i>	<i>0.12</i>	<i>0.14</i>
Hebdomadaires	48.99	35.56	30.35	26.55	24.58	23.49	21.83
t-stat	2.91	2.03	1.79	1.62	1.51	1.45	1.36
<i>p-value</i>	0.00	0.04	<i>0.07</i>	<i>0.11</i>	<i>0.13</i>	<i>0.15</i>	<i>0.17</i>
Mensuelles	34.50	15.19	1.98	-7.19	-13.46	-16.77	-17.99
t-stat	1.25	0.54	0.07	-0.26	-0.50	-0.64	-0.71
<i>p-value</i>	<i>0.21</i>	<i>0.59</i>	<i>0.94</i>	<i>0.79</i>	<i>0.61</i>	<i>0.52</i>	<i>0.48</i>

Analyse en Composantes Indépendantes

# facteurs	1	2 à 3	3 à 4	4 à 5	5 à 6	6 à 7	7 à 8
Quotidiennes	—	7.11	4.53	4.34	3.96	3.33	2.79
t-stat	—	1.31	0.78	0.74	0.66	0.56	0.46
<i>p-value</i>	—	<i>0.19</i>	<i>0.44</i>	<i>0.46</i>	<i>0.51</i>	<i>0.58</i>	<i>0.64</i>
Hebdomadaires	—	24.03	20.03	17.02	16.08	15.35	15.30
t-stat	—	1.73	1.48	1.30	1.25	1.19	1.18
<i>p-value</i>	—	<i>0.09</i>	<i>0.14</i>	<i>0.19</i>	<i>0.21</i>	<i>0.23</i>	<i>0.24</i>
Mensuelles	—	14.06	4.83	1.05	-5.51	-9.86	-12.82
t-stat	—	0.54	0.18	0.04	-0.22	-0.41	-0.55
<i>p-value</i>	—	<i>0.59</i>	<i>0.85</i>	<i>0.97</i>	<i>0.82</i>	<i>0.68</i>	<i>0.58</i>

TAB. 7 – Résultats du test de Connor-Korajczyk

Le tableau reporte la valeur du test de Connor-Korajczyk, en utilisant respectivement l'analyse en composantes principales et l'analyse en composantes indépendantes, pour les données quotidiennes, hebdomadaires et mensuelles. Les résultats obtenus confirment les résultats obtenus en testant la significativité de la prime résiduelle du marché. Le nombre minimale de facteurs nécessaires donné par le test est de quatre pour les données quotidiennes, trois pour les données hebdomadaires, et un pour les données mensuelles, en ce qui concerne l'analyse en composantes principales. Pour l'analyse en composantes indépendantes, le nombre donné par le test est uniformément de deux quelle que soit la fréquence des données utilisée.

est correcte. Cependant, d'autres sous-périodes existent, qui ne sont pas reportées ici, pour lequel ce nombre n'est pas suffisant. Pour l'analyse en composantes indépendantes, pour toutes les sous-périodes, y compris celles qui ne sont pas reportées ici, le test donne que le nombre de deux facteurs est suffisant et que l'ajout d'un facteur supplémentaire n'améliore pas la décomposition.

Les tests de Connor-Korajczyk confirment donc les résultats des tests obtenus en considérant le MEA dans la formulation de Wei, et conforte la confiance que l'on peut y accorder. En particulier, ils confortent les résultats des tests qui présentaient des valeurs proches des valeurs limites de rejet, et qui de ce fait auraient pu être discutables. La formulation du test de

Analyse en Composantes Principales

# facteurs	1 à 2	2 à 3	3 à 4	4 à 5	5 à 6	6 à 7	7 à 8
75/79	36.17	23.97	18.37	13.75	12.16	8.83	6.28
	2.59	1.64	1.22	0.90	0.81	0.61	0.43
78/82	11.60	5.25	2.48	1.40	0.34	1.61	1.38
	0.85	0.38	0.18	0.10	0.02	0.12	0.10
81/85	60.14	50.00	47.26	43.94	40.12	39.54	37.88
	1.46	1.21	1.16	1.10	1.02	1.02	0.99
84/88	-1.39	5.96	13.31	9.43	12.00	14.81	14.29
	-0.03	0.11	0.24	0.17	0.21	0.26	0.25
87/91	23.61	-24.54	-26.09	-33.96	-41.09	-45.24	-41.36
	0.55	-0.34	-0.36	-0.46	-0.56	-0.61	-0.55
90/94	20.95	-8.23	-24.93	-26.67	-26.17	-26.59	-28.10
	0.97	-0.36	-1.05	-1.12	-1.10	-1.12	-1.18
93/97	160.65	102.30	67.33	67.66	65.17	62.59	58.50
	1.53	0.98	0.63	0.64	0.62	0.60	0.57
Total	40.78	29.06	22.20	19.03	18.05	17.05	16.26
	4.09	2.63	1.99	1.71	1.62	1.55	1.48

Analyse en Composantes Indépendantes

# facteurs	1 à 2	2 à 3	3 à 4	4 à 5	5 à 6	6 à 7	7 à 8
75/79	—	12.83	11.78	8.91	11.09	10.74	12.25
	—	0.92	0.85	0.63	0.79	0.77	0.89
78/82	—	4.56	3.76	5.57	2.20	0.80	0.01
	—	0.33	0.27	0.41	0.16	0.06	0.00
81/85	—	6.19	7.67	5.69	5.85	5.56	4.97
	—	0.68	0.82	0.60	0.60	0.58	0.52
84/88	—	13.55	10.88	12.50	12.00	14.11	18.16
	—	0.34	0.27	0.30	0.27	0.32	0.39
87/91	—	-15.35	-36.41	-37.12	-40.07	-43.84	-44.58
	—	-0.33	-0.62	-0.61	-0.67	-0.73	-0.73
90/94	—	5.54	3.50	0.78	0.88	8.36	0.28
	—	0.26	0.17	0.04	0.04	0.41	0.01
93/97	—	3.99	2.51	2.97	2.18	0.11	1.02
	—	0.46	0.28	0.33	0.24	0.01	0.11
Total	—	7.11	4.53	4.34	3.96	3.33	2.79
	—	1.31	0.78	0.74	0.66	0.56	0.46

TAB. 8 – *Détails des résultats du test de Connor-Korajczyk par période*

Le tableau présente le détail des résultats du test de Connor-Korajczyk pour des périodes de cinq ans, séparées par intervalles de trois années. Pour chaque période, les quatre premières années servent à estimer la structure factorielle, et la cinquième année au calcul de la prime. Sur l'ensemble de la période, deux facteurs sont suffisants en employant l'analyse en composantes indépendantes alors que l'emploi de l'analyse en composantes principales nécessite au moins quatre facteurs. Si l'on conduit le test sur des sous-périodes, il ne paraît pas nécessaire d'employer une structure factorielle à plus d'un facteur.

Connor-Korajczyk est proche de celle du test du modèle de Wei qui a été effectué à la section précédente. Cependant, le test de prime résiduelle de Wei permet de conclure sur l'absence d'un facteur oublié qui soit rémunéré par le marché. La statistique de Connor-Korajczyk permet de tester qu'il n'y a pas de diminution significative de cette prime résiduelle si un facteur supplémentaire est ajouté à la décomposition. Comme nous l'avons discuté précédemment à la section 5.2.2.5, l'absence de rémunération d'un facteur ne permet pas nécessairement de conclure à l'absence de ce facteur, car sa rémunération dépend également de l'évolution du facteur au cours de la période. Le test de Connor-Korajczyk est donc une confirmation supplémentaire de l'absence de ce facteur oublié car il ne nécessite pas que ce facteur soit effectivement rémunéré ou non sur la période considérée.

5 Conclusion

Nous avons montré ici sur le marché suisse que l'emploi d'une nouvelle méthode de décomposition, l'analyse en composantes indépendantes, dans la recherche des facteurs communs, permet une amélioration significative des tests du Modèle d'Evaluation par Arbitrage.

Quatre tests ont été ici utilisés afin de déterminer le nombre de facteurs nécessaires à la structure factorielle. Les deux tests les plus utilisés empiriquement semblent inadaptés au Modèle d'Evaluation par Arbitrage. Les méthodes de détermination a priori du nombre de facteurs nécessaires, reposant sur l'analyse des matrices de variance-covariance, conduisent à surestimer le nombre de ces facteurs, qui serait supérieur à cinq voire dix. Cette détermination a priori suppose une structure factorielle *stricte*, avec des covariances entre les termes résiduels asymptotiquement nuls. Si l'on admet une hypothèse de structure factorielle *approximative*, moins de facteurs sont nécessaires pour annuler les risques résiduels du modèle. Les autres tests menés montrent qu'un nombre bien inférieur de facteurs est en fait nécessaire. Le critère d'examen de la significativité des primes les plus faibles, c'est-à-dire de celles dont les valeurs de Student sont les plus petites en valeurs absolues, ne permet pas d'obtenir de conclusion sur le nombre de facteurs, puisque l'absence de rémunération, ou la rémunération faible, d'un facteur sur une ou plusieurs périodes ne signifie pas qu'il puisse être omis de la structure factorielle. La seule conclusion que nous avons pu tirer de ce critère est que sur certaines sous-périodes, au moins deux facteurs étaient effectivement rémunérés, ce qui semble montrer la pertinence d'un modèle multifactoriel.

Seules des méthodes qui s'appuient dans un premier temps sur l'estimation de la structure factorielle pour un nombre de facteurs déterminé, et dans un deuxième temps sur la validation de cette structure par l'étude des primes de risque résiduelles, semblent de nature à déterminer efficacement le nombre de facteurs. La méthode adéquate paraît ainsi être l'étude de l'absence de primes résiduelles rémunérées, c'est-à-dire de l'absence de facteurs oubliés dans la structure factorielle estimée. Le test de prime résiduelle dans le modèle de Wei montre qu'avec l'analyse en composantes indépendantes, deux facteurs sont nécessaires, que ce soit pour les données quotidiennes ou hebdomadaires. Le test de Connor-Korajczyk confirme le test de prime résiduelle en montrant qu'un facteur supplémentaire n'apporte pas d'amélioration significative du modèle à deux facteurs, et ce quelle que soit la fréquence des données utilisées. Lorsque l'analyse en composantes principales est employée, les deux tests montrent qu'au moins deux facteurs supplémentaires sont nécessaires pour les données quotidiennes et un facteur supplémentaire pour les données hebdomadaires.

La méthode d'analyse en composantes indépendantes proposée semble donc plus efficace que l'analyse factorielle couramment employée, en particulier pour des données relevées avec une fréquence quotidienne, puisque le nombre de facteurs nécessaires à la décomposition est plus petit. D'autre part, la structure factorielle obtenue par la méthode d'analyse en composantes indépendantes est composée de deux facteurs quelle que soit la fréquence des données utilisées, contrairement à l'analyse en composantes principales dont le nombre varie de un à quatre, selon que les données sont de fréquence mensuelle, hebdomadaire ou quotidienne. Enfin, l'analyse des facteurs a permis d'identifier le premier facteur au portefeuille de marché, ou du moins à l'indice qui en est le proxy mais qui tient compte de tous les actifs traités sur le marché suisse des actions; dans le cas de l'analyse en composantes principales, le premier facteur est proche de l'indice mais ne peut pas y être identifié. Cette étude portant sur le marché suisse des actions emploie la partie essentielle et la plus pertinente des données disponibles sur ce marché. Une période de temps relativement longue pour ce type d'études a été employée, puisqu'au total vingt-deux ans de données ont été utilisées, depuis 1975, année à partir de laquelle des données fiables sont disponibles sur le marché suisse. Les données ont été utilisées suivant une fréquence quotidienne, hebdomadaire ou mensuelle, afin de vérifier la stabilité des résultats suivant la fréquence employée. Enfin, les actions ont été traitées individuellement; les études portant sur les marchés procèdent généralement à la constitution de portefeuilles, ce qui conduirait à un nombre très restreint de portefeuilles sur

un marché aussi étroit que l'est le marché suisse; d'autre part l'emploi d'actifs individuels est beaucoup plus efficace puisqu'il permet d'inclure toutes les interdépendances entre les actifs d'une part, et se place d'autre part dans le cas a priori le plus défavorable puisque la constitution de portefeuilles a pour effet de gommer les interdépendances et de «lisser» les séries de rendements.

Les tests menés sur les primes résiduelles permettent ainsi d'accepter le Modèle d'Évaluation à deux facteurs comme modèle d'évaluation sur le marché suisse pour la période allant du 1er janvier 1975 au 31 décembre 1997. L'étude des anomalies qui a été menée avec le modèle à deux facteurs dans Vesserau (1999) [35] montre que la prise en compte de ce second facteur est de nature à expliquer plusieurs des anomalies généralement observées lors de l'emploi du MEDAF classique. L'un des facteurs semble pouvoir être assimilé au portefeuille de marché (approximé ici par l'indice Vontobel de DATASTREAM); il reste à déterminer le second facteur intervenant dans cette décomposition. L'existence de ce modèle à deux facteurs permet d'envisager ce deuxième facteur comme une prime liée à la liquidité ou à la négociabilité des titres, comme semble l'indiquer le modèle proposé par Merton (1987) [25], où le MEDAF est étendu dans le cas où tous les investisseurs ne sont pas informés sur tous les actifs, ce qui mène à l'introduction d'un facteur supplémentaire dans la formulation de base du MEDAF, propre à chaque actif, et dépendant de sa liquidité et de sa taille.

Références

- [1] S. Amari, T.P. Chen, and A. Cichocki, *Stability analysis of learning algorithms for blind source separation*, Neural Networks **10** (1997), 1345–1351.
- [2] S. Amari, A. Cichocki, and H.H. Yang, *Recurrent neural networks for blind separation of sources*, Proceedings of International Symposium on Nonlinear Theory and its Applications, NOLTA-95 (1995), 37–42.
- [3] S. Beckers, P. Cummins, and C. Woods, *The estimation of multiple factor models and their applications: The swiss equity market*, Finanzmarkt und Portfolio Management **7** (1993), 24–45.
- [4] A.J. Bell and T.J. Sejnowski, *An information-maximization approach to blind separation and blind decomposition*, Neural Computation **7** (1995), 1129–1159.
- [5] S.J. Brown, *The number of factors in security returns*, Journal of Finance **44** (1989), 1247–1262.
- [6] J.F. Cardoso and B. Laheld, *Equivariant adaptive source separation*, IEEE Transaction on Signal Processing **44** (1996), 3017–3030.
- [7] G. Chamberlain, *A characterization of the distributions that imply mean-variance utility functions*, Journal of Economic Theory **29** (1983), 185–201.
- [8] N. Chen, *Some empirical tests of the theory of arbitrage pricing*, Journal of Finance **38** (1983), 1393–1414.
- [9] A. Cichoki, R. Unbehauen, L. Moczczynski, and E. Rummert, *A new on-line adaptive learning algorithm for blind separation of source signals*, Proceedings of International Symposium of Artificial Neural Networks, ISANN-94 (1994), 406–411.
- [10] P. Common, *Independent component analysis, a new concept?*, Signal Processing **36** (1994), 287–314.
- [11] G. Connor, *A unified beta pricing theory*, Journal of Economic Theory **34** (1984), 13–31.
- [12] G. Connor and R.A. Korajczyk, *Risk and return in an equilibrium APT: Application of a new test methodology*, Journal of Financial Economics **21** (1988), 373–394.
- [13] ———, *A test for the number of factors in an approximate factor model*, Journal of Finance **48** (1993), 1263–1291.
- [14] P. Dhrymes, I. Friend, and B. Gultekin, *A critical reexamination of the empirical evidence on the arbitrage pricing theory*, Journal of Finance **39** (1984), 323–346.
- [15] E. Fama and J. MacBeth, *Risk, return and equilibrium: Empirical tests*, Journal of Political Economy **38** (1973), 607–636.
- [16] M. Grinblatt and S. Titman, *Factor pricing in a finite economy*, Journal of Financial Economics **12** (1983), 497–507.
- [17] ———, *The relation between mean-variance efficiency and arbitrage pricing*, Journal of Business **60** (1987), 97–112.
- [18] J. Herault and C. Jutten, *Blind separation of sources, part 1: an adaptive algorithm based on neuromimetic architecture*, Signal Processing **24** (1991), 1–10.
- [19] G. Huberman, *A simple approach to arbitrage pricing theory*, Journal of Economic Theory **28** (1982), 183–191.
- [20] J.E. Ingersoll, *Some results in the theory of arbitrage pricing*, Journal of Finance **39** (1984), 1021–1039.

- [21] S.Y. Lee, *Some empirical tests on the arbitrage pricing theory: A portfolio approach*, PhD Thesis (1988).
- [22] B. Lehmann and D. Modest, *The empirical foundations of the arbitrage pricing theory*, Journal of Financial Economics **21** (1988), 213–254.
- [23] C. Marañon, *L'approche multifactorielle et les facteurs déterminants dans l'évaluation des actifs en suisse*, Unpublished manuscript (1994).
- [24] J. Mei, *A semiautoregression approach to the arbitrage pricing theory*, Journal of Finance **48** (1993), 599–620.
- [25] R.C. Merton, *A simple model of capital market equilibrium with incomplete information*, Journal of Finance **42** (1987), 483–510.
- [26] D. Morrison, *Multivariate statistical methods*, McGraw Hill, 1990.
- [27] E. Oja, *A simplified neuron model as a principal component analyser*, Journal of Math. Biology **15** (1982), 267–273.
- [28] E. Oja and J. Karhunen, *On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix*, Journal of Mathematical Analysis and Applications **104** (1985), 69–84.
- [29] ———, *Signal separation by nonlinear hebbian learning*, Computational Intelligence - a Dynamix System Perspective, New-York (M. Palaniswami, ed.), IEEE, 1995.
- [30] R. Roll and S.A. Ross, *An empirical investigation of the arbitrage pricing theory*, Journal of Finance **35** (1980), 1073–1103.
- [31] S.A. Ross, *The arbitrage theory of capital asset pricing*, Journal of Economic Theory **13** (1976), 343–362.
- [32] J. Shanken, *The current state of the arbitrage pricing theory*, Journal of Finance **47** (1992), 1569–1574.
- [33] N. Sinclair, *Security return data and blind factor analysis*, Working Paper 81-015, University of New South Wales Australian, Graduate School of Management, Australia, 1981.
- [34] C. Trzcinka, *On the number of factors in the arbitrage pricing model*, Journal of Finance **41** (1986), 347–368.
- [35] T. Vessereau, *Application des réseaux neuronaux en finance: l'analyse multifactorielle des actifs et l'évaluation non paramétrique des options*, Ph.D. thesis, Université de Genève, 1999.
- [36] J. Wei, *An asset-pricing theory unifying the CAPM and the APT*, Journal of Finance **43** (1987), 881–892.

Liste des publications au CIRANO *

Cahiers CIRANO / *CIRANO Papers* (ISSN 1198-8169)

- 99c-1 Les Expos, l'OSM, les universités, les hôpitaux : Le coût d'un déficit de 400 000 emplois au Québec — Expos, Montréal Symphony Orchestra, Universities, Hospitals: The Cost of a 400,000-Job Shortfall in Québec / Marcel Boyer
- 96c-1 Peut-on créer des emplois en réglementant le temps de travail? / Robert Lacroix
- 95c-2 Anomalies de marché et sélection des titres au Canada / Richard Guay, Jean-François L'Her et Jean-Marc Suret
- 95c-1 La réglementation incitative / Marcel Boyer
- 94c-3 L'importance relative des gouvernements : causes, conséquences et organisations alternative / Claude Montmarquette
- 94c-2 Commercial Bankruptcy and Financial Reorganization in Canada / Jocelyn Martel
- 94c-1 Faire ou faire faire : La perspective de l'économie des organisations / Michel Patry

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

- 2000s-43 Habit Formation with Recursive Preferences / Aylin Seckin
- 2000s-42 Habit Formation: A Kind of Prudence? / Aylin Seckin
- 2000s-41 Consumption with Liquidity Constraints and Habit Formation / Aylin Seckin
- 2000s-40 Consumption-Leisure Choice with Habit Formation / Aylin Seckin
- 2000s-39 Consumption with Habit Formation / Aylin Seckin
- 2000s-38 Consumption with Durability / Aylin Seckin
- 2000s-37 Le financement des biens publics par des contributions volontaires: Une évaluation à l'aide de l'économie expérimentale / Claudia Keser
- 2000s-36 Conventions and Local Interaction Structures: Experimental Evidence / Siegfried K. Berninghaus, Karl-Martin Ehrhart et Claudia Keser
- 2000s-35 Strategically Planned Behavior in Public Good Experiments / Claudia Keser
- 2000s-34 Why Do Experimental Subjects Choose an Equilibrium which Is Neither Payoff Nor Risk Dominant? / Claudia Keser et Bodo Vogt
- 2000s-33 The Measurement and Antecedents of Turnover Intentions among IT Professionals / Guy Paré et Michel Tremblay
- 2000s-32 Can the Theory of Incentives Explain Decentralization? / Michel Poitevin
- 2000s-31 IT Outsourcing Risk Management at British Petroleum / Benoit A. Aubert, Michel Patry, Suzanne Rivard et Heather Smith
- 2000s-30 A Resource-Based Analysis of IT Outsourcing / Vital Roy et Benoit Aubert
- 2000s-29 The Impact of Government-Sponsored Training Programs on the Labor Market Transitions of Disadvantaged Men / Lucie Gilbert, Thierry Kamionka et Guy Lacroix

* Vous pouvez consulter la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site Internet à l'adresse suivante :