

JEAN-FRANÇOIS MALÉCOT

## **L'APT et la question du nombre de facteurs**

*Journal de la société statistique de Paris*, tome 133, n° 4 (1992),  
p. 161-165

[http://www.numdam.org/item?id=JSFS\\_1992\\_\\_133\\_4\\_161\\_0](http://www.numdam.org/item?id=JSFS_1992__133_4_161_0)

© Société de statistique de Paris, 1992, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# L'APT ET LA QUESTION DU NOMBRE DE FACTEURS

par Jean-François MALÉCOT  
*Université Paris X-Nanterre, et  
Cereg de Paris Dauphine*

Au sujet des facteurs de l'APT, deux questions sont habituellement soulevées :

- Comment un investisseur est-il sûr d'avoir déterminé le nombre exact de facteurs évalués par le marché ?
- Quelles sont les variables pertinentes qui peuvent déterminer tel ou tel facteur de risque ?

L'objet des remarques qui suivent est d'inviter à la réflexion sur ces points sensibles relatifs aux tests empiriques de l'Arbitrage Pricing Theory et à la mise en pratique qui peut en être faite pour la gestion de portefeuille. A partir du moment où le modèle supposé est conforme aux évolutions structurelles du prix des actifs, la détermination et l'identification des facteurs devient un point clé de toute utilisation pratique qui, s'il n'est pas vraiment résolu, rend l'APT difficilement applicable avec rigueur.

Une controverse règne d'ailleurs sur cette question depuis les travaux de Dhrymes, Friend et Gultekin (1984) et la réponse de Roll et Ross (1984). En effet, selon Dhrymes, Friend et Gultekin, le nombre de facteurs s'accroît au fur et à mesure que le nombre de titres et d'observations s'élèvent dans les portefeuilles. Ainsi, pour des portefeuilles de 30 titres, un modèle à 5 facteurs serait acceptable (selon le critère de l'hypothèse nulle de 5 %) ; mais pour des portefeuilles de 60 titres, un modèle à 8 facteurs serait retenu et pour des portefeuilles de 90 titres, on aurait 13 facteurs. Par ailleurs, si le nombre de facteurs dépend également de la période de test (ce qui semble a priori raisonnable), il dépend également des sous-périodes choisies pour le test.

A cette critique, Roll et Ross (1984) répondent par la question de l'évaluation des facteurs : d'une part, tous les facteurs ne sont pas évalués, autrement dit associés à une prime de risque positive ; d'autre part, certains facteurs sont spécifiques à tel ou tel titre, ou tel ou tel ensemble de titres. Donc, plus le nombre de titres s'élève dans le portefeuille de test, plus on va trouver des facteurs spécifiques en sus des facteurs communs. Les résultats de Dhrymes, Friend et Gultekin sont donc tout à fait compréhensibles. Le seul test du modèle qui soit valide est alors de déterminer quels sont les facteurs qui font effectivement l'objet d'une prime de risque positive (ou une conséquence d'une telle hypothèse).

## L'APT ET LA QUESTION DU NOMBRE DE FACTEURS

Dans ce cadre, la question qui reste cependant non résolue est de savoir si les facteurs, leur nombre étant supposé à peu près stable d'un type de portefeuille à l'autre, sont vraiment communs. Les analyses factorielles de deux sous-ensembles de titres ne sont pas comparables entre elles et rien ne prouve que les facteurs extraits soient les mêmes d'une part, et fassent l'objet de la même évaluation d'autre part (au premier facteur du premier sous-ensemble de titres ne correspond pas forcément le premier facteur d'un autre sous-ensemble de titres, même si ces deux facteurs sont les plus importants en termes de variance expliquée). L'analyse en facteurs communs d'un portefeuille de titres, consécutive à l'hypothèse d'une communauté de facteurs déterminant la rentabilité des titres, ne permet donc pas vraiment de tester la communauté des facteurs. Si ces facteurs sont des forces économiques concrètes et causales, comme l'affirment Roll et Ross (1984), les tests statistiques ne permettent pas d'avoir une réponse sur leurs caractéristiques.

Il faut adopter alors une autre méthodologie, probablement moins rigoureuse sur le plan statistique en cherchant à identifier a priori certains facteurs, le plus souvent macroéconomiques. C'est d'ailleurs un problème épistémologique assez classique : un modèle théorique doit-il être testé sur ses hypothèses ou sur ses conclusions ? D'un point de vue théorique, si la proposition  $H$  est vraie, alors  $P$  est vraie ; mais si  $P$  est vrai, toujours sous certaines conditions expérimentales limitatives et souvent auxiliaires au modèle théorique, rien n'indique que  $H$  le soit aussi. Les mêmes conséquences n'ont pas forcément des causes communes.

La question est de savoir si l'incertitude sur le nombre de facteurs ne va pas perturber les tests lors de la deuxième étape : on comprend bien que tous les facteurs statistiquement « significatifs » ne le soient pas « financièrement ». Mais la question est : l'identification incertaine du nombre de facteurs est-elle sans influence sur les résultats des tests concernant les primes de risque ?

Nous allons exposer ce problème, en reprenant une formalisation résumée de l'APT telle qu'elle est présentée dans Lehman et Modest (1985).

Nous avons en effet, d'après l'APT :

$$R_{it} = E_i + \sum_{k=1}^K b_{ik} \tilde{\delta}_{kt} + \varepsilon_{it}$$

où  $R_{it}$  est la rentabilité pour le titre  $i$ ,  $E_i$  est l'espérance de rentabilité, la réalisation du facteur commun  $k$  pour le titre  $i$  et  $b_{ik}$  la sensibilité du titre à ce facteur  $\tilde{\delta}_{kt}$ .  $\tilde{\varepsilon}_{it}$  est le risque résiduel. On suppose que leur moyenne est nulle, leur variance finie et que la loi des grands nombres s'appliquant, ils sont indépendants d'un titre à l'autre. Par ailleurs, l'on a :

$$E[\tilde{\delta}_{kt}] = E[\tilde{\varepsilon}_{it} | \delta_{kt}] = 0$$

En écriture matricielle, nous avons :

$$R_t = E + B\tilde{\delta}_t + \tilde{\varepsilon}_t$$

la matrice  $B$  étant la matrice des « factors loadings » ( $N \times K$ ). Il est démontré que sous l'hypothèse d'une distribution Normale des rentabilités et des facteurs, la matrice empirique des variances-covariances des rentabilités suit une Loi de Wishart qui va servir de base à l'étude du maximum de la fonction log vraisemblance. Lors de cette étape, on sait qu'il faut estimer la décomposition suivante des rentabilités :

$$\Sigma = BB' + D$$

où  $D$  est une matrice diagonale.

Or le test habituellement utilisé pour déterminer le nombre de facteurs, décrit dans Lawley et Maxwell (p. 35), est un test du Chi Deux. Et comme le remarquent Lehman et Modest (1985), le test porte sur la structure factorielle (avec la matrice  $D$ ) et non sur la structure supposée. Comment s'écrit cette structure supposée ? Si le modèle est vrai, on peut estimer le vecteur  $\tilde{\delta}_t$  par les Moindres Carrés Généralisés, ce qui donnerait :

$$\tilde{\delta}_t = (B' \Omega^{-1} B)^{-1} B' \Omega^{-1} [R_t - E]$$

où la matrice  $\Omega$  est la matrice des variances-covariances des risques résiduels. La décomposition de la matrice  $\Sigma$  serait alors plutôt :

$$\Sigma = BB' + \Omega$$

Or la matrice des variances-covariances des risques résiduels ci-dessus n'est pas forcément diagonale et le ratio des déterminants (le déterminant intervient dans le maximum de vraisemblance) des deux expressions n'est pas égal à 1. Or s'il existe des facteurs communs et que les risques spécifiques de chaque titre soit légèrement corrélés d'une industrie à l'autre, le test du  $\chi^2$  dégagera systématiquement un nombre de facteurs significatifs beaucoup plus important. C'est ce que montrent Lehman et Modest (1985).

Comme la controverse entre Dhrymes, Friend, Gultekin (1984) d'une part et Roll et Ross (1984) d'autre part, porte sur l'utilisation de ce test qui aboutit à retenir un nombre de facteurs croissant avec le nombre de titres dans les portefeuilles, selon les premiers, il convient donc d'interpréter leurs résultats avec précaution. En fait, les tests ne portent pas sur la significativité statistique des primes de risque associées aux facteurs mais soit sur la nullité de la constante quand on régresse la rentabilité d'un portefeuille sur celle du marché (variante des rentabilités brutes) soit sur la somme des factors loadings (on teste si elle est différente de un) et la non nullité de la constante. Lehman et Modest (1988) présentent des résultats les plus complets sur cette question. Leurs résultats ne semblent pas indiquer de différence selon que l'on suppose un modèle à 5, 10 ou 15 facteurs ; quoique que cela ne semble pas influencer les résultats, surtout si l'on compare avec les effets tailles (très petites firmes versus très grandes firmes), le nombre de facteurs n'est pas déterminé.

D'ailleurs, on peut faire observer que l'on dispose d'un moyen alternatif, permettant de tester le nombre de facteurs, tout en tenant compte des dispersions des estimations des coefficients de la matrice  $B$ . En effet, Chatterjee et Pari (1989) ont récemment avancé l'idée d'une procédure reposant sur une simulation du type Boots-

trap de la distribution des coefficients de la matrice  $B$ , et ce par une génération de sous-échantillons aléatoires répliquant l'échantillon initial. A l'aide de ces distributions simulées, ils testent la signification des coefficients par un test du type Student (et de Hotelling pour le test sur plusieurs facteurs). Leurs résultats montrent d'ailleurs que le nombre de facteurs ne paraît pas varier avec le nombre de titres dans les portefeuilles (ils montrent également qu'il n'y a qu'un seul facteur pour lequel tous les titres ont un  $t$  de Student significatif). Ils observent également que les résultats semblent fortement influencés par un petit nombre de titres pour lesquels les fluctuations d'échantillonnages semblent très fortes.

En conclusion, la question du nombre de facteurs en présence reste sans réponse, du moins à ce jour. Cependant, on peut conduire un test empirique de l'APT sans le conditionner à cette question. La recherche des facteurs pourrait alors s'orienter vers d'autres approches comme celle de Burtmeister et McElroy (1988) par exemple, où les facteurs sont (partiellement) identifiés à l'aide de variables macroéconomiques.

L'approche pronée par Burtmeister et McElroy (1988a, 1988b) consiste à déterminer a priori une série pertinente de facteurs macroéconomiques. Burtmeister et McElroy en retiennent cinq : le spread entre les taux des bons du Trésor et des obligations privées, le spread entre les obligations d'État et les bons du Trésor, le taux d'inflation anticipée et le taux de croissance (réel) des ventes de produits finis. Quant au cinquième, il s'agit d'un facteur résiduel, soit la part non expliquée d'un indice (ici le SP 500) par les quatre premiers facteurs.

La méthode consiste à estimer un système d'équations simultanées avec des contraintes sur les coefficients et sur la matrice de variance-covariance, afin que les estimateurs soient cohérents avec les hypothèses de l'APT. Plusieurs techniques sont présentées dans (1988a). Les résultats présentés par Burtmeister (1988b) paraissent tout à fait cohérents, et statistiquement supérieurs à ceux d'un modèle avec un seul indice. Par la suite, des estimations par secteurs, par titres sont possibles avec un « profil » explicatif du risque dans chacun des cas. Des stratégies de gestion passive ou active peuvent ensuite être formulées aisément.

## RÉFÉRENCES

- CHATTERJEE S., PARI R.A. (1990), "Bootstrapping the Number of Factors in the Arbitrage Pricing Theory", *J. of Financial Research*, 15-21.
- BURTMEISTER E., MCELROY M.B. (1988a) "Joint Estimation of Factor Sensitivities and Risk Premia for the Arbitrage Pricing Theory", *J. of Finance*, 721-733.
- BURTMEISTER E., MCELROY M.B. (1988b) "Sorting out Risks Using Known APT Factors", *Financial Analyst Journal*, 29-42.
- DHRYMES P.J., FRIEND I., GULTEKIN N.B. (1984) "A Critical Examination of the Empirical Evidence on the Arbitrage Pricing Theory", *J. of Finance*, 323-346.
- LAWLEY D.N., MAXWELL A.E. (1971) *Factor Analysis as a Statistical Method*, Elsevier.

## L'APT ET LA QUESTION DU NOMBRE DE FACTEURS

LEHMANN B.N., MODEST D.M. (1985) "*The Empirical Foundations of the Arbitrage Pricing Theory II : the Optimal Construction of Basis Portfolio*", Document de recherche non publié.

LEHMANN B.N., MODEST D.M. (1988) "The Empirical Foundations of the Arbitrage Pricing Theory", *J. of Financial Economics*, 21-254.

ROLL R., ROSS S.A. (1984) "A Reply", *J. of Financial Economics*, 347-350.