

# UN ESSAI DE CLASSIFICATION DE TITRES BOURSIERS FONDÉE SUR L'ANALYSE FACTORIELLE

par

**Bernard ZARCA**

## PLAN DE L'ÉTUDE

Introduction .....	48
1. Définition des caractéristiques .....	49
2. Collecte de l'Information .....	52
3. Méthode de traitement .....	53
4. Exposé et Analyse des résultats .....	55
Conclusions .....	64

N.B. — Cette étude a été faite dans le cadre de l'E.R.A. n° 86 du C.N.R.S. et paraîtra simultanément dans la *Revue Économique*, mars 1970.

## INTRODUCTION

Nous précisons, dans cette introduction, les objectifs et les limites de notre étude.

Partant de l'idée qu'il existe des liaisons entre les caractéristiques boursières des titres et les caractéristiques économique-financières des sociétés qui les émettent, nous voudrions les expliciter et, ce faisant, construire une classification de certaines actions françaises cotées à la Bourse de Paris.

Supposer l'existence de liaisons entre variables quantitatives, ce n'est encore rien dire des relations entre les concepts qu'elles recouvrent ni de la nature de ces liaisons. Nous ne voulons pas élaborer une théorie économique du fonctionnement simultané de la Bourse et de l'Entreprise. Nous ne nous appuyons pas sur un corps structuré d'hypothèses concernant ce fonctionnement. C'est même dans la mesure où un tel corps n'existe pas que nous formons ce projet de classification dont l'intérêt pratique apparaît clairement tandis que sa valeur théorique peut échapper par sa minceur : ordonner un savoir parcellaire et faciliter ainsi la constitution de ce corps structuré.

Point d'hypothèse, point d'explication causale. Seulement une description qui, pour construite qu'elle soit — et ce par la technique statistique qu'elle met en œuvre — et « non-naïve », ne fût-ce que par le choix limité des variables prises en compte, n'en reste pas moins empirique.

Certes, nous soupçonnons bien que la politique économique-financière de l'entreprise influence la position boursière du titre qu'elle a émis et, qu'en retour, la conjoncture boursière infléchit les décisions économique-financières de l'entreprise si bien que nos deux groupes de caractéristiques entretiennent des relations circulaires. C'est vers l'étude d'un processus dynamique qu'il faudrait s'orienter. Seule la limitation actuelle de notre information ne nous a pas permis d'entreprendre une telle étude, mais le travail que nous allons présenter n'exclut pas une telle orientation, bien au contraire, il montrera qu'elle est possible.

Les limites de notre étude ainsi précisées, définissons le double objectif que constitue la construction d'une classification.

Étant donné un ensemble (ici l'ensemble fini de cent soixante-deux actions françaises cotées à la Bourse de Paris) construire une classification sur cet ensemble c'est :

- trouver un critère de partition de cet ensemble,
- opérer une partition de cet ensemble.

En général, on opère une partition d'un ensemble en se servant d'un critère déjà connu. Ainsi, par exemple, prenant pour critère : « le compartiment boursier du titre », il est facile de partitionner un ensemble donné d'actions. On divisera l'ensemble en parties deux à deux disjointes, telles que les actions d'une même partie appartiennent au même compartiment boursier. Remarquons que ce critère simple est d'ordre qualitatif et que nous n'avons en rien jugé de sa pertinence.

Or c'est bien dans la mesure où nous cherchons un critère pertinent au problème posé au début de cette introduction et que nous ne pouvons pas nous appuyer sur un corps déjà structuré d'hypothèses pour le choisir « a priori » qu'il nous faut bien le construire.

La méthode mathématique dont nous nous servons permet de construire une classification — approchée — en une seule étape. C'est-à-dire que nous trouvons simultanément un critère de partition et la partition elle-même (approchée, car certains éléments seront à la frontière de deux classes ou seront isolés, ou bien encore et dans le pire des cas, formeront une classe non définie).

La « valeur » du critère pour les éléments d'une classe donnée définit cette classe en compréhension. Les éléments qui lui appartiennent la définissent en extension.

Nous reviendrons, lors de l'exposé des résultats, sur l'idée de classification approchée. Disons tout de suite qu'il n'y a pas de méthode de classification automatique miraculeuse, sorte de boîte noire où l'information entrerait sans ordre aucun et sortirait sous forme de petits paquets bien fermés et étiquetés. Il nous faut nous-même ficeler plus ou moins lâchement les paquets et écrire soigneusement les étiquettes.

Nous avons dit que l'intérêt pratique d'une telle classification apparaissait clairement. Donnons en, pour terminer cette introduction, un exemple en ne cachant pas qu'il fut le point de départ de ce travail : on voudrait établir des correspondances entre la possession de certains titres repérés par leurs caractéristiques économique-financières et boursières et une certaine politique de gestion du portefeuille d'une part, un profil socio-économique de porteurs de valeurs mobilières d'autre part.

Nous avons retenu huit cent titres boursiers environ représentant à eux tous 80 % de la capitalisation boursière de Paris. Notre étude pour être exhaustive eût dû porter sur l'ensemble de ces titres. En fait, nous avons pu, dans un temps relativement court, relever les renseignements nous permettant de construire toutes les caractéristiques retenues de cent soixante-deux actions françaises, parmi lesquelles figurent les titres les plus « connus ». Notre travail est donc fragmentaire et peut très bien se prolonger.

## 1. DÉFINITION DES CARACTÉRISTIQUES

Nous tentons de construire une classification qui tienne compte de trois séries de caractéristiques des titres :

1. — Caractéristiques du marché du titre.
2. — Caractéristiques du titre.
3. — Caractéristiques économique-financières de la société émettrice.

Et qui explicite les relations entre les caractéristiques boursières (1 et 2) et les caractéristiques économique-financières (3).

Ces caractéristiques ont été choisies parmi celles qu'utilisent les analystes financiers et ont un intérêt pratique certain pour les porteurs qui étudieraient le marché boursier en vue de placements ou de dé-placements. Cela ne veut pas dire que nous préjugeons de leur pouvoir discriminant.

Certes, nous croyons qu'en général leur valeur varie d'un titre à l'autre. Mais, il est fort possible que certaines d'entre elles reflètent, au contraire, certaines constantes qu'il n'est pas négligeable de reconnaître, quitte à les écarter ensuite, pour ne garder que les caractéristiques à fort pouvoir discriminant. D'ailleurs, comme nous y insisterons ultérieurement, notre méthode de classification ne privilégie aucune caractéristique. Elle permet, au contraire, de trouver celles qu'on pourrait ensuite privilégier dans un but précis.

Cependant il n'est pas inutile de préciser que nous avons retenu les caractéristiques que l'information dont nous disposions nous permettait de construire. C'est-à-dire que la liste des caractéristiques aurait pu être plus riche, et surtout les caractéristiques elles-mêmes mieux construites (aussi avons nous la claire conscience que ces limites font de notre travail un essai méthodologique). Ainsi, par exemple, les caractéristiques retenues correspondent à une seule année, à savoir 1965 ou plus rarement à une évolution se terminant en 1965. Il eût sans doute été intéressant d'introduire plusieurs indices de croissance et de mesurer la même caractéristique annuelle sur plusieurs années consécutives. Ceci nous eût permis de construire une classification plus stable — car un titre pourrait très bien changer de classe si nous refaisions une étude identique à celle-ci pour 1966 — et de mieux préparer l'étude du processus dynamique d'interaction entre la Bourse et l'Entreprise. Mais, outre le fait que l'information à recueillir eut été beaucoup plus volumineuse et, par suite, le pourcentage de titres pour lesquels elle eut été incomplète beaucoup plus fort, nous eussions quelque peu dépassé notre but initial puisque cette étude devra servir à une classification de portefeuilles considérés en 1965.

Nous donnons, ci-dessous, la liste des caractéristiques retenues ainsi que leur définition et leur mesure qui, on le verra, sera quelquefois approximative.

### 1.1. CARACTÉRISTIQUES DU MARCHÉ DU TITRE

Elles sont au nombre de trois :

**MV 1 = Le marché potentiel du titre** : est défini comme étant la part de la capitalisation boursière non détenue par des sociétés autres que les petites entreprises et que les porteurs institutionnels, et qui en détiennent une part assez importante pour qu'on puisse supposer une moindre mobilité des titres correspondants. C'est donc la masse des titres détenue par les particuliers, les petites entreprises et les porteurs institutionnels et sur lesquels s'effectuent les principales transactions.

**MV 2 = Le marché réel du titre** est défini comme étant la valeur globale des transactions opérées sur lui durant l'année considérée.

**MV 3 W Le rapport du marché réel au marché potentiel** : relative la deuxième caractéristique. Nous parlerons, à ce propos, de **taux de transaction**.

### 1.2. CARACTÉRISTIQUES DU TITRE

Elles indiquent sa position en Bourse : sa rentabilité, sa stabilité, ses potentialités. Un porteur averti pourrait en tenir compte pour gérer son portefeuille. Elles auraient pu être plus nombreuses et plus rigoureusement mesurées. Nous nous sommes limités aux suivantes :

**TC 1 : Le taux de fluctuation des cours** du titre est définie, pour l'année considérée, comme étant le rapport de l'écart entre les cours extrêmes de l'année à son cours moyen annuel.

C'est un indicateur « grossier » de stabilité.

**TC 2 : Le taux de variation du cours au 31 décembre** du titre est défini par :

$$TC 2 = \frac{(\text{cours au 31-12-1965}) - (\text{cours au 31-12-1964})}{(\text{cours au 31-12-1964})}$$

C'est un indicateur de plus value à un an. Il eût été préférable de comparer les valeurs moyennes de ces deux années consécutives, mais nous ne disposons pas toujours de la valeur moyenne en 1964.

**TC 3 : L'anticipation à la croissance** de la valeur d'un titre est définie, pour l'année considérée, comme étant le rapport du cours moyen annuel à la valeur de liquidation du titre.

Rappelons que la valeur de liquidation du titre est égale au rapport de l'actif net de la société au nombre de titres émis et que la valeur de cet actif net est égale à celle des fonds propres de la société.

L'appellation « anticipation à la croissance » reflète l'idée que la Bourse anticipe la croissance d'une société, celle-ci ayant effectué des investissements, et que ce ratio est d'autant plus fort que la croissance anticipée l'est elle-même. L'anticipation est favorable si ce ratio est voisin ou supérieur à l'unité, défavorable s'il est inférieur à l'unité.

A supposer que l'actif net reflète bien la croissance de la société (un biais pouvant s'introduire du fait qu'en 1965 il avait été réévalué pour 20 % des sociétés cotées en Bourse) ce ratio est un indicateur de potentialité du titre.

**TC 4 : Le dividende brut** est le rapport annuel du titre avant impôt.

**TC 5 : Le taux de rendement effectif** du titre est défini comme étant le rapport du dividende brut au cours du 31 décembre de l'année considérée. C'est un indicateur de rentabilité.

**TC 6 : Le taux de progression du dividende en cinq ans** d'un titre est défini par la relation :

$$TC 6 = \frac{(\text{dividende brut 1965}) - (\text{dividende brut 1960})}{(\text{dividende brut 1960})}$$

Cet indicateur permet de saisir grossièrement la tendance de la société à distribuer des dividendes. Sa relation avec le taux d'accroissement du chiffre d'affaires sera particulièrement intéressante.

1. 3. **LES CARACTÉRISTIQUES ÉCONOMICO-FINANCIÈRES** dont on cherchera les relations avec les précédentes, sont des indicateurs du dynamisme, de l'indépendance, de la rentabilité, et de la croissance de l'entreprise.

**1 E 1 : Le taux d'investissement** est égal aux rapports des investissements bruts au chiffre d'affaires.

C'est un indicateur du dynamisme de l'entreprise.

Remarquons que le chiffre d'affaires est mesuré hors taxes alors que l'investissement est mesuré toutes taxes comprises.

**1 E 2 : Le taux d'autofinancement brut** est défini comme étant le rapport de l'autofinancement brut aux investissements bruts de l'année.

C'est un indicateur de l'indépendance de l'entreprise.

L'autofinancement sera mesuré approximativement par la différence entre le cash flow courant et les dividendes versés, ce qui ne tient pas compte des profits et pertes exceptionnels.

Rappelons que le **cash flow courant** est égal au revenu brut courant diminué des frais financiers et des impôts sur bénéfice.

L'autofinancement brut comprend le bénéfice net non distribué, la dotation aux amortissements, les réserves pour fluctuations des cours et renouvellement des stocks, l'autofinancement exceptionnel.

**EI 3 : Le taux d'accroissement moyen du chiffre d'affaires sur cinq ans** : est défini par la formule

$$EI 3 = \sum_{i=61}^{65} \frac{CA_i - CA_{i-1}}{CA_{i-1}}$$

où  $CA_i$  est le chiffre d'affaires de l'année  $i$ .

C'est un indicateur de croissance de l'entreprise.

**EI 4 : La marge bénéficiaire brute** est définie comme étant le rapport du cash flow courant au chiffre d'affaires de l'année considérée.

**EI 5 : Le taux de rotation des actifs** est défini comme étant le rapport du chiffre d'affaires aux fonds propres de la société.

Si ce taux est trop élevé il peut être le signe d'une trop grande dépendance de la société vis-à-vis de ses créiteurs. Inversement s'il est trop faible il peut être à l'origine d'une rentabilité faible. C'est donc un indicateur de bonne gestion.

Rappelons que les fonds propres sont égaux aux capitaux permanents diminués des dettes à plus d'un an.

**EI 6 : La fraction du bénéfice mise en réserve** est définie comme étant la part du bénéfice brut non distribuée.

En fait, on posera :

$$EI 6 = \frac{\text{cash flow courant} - \text{dividendes}}{\text{cash flow courant}}$$

Le bénéfice brut étant égal au cash flow courant aux profits et pertes exceptionnels près.

C'est un indicateur de la tendance à l'autofinancement, et donc indirectement de la politique de distribution des bénéfices de l'entreprise.

**EI 7 : La rentabilité par rapport aux fonds propres** est définie comme étant le rapport :

$$EI 7 = \frac{\text{cash flow courant}}{\text{fonds propres}}$$

C'est un indicateur de rentabilité de l'entreprise.

## 2. COLLECTE DE L'INFORMATION

Nous citons dans ce paragraphe les différentes sources où nous avons recueilli les données nous permettant de construire les caractéristiques définies dans le paragraphe précédent :

**Première source** : Journal de la Cote Officielle.

Nous y avons relevé les cours des titres au 31-12-1964 et 31-12-1965.

**Deuxième source** : S.E.D.E.S.

Nous y avons relevé le nombre de titres émis avant le 1-1-1966.

**Troisième source** : D.A.F.S.A.

Nous avons relevé sur les fiches D.A.F.S.A. 1966 ou, dans la mesure où elles n'existaient pas cette année là, sur celles de 1965 les renseignements suivants :

- chiffre d'affaires hors taxes 1965,
- taux de progression du chiffre d'affaires 1960-1961, 1961-1962, 1962-1963, 1963-1964, 1964-1965,
- dividende brut par action pour les années 1960 et 1965,
- cash flow courant par titre en 1965,
- fonds propres de la société émettrice en 1965,
- investissements bruts en 1965,
- valeur moyenne du titre en 1965.

**Quatrième source :** DOCFIN 1968 « Tout sur l'Économie et les Finances de la France » publié par la S.E.F.

Si nous ne possédons pas les documents D.A.F.S.A. pour les titres retenus, nous y avons relevé les renseignements précédents concernant ces derniers.

Lorsqu'il y avait incompatibilité entre certains renseignements (par exemple, lorsqu'une valeur moyenne était située en dehors des cours extrêmes de l'année), nous nous sommes reportés au Journal de la Cote Officielle de janvier 1966 et nous avons recalculé, en collaboration avec les services de documentation de la Chambre Syndicale des Agents de Change, certaines valeurs moyennes. Nous avons pu ainsi corriger quelques erreurs afin de n'exploiter que des données cohérentes entre elles.

**Cinquième source :** Les services de documentation de la Banque de France nous ont fourni, dans la mesure du possible, la valeur de la masse des transactions opérées sur chaque titre au cours de l'année 1965 (caractéristique MV 2).

**Sixième source :** « Les Liaisons Financières des Entreprises Françaises. Actionnaires et Participations en France et à l'Étranger », 1966, 1<sup>ère</sup> édition, publié par la S.E.F.

Nous y avons relevé, pour chaque société étudiée, le pourcentage de titres détenus en 1965 par d'autres sociétés, autres que les petites entreprises et les porteurs institutionnels, et qui en sont les principales détentrices, ceci afin de construire la caractéristique MV 1.

Toutes ces données ont été mises sur cartes afin de permettre un traitement sur ordinateur. Celui-ci a été effectué sur la CDC 3 600 du Centre Blaise-Pascal.

Nous en exposons maintenant la méthode.

### 3. MÉTHODE DE TRAITEMENT

Une fois le fichier constitué, nous avons effectué sur ordinateur le traitement suivant :

- Calcul pour chaque titre des caractéristiques définies précédemment à partir des données mises sur cartes.
- Analyse factorielle en composantes principales de la matrice de corrélations des rangs de Spearman, chaque titre comptant pour une observation.
- Sortie de graphiques donnant la répartition des caractéristiques et des titres dans les plans des premiers facteurs et permettant de déceler les regroupements préférentiels des caractéristiques d'une part, de regrouper les titres en classes d'autre part.

Nous ferons quelques remarques sur l'analyse factorielle en composantes principales dont on trouvera un exposé dans le livre de T.W. Anderson cité en bibliographie.

Un aperçu intuitif de la méthode peut être formulé ainsi :

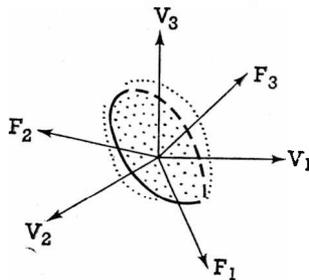
Chaque observation  $i$  étant repérée dans  $\mathbb{R}^J$  (\*) par l'ensemble  $(r^1, \dots, r^J)$  de ses  $J$  rangs en chacune des variables considérées  $V_1, \dots, V_J$ , il s'agit de trouver un nouveau repère de  $\mathbb{R}^J$   $(F_1, \dots, F_J)$  tel que la projection du nuage des observations sur l'axe  $F_1$  soit la plus étalée possible (ce qu'on exprime encore en disant que le pourcentage de la variance du nuage « expliqué » par  $F_1$  est le plus grand possible), puis celle sur l'axe  $F_2$  une fois  $F_1$  trouvé, etc...

Cette recherche se fait sous la contrainte suivante :  $F_1, \dots, F_J$  sont des combinaisons linéaires de  $V_1, \dots, V_J$ , de norme inclidienne 1 et telles que leurs corrélations deux à deux soient nulles.

On montre que le repère  $(F_1, \dots, F_J)$  est constitué d'axes orthogonaux et que la variance globale du nuage d'observations est égale à la somme des variances de ses projections sur les  $J$  nouveaux axes.

**FIGURE 1.**

**La projection du nuage sur le plan  $(F_1, F_2)$  est la plus étalée possible.**



Faisons maintenant deux remarques d'ordre général :

— Il est important de noter que le changement de repère se fait indépendamment des distributions des différentes variables prises en compte. L'analyse factorielle aux composantes principales présente l'avantage de ne point postuler de modèle pour la recherche des axes factoriels, ce qui lui ôte tout caractère suspect, surtout lorsqu'on sait combien les distributions réelles s'éloignent « quelquefois » des distributions théoriques dont on voudrait qu'elles soient proches.

— L'analyse postule que le coefficient de corrélation est un indicateur fiable de la liaison entre deux variables.

On peut d'abord objecter que cet indicateur est trop synthétique et qu'il ne résume qu'imparfaitement l'information contenue dans deux séries statistiques. Nous pouvons cependant le retenir dans la mesure où, pour deux variables  $X$  et  $Y$ , nous cherchons à savoir si, aux aléas près,  $Y$  croît lorsque  $X$  croît. Pourquoi avons-nous choisi le coefficient de corrélation des rangs et non le coefficient de corrélation linéaire ? D'abord, pour comparer

(\*) Remarque :  $\mathbb{R}^J$  est l'espace à  $J$  dimensions dans lequel chaque point a  $J$  coordonnées.

deux coefficients de corrélation linéaires,  $\text{cor}(X, Y)$  et  $\text{cor}(X, Z)$  par exemple, et dire que la liaison entre  $X$  et  $Y$  est plus forte que celle entre  $X$  et  $Z$ , il faut supposer que ces liaisons sont bien, aux aléas près, linéaires. Or il se peut fort bien, et quelques graphiques nous le prouvent, qu'en fait la liaison entre  $X$  et  $Y$  soit quadratique alors que celle entre  $X$  et  $Z$  est linéaire. En choisissant le coefficient de corrélation des rangs, nous perdons sans doute de l'information, mais une information difficile à analyser. Nous n'avons pas à faire ainsi l'hypothèse de linéarité. Ensuite, pour tester la significativité d'un coefficient de corrélation linéaire il faut supposer que les variables sont normales alors que nous exposerons un moyen simple de tester la significativité du coefficient de corrélation des rangs indépendamment de toute hypothèse sur les lois de distribution des variables. Enfin pour que les projections du nuage d'observations sur les plans des premiers facteurs  $(F_1, F_2)$  et  $(F_3, F_4)$  nous permettent d'induire d'une classification opérée sur elles une classification des points du nuage il faut montrer que ces plans ne sont pas quelconques au sens où ils auraient pu être obtenus à partir de données choisies au hasard, soit que le pourcentage de la variance du nuage « expliqué » par ces premiers facteurs est statistiquement significatif. Or il n'existe pas de tel test dans le cas où l'on part de valeurs quantifiées et donc de corrélations linéaires, alors qu'on peut construire un test non paramétrique simple dans le cas où les observations sont définies dans  $\mathbb{R}^J$  par leurs rangs en chacune des variables et que l'on analyse la matrice des corrélations des rangs de celles-ci.

Nous pensons avoir assez justifié ainsi l'utilisation de méthodes non paramétriques. Disons, pour terminer, quelle est l'utilité de la projection du nuage d'observations sur le plan des premiers facteurs  $(F_1, F_2)$  où il s'étale au maximum. Nous faisons une partition de cette projection du nuage fondée sur des relations de voisinage et non sur un système d'inégalités numériques. Aussi sera-t-elle approchée. L'esprit est malheureusement incapable d'opérer une telle partition dans  $\mathbb{R}^J$  ( $J > 3$ ) et c'est seulement en combinant les partitions que l'on peut opérer dans les plans  $(F_1, F_2)$ ,  $(F_3, F_4)$  que l'on arrive à regrouper les observations en classes.

Suivant la valeur de  $V_1, \dots, V_J$  en  $F_1, F_2, F_3, F_4$  et la disposition des classes d'observations dans les plans  $(F_1, F_2)$  et  $(F_3, F_4)$  on peut savoir lesquelles des variables  $V_1, \dots, V_J$  ont une forte ou une faible valeur relative pour les éléments de chacune de ces classes. Ainsi s'associe à une partition spatiale un ensemble de significations qui définit le critère de classification.

#### 4. PRÉSENTATION ET ANALYSE DES RÉSULTATS

##### A) Étude des relations entre variables prises deux à deux

Nous ne retiendrons dans ce paragraphe que les couples de variables dont le coefficient de corrélation est non seulement significativement différent de zéro mais encore fortement positif ( $> 0,50$ ) ou fortement négatif ( $< -0,50$ ) de telle sorte que nous pourrions inférer qu'une variable croît (décroît) lorsque l'autre croît, sans préjuger du type de croissance (linéaire, logarithmique, par paliers, etc...).

D'un point de vue strictement statistique comment peut-on conclure que le coefficient de corrélation de rangs  $r$  est significativement différent de zéro ? Pour un grand nombre d'observations  $N$ , et 162 est un tel grand

nombre, la variable  $t = r \sqrt{\frac{N-2}{1-r^2}}$  est distribuée comme un  $t$  de Student

(voir Siegel, 1956). Pour  $N = 162$ , on peut affirmer avec moins de 5 chances sur 1 000 d'être dans l'erreur que

$r$  est significativement positif si  $t = r \sqrt{\frac{N-2}{1-r^2}} > 2,580$  soit  $r > 0,20$ ,

$r$  est significativement négatif si  $t = |r| \sqrt{\frac{N-2}{1-r^2}} > 2,580$  soit  $r < -0,20$ .

Les coefficients  $r > 0,50$  ou  $r < -0,50$  sont donc forts et peuvent être interprétés en termes de croissance. Cependant, ils ne nous permettront pas d'affirmer qu'il y a un lien de causalité entre deux variables. Si un phénomène (I) est une condition de l'existence d'un phénomène (II) nous dirons que la variable (II) qui mesure ce dernier est une fonction de la variable (I) qui mesure le premier. Mais l'affirmation d'un lien de causalité entre ces phénomènes ne peut se faire qu'à la suite d'hypothèses sur leur mode d'interaction, éventuellement dépendant d'autres phénomènes.

Nous donnerons donc les premiers résultats suivants :

1<sup>o</sup> Le marché réel est une fonction croissante du marché potentiel.

$$MV 2 \nearrow MV 1$$

$$\text{Cor}(MV 1, MV 2) = 0,77$$

Ceci étant aisément prévisible. A contrario, il sera intéressant de caractériser les entreprises pour lesquelles  $MV 2$  est nettement inférieur à la moyenne des  $MV 2$  de celles des entreprises qui ont même  $MV 1$  qu'elles.

2<sup>o</sup> Le taux de fluctuation des cours décroît lorsque le taux de variation des cours décroît.

$$TC 2 \searrow TC 1$$

$$\text{Cor}(TC 1, TC 2) = -0,65$$

Il est délicat d'interpréter cette relation d'une part parce que le taux de fluctuations ne rend pas compte du dynamisme de celles-ci et ne permet pas de préjuger de leur durée ni de leur périodicité, d'autre part, parce que, la tendance générale de l'année 1965 étant à la baisse, on ne peut conclure pour les quelques cas de tendance à la hausse qui sont d'ailleurs de faible amplitude.

On peut affirmer que le taux de fluctuation décroît lorsque le taux de variation croît en restant négatif. Cependant le coefficient de corrélation des rangs entre taux de fluctuation et valeur absolue du taux de variation est de  $+0,84$ . Ceci montre que le taux de fluctuation croît lorsque l'écart relatif entre les cours de début et de fin d'année croît en valeur absolue, que la tendance soit à la hausse ou qu'elle soit à la baisse.

Cette relation est alors tout à fait prévisible. Car à supposer que la croissance ou la décroissance des cours soit régulière au long de l'année — ce qui est souvent le cas — les cours extrêmes se confondent, à la limite, avec les cours en début et en fin d'année (fig. 2) et la corrélation est alors parfaite.

En fait, nous nous trouvons, très probablement, dans le cas intermédiaire de la figure 3.

FIGURE 2

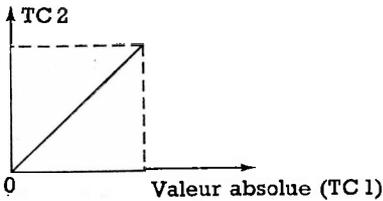
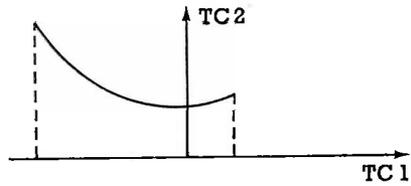


FIGURE 3



3° Le taux d'autofinancement décroît lorsque le taux d'investissement croît :

$$EI 2 \searrow EI 1$$

$$\text{Cor} (EI 1, EI 2) = -0,59$$

Cette relation est aisément interprétable. Elle signifie qu'à chiffre d'affaires égal, plus une société investit plus elle fait appel à l'extérieur pour financer ses investissements.

4° La marge bénéficiaire brute croît lorsque le taux d'investissement croît :

$$EI 4 \nearrow EI 1$$

$$\text{Cor} (EI 1, EI 4) = +0,63$$

Pour interpréter cette relation, il faut supposer que lorsque le taux d'investissement augmente, la productivité du travail augmente plus vite que les charges financières, ce qui entraîne des coûts salariaux unitaires plus faibles. Pour un même chiffre d'affaires le cash flow est alors plus élevé et la marge bénéficiaire brute l'est aussi.

5° La marge bénéficiaire brute décroît lorsque le taux de rotation des actifs croît :

$$EI 4 \searrow EI 5$$

$$\text{Cor} (EI 4, EI 5) = -0,74$$

Les sociétés ayant un fort taux de rotation des actifs ne sont pas celles qui ont une forte marge bénéficiaire brute et réciproquement. L'analyse factorielle permettra d'explicitier ce résultat.

6° Remarquons que la rentabilité par rapport aux fonds propres EI 7 a une corrélation positive mais faible (de l'ordre de 0,30) avec chacune des deux variables précédentes. Or l'on peut écrire :

$$EI 7 = EI 4 \times EI 5$$

relation qui montre qu'un même taux de rentabilité peut être obtenu avec une forte marge bénéficiaire et un faible taux de rotation des actifs ou, inversement, avec une faible marge et un fort taux de rotation. Nous verrons que les plus fortes rentabilités sont obtenues, pour les sociétés faisant partie de l'échantillon, dans le cas d'une forte marge et d'un faible taux de rotation.

## B) Résultats de l'analyse factorielle

Nous distinguerons trois paragraphes :

- 1) Test de la significativité des résultats.
- 2) Regroupements préférentiels des variables (caractéristiques).
- 3) Regroupements des observations (titres).

1) Avant de regrouper les variables et les observations en fonction des proximités de leurs projections sur les plans ( $F_1, F_2$ ) et ( $F_3, F_4$ ) nous devons montrer que cette opération est licite c'est-à-dire que les configurations dans ces plans résument celles de l'espace à 16 dimensions dont on est parti. Soit encore que ces plans sont préférentiels au sens où le hasard ne nous aurait pas, à lui seul, permis de les construire.

En termes statistiques, il faut pouvoir tester la significativité des pourcentages de variance globale du nuage d'observations expliqués par chacun des quatre premiers facteurs.

Nous avons fait le test non paramétrique suivant :

Nous générons seize séries de cent soixante-deux nombres au hasard, auxquels nous associons leur rang dans la série. Nous avons ainsi cent soixante-deux observations « au hasard » définies chacune par l'ensemble de ses rangs dans chacune des seize séries. Chaque série peut être considérée comme l'ensemble des rangs associés aux cent soixante-deux valeurs tirées d'une variable aléatoire uniforme (distribuée au hasard). Nous faisons alors l'analyse factorielle de la matrice de corrélation des rangs de ces seize variables.

Nous recommençons vingt fois cette opération.

Si alors le pourcentage de variance globale expliquée par le  $i$ -ème facteur dans l'analyse faite sur les données de l'étude est supérieur à chacun des pourcentages de variance expliqué par le  $i$ -ème facteur dans les analyses faites sur données générées au hasard, nous pouvons affirmer avec moins de 5 % de chances d'être dans l'erreur (1/20) que ce pourcentage est significatif et qu'il lui correspond une direction préférentielle du nuage d'observations.

Nous avons obtenu les résultats suivants :

	Données de l'étude	Données générées au hasard
	Pourcentage de variance expliquée	Pourcentage de variance expliquée maximum
	—	—
1 <sup>er</sup> facteur .....	18,12	10,72
2 <sup>e</sup> — .....	15,06	9,87
3 <sup>e</sup> — .....	13,75	8,92
4 <sup>e</sup> — .....	10,53	8,49
5 <sup>e</sup> — .....	9,40	8,00
6 <sup>e</sup> — .....	7,43	7,84

Au seuil considéré, nous pouvons conclure que les pourcentages de variance expliqués par les trois premiers facteurs sont nettement significatifs ; ceux expliqués par les deux suivants le sont encore mais moins nettement ; ceux expliqués par le sixième ainsi que les suivants ne le sont plus.

Les plans ( $F_1, F_2$ ) et ( $F_3, F_4$ ) sont donc des plans préférentiels. Nous pouvons y faire des regroupements en fonction des proximités qui résume-

ront la configuration, dans l'espace à seize dimensions, des observations ainsi que la matrice de corrélations des variables.

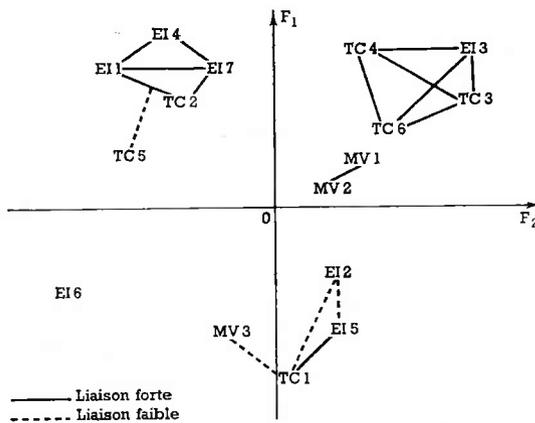
2) Les regroupements préférentiels de variables dans les plans  $(F_1, F_2)$  et  $(F_1, F_4)$  où chaque variable est repérée par ses coefficients de corrélation avec  $F_1$  et  $F_2$  d'une part, avec  $F_3$  et  $F_4$  d'autre part, vont nous permettre de construire un critère de classification.

Dans le plan  $(F_1, F_2)$  (voir schéma 1) nous pouvons regrouper en fonction de leur proximité :

- 1) TC 3 — TC 4 — TC 6 — EI 3
- 2) EI 1 — EI 4 — EI 7 — TC 2
- 3) MV 1 — MV 2
- 4) EI 5 — TC 1 — MV 3
- 5) EI 2
- 6) EI 6
- 7) TC 5

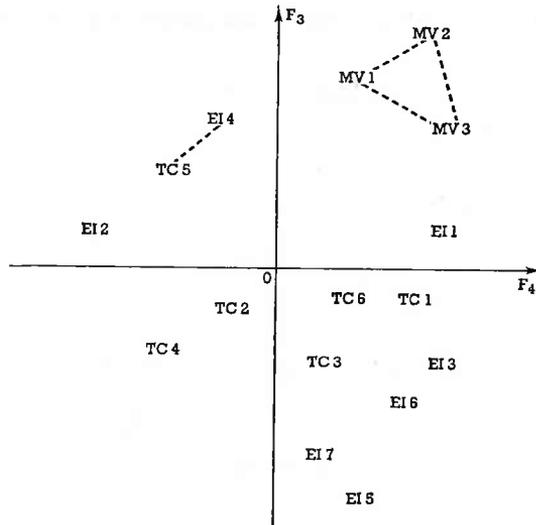
Le schéma 1 indique de plus la position relative des groupes :

SCHÉMA 1



Ce schéma autorise à tirer les conclusions suivantes, nuancées éventuellement par la prise en compte des proximités dans le plan ( $F_3, F_4$ ) (voir schéma 2) :

**SCHÉMA 2**



S'associent préférentiellement :

—	Caractéristiques boursières	Caractéristiques des sociétés
G <sub>1</sub>	Anticipation à la croissance Dividende Taux d'accroissement du dividende	Taux d'accroissement du chiffre d'affaires moyen
G <sub>2</sub>	Taux de variation des cours Rendement	Taux d'investissement Marge bénéficiaire brute Rentabilité par rapport aux fonds propres
G <sub>3</sub>	Marché réel Marché potentiel	
G <sub>4</sub>	Taux de transaction Taux de fluctuation des cours	Taux de rotation des actifs Taux d'autofinancement
G <sub>5</sub>		Pourcentage du bénéfice mis en réserve.

On peut alors opposer les groupes suivants parce que, dans le plan  $F_1, F_2$ ) ils sont symétriques par rapport à l'origine :

G<sub>1</sub> et G<sub>5</sub>

G<sub>2</sub> et G<sub>4</sub>

Ce qui veut dire qu'à de fortes valeurs des variables du groupe 1 (2) correspondent de faibles valeurs de celles du groupe 5 (4) et réciproquement.

On définit ensuite les critères de classification suivants qui tiennent compte à la fois des regroupements préférentiels et des oppositions entre groupes de variables :

(Remarque : lorsqu'on dira qu'un titre a une forte ou une faible valeur pour l'une quelconque des variables considérées il faudra comprendre que cette valeur est forte ou faible par rapport à l'ensemble des valeurs que prend cette variable pour les titres de l'échantillon. Il s'agit donc d'un point de vue ordinal et non pas quantitatif.)

Critère  $C_1$  : Titres dont le dividende est en forte progression (étant lui-même important et correspondant sans doute à un fort nominal) et émis par des sociétés en expansion (fort taux d'accroissement du chiffre d'affaires moyen) bien qu'elles ne mettent en réserve qu'une faible part de leur bénéfice.

Critère  $C_2$  : S'obtient par négation des propositions définissant  $C_1$  (en remplaçant fort par faible et vice versa).

Critère  $C_3$  : Titres de fort rendement, ayant acquis les plus fortes plus values en un an (on dira plutôt les plus faibles moins values) émis par des sociétés dynamiques (fort taux d'investissement) et rentables (forte rentabilité par rapport aux fonds propres) dont la marge bénéficiaire brute est importante alors que le taux d'autofinancement et le taux de rotation des actifs sont faibles.

Critère  $C_4$  : S'obtient par négation des propositions définissant  $C_3$ .

Nous noterons qu'une forte rentabilité correspond à une forte marge et à un faible taux de rotation des actifs, alors que l'inverse eût été également possible (faible marge, fort taux de rotation). Mais, de plus, elle correspond à un fort taux d'investissement et un assez fort marché réel alors qu'une faible rentabilité correspond à un fort taux d'autofinancement. Aussi peut-on supposer que les sociétés ayant une forte rentabilité ont une position de monopole et ne craignent pas de faire appel à l'extérieur pour financer leurs investissements, alors que celles dont la rentabilité est faible sont en situation de concurrence, ce qui les oblige à avoir une gestion « sévère » (fort taux de rotation, fort taux d'autofinancement).

Chacun des critères précédents définit une classe en compréhension. Cependant les positions relatives des groupes de variables ne nous autorisent pas à délimiter quatre classes disjointes de titres correspondant aux quatre cadrans du schéma 1. En effet, si un titre se trouve dans la partie centrale du plan ( $F_1, F_2$ ) ou à la frontière de deux cadrans, il nous faudra tenir compte de sa position dans le plan ( $F_3, F_4$ ) pour l'affecter à une classe donnée. Ceci nous amène à parler du regroupement des observations.

3) Nous pouvons regrouper les observations en classes correspondant à chacun des critères construits au paragraphe précédent dans la mesure où nous projetons simultanément les variables et les observations sur le plan ( $F_1, F_2$ ) (graphique 1). Cette projection simultanée est obtenue par un changement d'échelle :

— La valeur d'un facteur pour une observation est divisée par 2,8 fois l'écart-type de ce facteur, ce qui, en supposant que cette valeur ne dépasse que très rarement 2,8 écarts-type, la ramène dans l'intervalle  $[-1, +1]$ .

— La valeur d'un facteur pour une variable est transformée en coefficient de corrélation entre ce facteur et cette variable, coefficient compris dans

l'intervalle  $[-1, +1]$ , ce qui modifie légèrement la direction définie par cette variable par rapport aux axes factoriels, mais conserve les proximités entre variables.

Aussi une plus petite distance entre projections d'une variable et d'une observation ne signifie pas toujours nécessairement une plus forte valeur de cette variable pour cette observation, bien que l'affectation d'une observation à une classe correspondant à un critère de classification donné reste légitime. La délimitation des classes a tenu compte des positions des points frontières dans le plan  $(F_3, F_4)$  ; de là provient la non régularité de leurs frontières. D'ailleurs certains points ne peuvent être affectés à aucune classe dans la mesure où leur position dans le plan  $(F_3, F_4)$  ne permet pas de décider d'une telle affectation. Ce sont des titres isolés, moyens selon tous les points de vue considérés dans l'étude (points cerclés sur le graphique 1).

Nous n'analyserons pas chaque titre individuellement selon son appartenance à telle ou telle classe — nous appellerons celles-ci  $C_1, C_2, C_3, C_4$  comme les critères qui les définissent — mais nous comparerons ces classes d'après leur composition correspondant aux variables suivantes :

— Chiffre d'affaires.

— Secteur industriel.

— Nous avons distingué des sociétés dont le chiffre d'affaires est inférieur à 30 000 000 NF de celles dont le chiffre d'affaires est plus grand que ce chiffre.

Il apparaît que le pourcentage de titres correspondant à des « petits chiffres d'affaires » est plus fort dans les classes  $C_1$  et  $C_3$ .

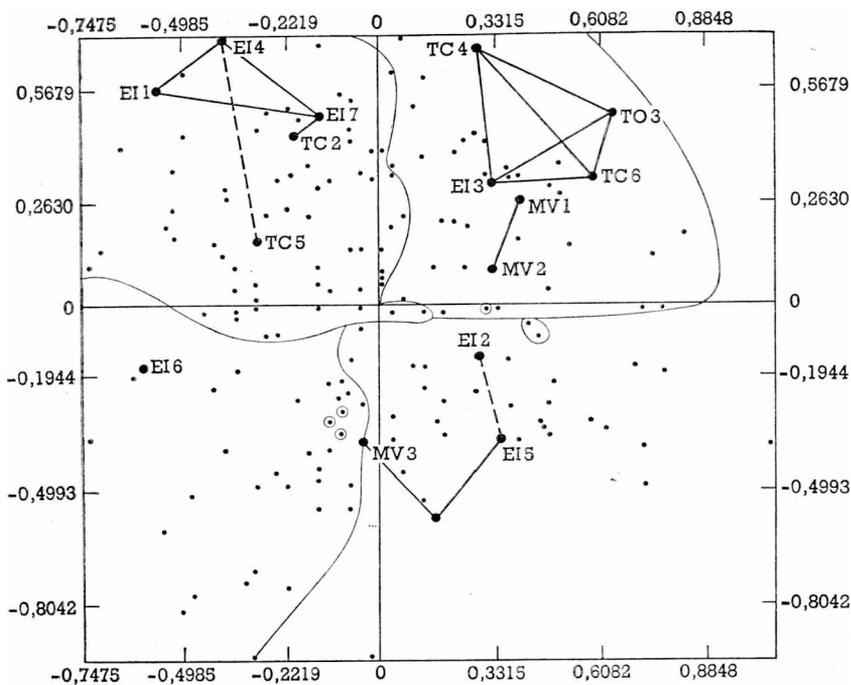
Classes	Pourcentage de « petits chiffres d'affaires »
$C_1$ .....	76
$C_2$ .....	9
$C_3$ .....	78
$C_4$ .....	36

Les sociétés faisant de gros chiffres d'affaires ont, relativement, plus tendance à s'autofinancer ou à mettre leurs bénéfices en réserve ce qui diminue le rendement de leurs titres et, d'autre part, sont celles dont les marchés potentiel et réel sont relativement moins importants bien que le taux de transaction soit relativement fort. Ce sont donc des sociétés où les participations d'autres sociétés sont les plus importantes.

Les sociétés dont le chiffre d'affaires est relativement faible ont un plus gros marché boursier — donc font plus appel aux capitaux extérieurs — mais semblent plus rentables, de leur point de vue, et de celui de la Bourse.

Ces conclusions n'expriment cependant qu'une tendance moyenne. On ne peut expliquer par les différences de chiffre d'affaires seules, les différences de performances des sociétés étudiées.

— L'échantillon des titres étudiés ne recouvre pas tous les secteurs industriels ; pour certains secteurs, la quantité de titres est trop faible pour conclure de façon significative. Aussi nous donnerons les chiffres partiels suivants :



Secteurs	Nombre de titres appartenant à :	
	C <sub>1</sub> ou C <sub>3</sub>	C <sub>2</sub> ou C <sub>4</sub>
Industries agricoles et alimentaires .....	10	4
Premières transformations et travail des métaux .....	6	1
Constructions électriques et électroniques .....	10	5
Industries chimiques et du caoutchouc .....	14	4
Extraction et fabrication des matériaux de construction ..	11	0
Services divers .....	1	10
Entreprises commerciales .....	3	10

Les cinq premiers secteurs sont les mieux classés parmi ceux considérés dans l'échantillon (titre appartenant en majorité à C<sub>1</sub> ou C<sub>3</sub>). Les derniers sont les moins bien classés (titres appartenant en majorité à C<sub>2</sub> ou C<sub>4</sub>).

Sans doute eût-il été possible de hiérarchiser les secteurs industriels si nous disposions d'un échantillon plus grand et de données pluri-annuelles ; et même de composer, par simulation, des portefeuilles bien diversifiés correspondant à des objectifs fixés. Cette étude peut donc se prolonger dans cette direction.

## CONCLUSIONS ET PERSPECTIVES

A ce stade de la recherche nous ne pouvons proposer un modèle causal du processus d'interaction entre la Bourse et l'économie. Nous pouvons suggérer, cependant, qu'un modèle stochastique peut être construit en partant de l'hypothèse générale suivante :

La politique de l'entreprise définie à la fin de l'année  $n - 1$  rendrait compte de ses résultats d'exercice de l'année  $n$  et de sa position en Bourse au cours de l'année  $n$ , dans une conjoncture générale de l'économie  $C_n$  au cours de l'année  $n$ . Cependant la politique de distribution des dividendes à la fin de l'année  $n$  et la politique économique-financière définie alors pour l'année  $n + 1$  seraient infléchies par l'évolution boursière au cours de l'année  $n$ . L'entreprise aurait donc une perspective à long terme qu'elle modulerait chaque année en fonction de sa plus ou moins grande dépendance du marché boursier. Il s'agirait alors de chercher quelles modulations choisies donnent les meilleurs résultats.

Le problème est complexe. De notre étude nous retiendrons surtout que les décisions d'investissement, donc le dynamisme de l'entreprise, semblent être un facteur très important d'explication de ses résultats tant économiques que boursiers (rentabilité, rendement, plus value) ; que l'expansion, mesurée par le taux d'accroissement moyen du chiffre d'affaires est source de meilleurs revenus des porteurs (croissance des dividendes). La bourse apparaît enfin comme un facteur très important de développement puisque ce sont les sociétés dynamiques et en expansion qui font le plus appel au marché boursier. L'autofinancement, s'il est source d'indépendance, semble caractériser les sociétés bien assises mais peu dynamiques !

Nous espérons pouvoir prolonger cette étude dans les deux directions suivantes à condition d'élargir l'échantillon et de mesurer les variables sur plusieurs années consécutives :

1. Étudier la composition de portefeuilles réels en titres appartenant à chacune des classes construites. La structure d'un portefeuille étant alors définie par les pourcentages du montant du portefeuille revenant à chaque classe, il s'agira de comparer cette structure au profil socio-économique du porteur.

2. Élaborer un modèle d'interaction de l'économie et de la Bourse.

### BIBLIOGRAPHIE

- Sur l'analyse factorielle en composantes principales et les statistiques non paramétriques.
- V. T. W. ANDERSON, *An Introduction of Multivariate Statistical Analysis* (John Wiley et Sons, Inc, 1958), pages 270 à 286).
- S. SIEGEL, *Non Parametric Statistic for the Behavioural Sciences* (Mc Graw Hill, 1956).
- L. LEBART, « Introduction à l'analyse des données » (2<sup>e</sup> partie, *Consommation*, n° 4, 1969).
- Sur les finances de l'entreprise et la Bourse :
- G. ANGÉNIEUX, *Les ratios dans l'entreprise* (Dunod, 1960).
- G. DEPALLEUS, *Gestion financière de l'entreprise* (Sirey, 1967).
- GRAHAM, DODD et GOTTLE, *Security Analysis. Principles and Techniques* (Mc Graw Hill, 1962).