Analyse de la structure des marchés et étude du concept de concurrence entre les marques : une application sur données de panel

Philippe Aurier

▶ **To cite this version:**

Philippe Aurier. Analyse de la structure des marchés et étude du concept de concurrence entre les marques : une application sur données de panel. Gestion et management. HEC PARIS, 1990. Français. NNT : 1990EHEC0010 . pastel-00994956

HAL Id: pastel-00994956
https://pastel.hal.science/pastel-00994956

Submitted on 22 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

141 552

015490

ECOLE DES HAUTES ETUDES COMMERCIALES
JOUY-EN-JOSAS

ANALYSE DE LA STRUCTURE DES MARCHES ET ETUDE DU CONCEPT DE CONCURRENCE ENTRE LES MARQUES : UNE APPLICATION SUR DONNEES DE PANEL

THESE

POUR LE DOCTORAT ES SCIENCES DE GESTION
conforme au nouveau régime defini par l'arrêté du 23 novembre 1988

présentée et soutenue publiquement

le 2 octobre 1990

par

Philippe AURIER

Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

JURY

Président du Jury
Jean-Louis CHANDON, *rapporteur*
Professeur à l'Université d'Aix-Marseille III

Directeur de Recherche
Yves EVRARD
Professeur HEC

Gilles LAURENT, *rapporteur*
Professeur au Groupe HEC

Bernard PRAS
Professeur à l'Université Paris IX Dauphine

Pierre-Louis DUBOIS
Professeur à l'Université de Montpellier II

Bruno COLIN
A.C. Nielsen

A ma soeur Catherine,
...à mes Parents.

Remerciements

J'exprime ma profonde reconnaissance à Monsieur Yves Evrard, Professeur au Groupe HEC, qui m'a conseillé et encouragé tout au long de ce travail.

Mes remerciements les plus sincères vont à Messieurs Jean-Louis Chandon, Professeur à l'Université d'Aix-Marseille III et Gilles Laurent, Professeur au Groupe HEC, qui ont accepté d'examiner mon travail et de participer à ce jury.

Je remercie chaleureusement Messieurs Pierre-Louis Dubois, Professeur à l'Université de Montpellier II et Bernard Pras, Professeur à l'Université de Paris IX Dauphine, pour leur participation au jury.

Je tiens à remercier particulièrement Messieurs Alain Pioche, Directeur du Développement Marketing à la Société Nielsen et Bruno Colin, Directeur de Projet Single Source à la Société Nielsen, qui m'ont permis d'effectuer la partie appliquée de cette recherche en me donnant accès aux données du Panel Scan 5000 ainsi qu'aux moyens informatiques de la société. Je leur exprime toute ma gratitude. Je remercie Monsieur Bruno Colin pour sa participation à ce jury.

Je remercie Jean-Louis Turrière, directeur des études à l'ESC Montpellier pour les facilités matérielles qu'il a su m'accorder tout au long de ce travail.

J'exprime ma plus profonde reconnaissance à mes amis Isabelle Arnal et Michel Aslanian, de la Société Setia, qui ont mis à ma disposition les matériels informatiques et m'ont aidé lors de la mise en forme de ce travail.

Ma reconnaissance va aussi aux ingénieurs du CNUSC à Montpellier, qui m'ont conseillé dans la partie informatique de cette recherche.

Enfin, j'exprime toute ma sympathie à mes amis qui m'ont encouragé tout au long de ce travail.

0154 490

Analyse de la Structure des Marchés et Etude du Concept de Concurrence entre les Marques: une Application sur Données de Panels.

Résumé: Une des phases essentielles dans l'analyse de la structure d'un marché est d'établir une mesure des relations entre les marques. Cette recherche a pour objectif d'étudier la validité de plusieurs mesures de la relation entre marques qui ont été proposées dans la littérature.

La recherche comporte deux parties. Dans la première, après avoir proposé une méthodologie pour l'analyse de la structure des marchés, nous présentons les concepts qui sont liés indirectement (Situation d'Usage, Recherche de Variété, Modèles de choix) ou directement (Substituabilité, Concurrence, Définition des Marchés) au domaine de l'analyse de la structure des marchés. Nous présentons ensuite un panorama exhaustif des modèles d'analyse des relations entre marques, modèles qui procèdent à partir de données de jugement ou données de comportement des consommateurs.

Dans la deuxième partie, nous nous centrons sur l'étude de la concurrence entre les marques à partir de données de comportement. La recherche est menée sur les données du panel SCAN 5000 de la société Nielsen, pour quatre marchés de produits de consommation fréquente. Nous considérons dans cette recherche que la concurrence entre deux marques est un concept théorique non-observable. et que, par concéquent, les mesures qui ont été proposées dans la littérature n'en sont que des opérationnalisations particulières. Pour apprécier la validité de ces mesures nous utilisons les techniques qui ont été développées en Psychométrie (la théorie de la mesure) et appliquées à la mesure de phénomènes psychologiques ou sociologiques. Nous proposons en outre une mesure de la concurrence permettant de prendre en compte la non-symétrie des relations de concurrence et possédant les proriétés d'une distance. D'autre part, nous étudions l'influence de plusieurs variables (Revenu, CSP, Quantité Consommée, Taille du Foyer, Point de Vente, Période d'observation) sur la mesure de la concurrence.

Mots Clefs: Analyse de la Structure des Marchés, Substituabilité, Concurrence, Positionnement, Théorie de la Mesure.

ANALYSE DE LA STRUCTURE DES MARCHES ET ETUDE DU CONCEPT DE CONCURRENCE ENTRE LES MARQUES, UNE APPLICATION SUR DONNEES DE PANELS

PREMIERE PARTIE ANALYSE DE LA STRUCTURE DES MARCHES: LE POINT SUR LES ENJEUX THEORIQUES

INTRODUCTION GENERALE 1

IMPORTANCE DE LA DEFINITION ET DE LA STRUCTURE DES MARCHES POUR LE MANAGER 3
DIFFERENTES APPROCHES POUR L'ANALYSE DE LA STRUCTURE DES MARCHES 7
METHODOLOGIE D'ANALYSE DE LA STRUCTURE DES MARCHES 12
PLAN DE LA RECHERCHE

CHAPITRE I: LES CONCEPTS FONDAMENTAUX 17

SECTION 1 LA SITUATION D'USAGE 18

SS1 LES VARIABLES SITUATIONNELLES DANS LA LITTERATURE MARKETING 18
SS2 LE CHAMP COUVERT PAR LES VARIABLES DE SITUATION 21
SS2-1 Les variables de situation ayant un effet direct sur le comportement la Situation d'Achat 21
SS2-2 Les variables de situation ayant un effet indirect sur le comportement 21

SS3 SITUATION D'USAGE ET ANALYSE DE LA STRUCTURE DES MARCHES 24
SS3-1 Effet de la situation d'usage sur la mesure de la concurrence entre les marques 25
SS3-2 Effet de la situation d'usage sur la structure des marchés 26
SS4 PRISE EN COMPTE DE LA SITUATION D'USAGE DANS LES MODELES D'ANALYSE DE LA STRUCTURE DES MARCHES 26
SS4-1 Prise en compte directe de la situation d'usage 27
SS4-2 Prise en compte indirecte de la situation d'usage 29

Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

SECTION 2 LA RECHERCHE DE VARIETE 30

SS1 RECHERCHE DE VARIETE : PRESENTATION ET DEFINITIONS 30
SS1-1 Approche stochastique 32
SS1-1-1 Modèles d'ordre 0 32
SS1-1-2 Modèles d'ordre 1 33
SS1-1-3 Modèles d'ordre 2 33
SS1-2 Approche déterministe 34
SS1-2-1 Les facteurs dérivés 34
SS1-2-2 Les facteurs directs 35
SS2 MESURE DE LA RECHERCHE DE VARIETE 37
SS2-1 Modèles stochastiques 37
SS2-2 Modèles déterministes 39
SS2-2-1 Mesure directe 39
SS2-2-2 Mesure hybride 40
SS3 IMPLICATIONS POUR L'ANALYSE DE LA STRUCTURE DES MARCHES 42

SECTION 3 THEORIE DES CHOIX DES CONSOMMATEURS 46

SS1 MODELES STOCHASTIQUES DE CHOIX 47
SS1-1 Modèles d'ordre 0 48
SS1-2 Modèles de Markov à deux états 48
SS1-3 Modèle d'apprentisage 49
SS2 MODELES DE CHOIX PROBABILISTES 52
SS3 MODELES DE CHOIX DETERMINISTES 55

SECTION 4 SIMILARITE SUBSTITUABILITE CONCURRENCE 62

SS1 LES CONCEPTS DE SIMILARITE SUBSTITUABILITE CONCURRENCE DANS LA LITTERATURE 62
SS1-1 Similarité 62
SS1-2 Substituabilité 64
SS1-3 Concurrence 65
SS2 NIVEAUX D'ANALYSE DES CONCEPTS DE SIMILARITE SUBSTITUABILITE ET CONCURRENCE 65
SS3 COMPARAISON ENTRE LES CONCEPTS 67
SS3-1 Similarité et concurrence 67
SS3-2 Similarité et substituabilitér 69
SS3-3 Substituabilité et concurrence 72

SECTION 5 LA DEFINITION DES MARCHES 74

SS1 DEFINITION DES MARCHES: DIFFERENTS POINTS DE VUE 74
SS1-1 Approche économique 74
SS1-2 Approche management 75
SS1-2-1 Définition à partir du produit 75
SS1-2-2 Définition orientée vers le consommateur 75
SS1-2-2-1 Définition du marché par rapport aux unités qui achètent 76
SS2-1-2-2 Définition du marché par rapport aux produits (unités achetées) 76
SS1-2-2-3 Dédinition du marché par rapport aux produits et aux consommateurs 77
SS2 L'IMPORTANCE DE L'UNIVERS CONCURRENTIEL 78
SS2-1 Etude de la modification des relations de concurrence due à l'introduction (ou au retrait) d'un produit (marque) dans le champ concurrentiel 79
SS2-2 Effet de la modification du nombre d'alternatives étudiées sur les résultats de l'analyse statistique 82
SS3 METHODES DE DEFINITION DES MARCHES 83
SS3-1 Méthodes procédant à partir de données de comportement 83
SS3-2Méthodes procédant à partir de données de jugement 84

CHAPITRE II: LA MESURE DE LA RELATION ENTRE LES MARQUES 84

SECTION 1 MESURE DE LA CONCURRENCE PAR LES ELASTICITES CROISEES 91

SS1 ESTIMATION DES ELASTICITES EN COUPES INTANTANEES (le modèle de Hagerty- Carmn et Russell 1988) 93
SS2 ESTIMATION DES ELASTICITES A PARTIR DE DONNEES LONGITUDINALES 94
(les modèles de Vanhoncker 1984, Cooper 1988)
SS3 ESTIMATION DES ELASTICITES A PARTIR DES FONCTIONS D'UTILITE DES CONSOMMATEURS 97
(les modèles de Kamakura et Russell 1989, Bucklin et Srinivasan 1989, Allenby 1988)
SS4 MODELE D'ANALYSE DE LA STRUCTURE DES MARCHES DERIVE DU CONCEPT D'ELASTICITE CROISEE 99
(extension du modèle DEFENDER, Shugan 1987)

SECTION 2 INDICES CONSTRUITS A PARTIR DES TRANSFERTS ENTRE MARQUES 104

SS1 PROBLEMES METHODOLOGIQUES LIES A L'UTILISATION DES HISTORIQUES D'ACHATS DES CONSOMMATEURS 104
SS2 INDICES DETERMINISTES 106
(Indices de Lehmann 1972, Totten et Elrod 1988)
SS3 INDICES STOCHASTIQUES 109
SS3-1 Indices construits à partir de modèles de choix d'ordre 0 109
SS3-1-1 Le modèle de Hendry 109
SS3-1-1-1 Constante de changement 110
SS3-1-1-2 Structure théorique 111
SS3-1-1-3 Indices théoriques dans le modèle Hendry 113
SS3-1-1-4 Indices de concurrence du modèle Hendry 114
SS3-1-1-5 Développements récents du modèle Hendry 115
SS3-1-2 Le modèle de Rao et Sabalava 116
SS3-1-3 Segmentation des consommateurs et analyse de la structure des marchés 118
Modèle de Grover et Dillon 118
Modèle de Grover et Srinivasan 122
SS3-1-4 Modèle de concurrence expliqué par les variables de contrôle du mix 123
SS3-2 Indices construits a partir de modèles de choix d'ordre 1: Prise en compte de la recherche de variété 127
SS3-2-1 Le modèle de Givon 128
SS3-2-2 Le modèle de Lattin et McAlister 133

SECTION 3 AUTRES APPROCHES DE LA MESURE DE LA CONCURRENCE UTILISANT DES DONNEES D'HISTORIQUES D'ACHATS 140

SS1 UTILISATION DES PARTS D'ACHATS COMME MESURE DE LA CONCURRENCE (Merunka et Bourgeat) 140

SS2 UTILISATION DES PERIODES INTER-ACHATS COMME MESURE DE LA CONCURRENCE (Modèle d'incidence: Fraser et Bradford) 141

SECTION 4 ETUDE DE LA RELATION ENTRE LES MARQUES A PARTIR DE DONNEES DE JUGEMENT DES CONSOMMATEURS 147

SS1 MESURES DE LA RELATION ENTRE LES MARQUES A PARTIR DES PERCEPTIONS 148

SS1-1 Mesures procédant à partir des similarités perçues 148

SS1-2 Utilisation ded adéquations perçues entre produits et usages 149

SS2 MESURE DE LA RELATION ENTRE LES MARQUES A PARTIR DES ENSEMBLES EVOQUES (Chandon et Strazziéri) 152

SS3 MESURE DE LA RELATION ENTRE LES MARQUES A PARTIR DES CHOIX : LE MODELE PRODEGY 155

TABLEAUX RECAPITULATIFS DES APPROCHES UTILLISANT LES COMPORTEMENTS OU LES JUGEMENTS DES CONSOMMATEURS 160

SECTION 5 MODELES DE PREFERENCES ET MESURE DE LA RELATION ENTRE LES MARQUES 165

SS1 MODELES DE PREFERENCES PROCEDANT A PARTIR DE DONNEES DE JUGEMENT 165

SS1-1 Les modèles de préférences probabilistes 167

SS1-2 Utilisation des modèles d'utilité dans les modèles de préférences 169

SS2 MODELES DE PREFERENCES PROCEDANT A PARTIR DE DONNEES DE CHOIX 170

SS2-1 Modélisation des choix multiples 173

SS2-2 Modèles de préférences ne faisant pas l'hypothèse IIA 176

SS2-3 Le modèle Choice Map (Elrod 1988) 179

TABLEAU RECAPITULATIFS DES PRINCIPAUX MODELES DE CHOIX ET DE PREFERENCES

CONCLUSION DU CHAPITRE II 184

CHAPITRE III: ANALYSE DE LA STRUCTURE DES INDICES: ASPECTS METHODOLOGIQUES LIES A LA REPRESENTATION DE LA STRUCTURE DES INDICES 187

SECTION 1 METHODES CONFIRMATOIRES ET EXPLORATOIRES 188

SS1 APPROCHES CONFIRMATOIRES 188
SS1-1 Confirmation d'indices 188
SS1-2 Confirmation de structures d'indices 188
SS2 APPROCHES EXPLORATOIRES 191

SECTION 2 PRISE EN COMPTE DE LA NON-SYMETRIE DES RELATIONS ENTRE MARQUES 192

SS1 TRAITEMENT DE LA NON-SYMETRIE PAR SEPARATION DES LIGNES ET DES COLONNES DE LA MATRICE D'INDICES EN DEUX ENTITES DISTINCTES 193
SS2 ANALYSE SPATIALE DES MATRICES ASSYMETRIQUES 194
SS3 ANALYSE NON SPATIALE DES MATRICES ASYMETRIQUES 196

SECTION 3 ANALYSE DES MATRICES TERNAIRES D'INDICES 200

SECTION 4 DEVELOPPEMENTS RECENTS EN MATIERE DE CONSTRUCTION DE PARTITIONS ET DE TRAITEMENT DES DONNEESQUALITATIVES: APPLICATIONS A L'ANALYSE DE LA STRUCTURE DES MARCHES 202

SS1 NOUVELLES METHODES D'ETUDE DE LA PARTITION DES MARCHES 202
SS2 LE TRAITEMENT DES DONNEES QUALITATIVES 205

CONCLUSION DE LA PREMIERE PARTIE 208

DEUXIEME PARTIE: ETUDE DE LA VALIDITE DU CONCEPT DE CONCURRENCE ENTRE LES MARQUES ET DES VARIABLES QUI INFLUENCENT CE CONCEPT

INTRODUCTION DE LA DEUXIEME PARTIE 214

LE CHOIX D'UN DOMAINE DE RECHERCHE EMPIRIQUE 214
LES OBJECTIFS DE LA RECHERCHE 216
PLAN DE LA RECHERCHE 219

CHAPITRE IV PRESENTATION DE LA RECHERCHE 220

SECTION 1 LE PANEL SCAN 5000 221

SS1 ESPRIT ET ENVIRONNEMENT DU PANEL SCAN 5000 221
SS2 DESCRIPTION DU CONTENU DE SCAN 5000 225

SECTION 2 MARCHES MARQUES ET FACTEURS MANIPULES 228

SS1 LE CHOIX DES MARCHES 228
SS2 LE CHOIX DES MARQUES (UNITE D'ANALYSE) 230
SS2-1 Contraintes d'ordre théorique 230
SS2-2 Contraintes techniques 230
SS2-3 Contraintes de fiabilité statistique 230
SS2-4 Contraintes pratiques 231
SS3 LES VARIABLES MANIPULEES 231
SS3-1 Codification des variables quantitatives 233
SS3-2 Codification des variables qualitatives 234

SECTION 3 LES INDICES ETUDIES 236

SS1 INDICE AGREGE DETERMINISTE 236
SS2 INDICE INDIVIDUEL DETERMINISTE 237
SS3 INDICE AGREGE STOCHASTIQUE 239
SS4 INDICE INDIVIDUEL STOCHASTIQUE 241
SS5 INDICES CONSTRUIT A PARTIR DE PSEUDOS DISTANCES DU CHI2 244
SS6 INDICES DE TRANSFERTS BRUTS ENTRE MARQUES 249

SECTION 4 PRESENTATION DES CALCULS ET DE LEUR MISE EN FORME 251

SS1 CALCUL ET MISE EN FORME DES RESULTATS BRUTS 251
SS1-1 Extractions des bases et calculs 251
SS1-2 Mise en forme des résultats bruts 252
SS1-2-1 Définition du mode de comparaison de deux mesures 252
SS1-2-2 Définition de la similarité entre deux mesures 253
SS2 CREATION ET DESCRIPTION DU CONTENU DES FICHIERS DE CORRELATIONS 253
SS3 CLASSIFICATION DES CORRELATIONS ANALYSEES 256
SS4 ANALYSE DESCRIPTIVE DU CONTENU DES FICHIERS DE CORRELATIONS 258
SS4-1 Etude des tendances centrales des corrélations 259
SS4-2 Etude de la distribution des corrélations 261

CHAPITRE V ETUDE DE LA VALIDITE DES INDICES 264

SECTION 1 MESURE ET VALIDITE: ELEMENTS FONDAMENTAUX 266

SS1 LE PROBLEME DE LA MESURE DES CONCEPTS 266
SS2 THEORIE DE LA MESURE EN SCIENCE SOCIALE 268
SS3 FIABILITE DES MESURES 269
SS3-1 Les différentes approches de la fiabilité 269
SS3-1-1 Approche classique 269
SS3-1-2 La théorie du domaine d'échantillonnage 270
SS3-1-3 Le modèle des tests parallèles 270
SS3-1-4 La théorie des facteurs 271
SS3-2 Mesure de la fiabilité 271
SS3-2-1 Mesure de la consistance interne 272
SS3-2-2 Méthodes des formes alternatives 272
SS3-2-3 Test-Retest 272
SS3-3 Généralisation du concept de fiabilité 273
SS4 VALIDITE DES MESURES 274
SS4-1 Etude de la validité interne 275
SS4-2 Validités convergente et discriminante 276
SS5 VALIDITE EXTERNE 279
SS5-1 Validité nomologique 279
SS5-2 Validité prédictive 279
SS5-3 Validité faciale 280

SECTION 2 FIABILITE DES MESURES ET ANALYSE DE LASTRUCTURE DES MARCHES 283

SS1 COHERENCE INTERNE DE LA MESURE DE LA CONCURRENCE 283
SS1-1 Etude des corrélations moyennes 283
SS1-2 Mesure de la cohérence interne avec lecoefficient alpha de Cronbach 287
SS1-3 Etude de la cohérence interne par l'analyse factorielle 292
SS2 ETUDE DE LA FIABILITE DANS LE TEMPS (TEST-RETEST) 298
CONCLUSIONS SUR L'ETUDE DE LA FIABILITE DE LA MESURE DE LA CONCURRENCE 302

SECTI0N 3 ETUDE DE LA VALIDITE DE TRAITS 306

SS1 APPROCHE CLASSIQUE DE LA VALIDITE DE TRAITS 306
SS1-1 Etude du critère de validité convergente 309
SS1-2 Etude du critère de validité discriminante 312
SS2 AUTRES APPROCHES DE LA VALIDITE DE TRAITS 314
SS2-1 Approche exploratoire 314
SS2-1-1 Utilisation de l'analyse factorielle 315
SS2-1-2 Utilisation jointe de l'analyse typologique et de l'analyse des similarités 317
SS2-2 Approche confirmatoire de la validité de traits 321
SS2-2-1 Analyse factorielle confirmatoire 321
SS2-2-2 Utilisation du modèle Lisrel 324

SECTION 4 ETUDE DE LA VALIDITE EXTERNE 331

SS1 PRESENTATION DU PROBLEME 331
SS1-1 Validité nomologique 331
SS1-2 Validité prédictive 331
SS1-3 Validité faciale 332
SS2 ETUDE DE LA VALIDITE FACIALE 332
SS2-1 Mise en oeuvre de l'étude 332
SS2-2 Qualité des représentations obtenues 334
SS2-3 Protocole d'évaluation des indices 338
SS2-4 Les résultats obtenus 339

SECTION 5 ETUDE DES FAMILLES D'INDICES 340

SS1 ETUDE DES FAMILLES D'INDICES A PARTIR DE L'ANALYSE DES SIMILARITES 342
SS2 ETUDE DES FAMILLES D'INDICES A PARTIR DE L'ANALYSE TYPOLOGIQUE 344
SS3 ETUDE DES FAMILLES D'INDICES A PARTIR DE L'ANALYSE FACTORIELLE 345

CONCLUSION DU CHAPITRE V 347
CONCLUSIONS RELATIVES A LA MESURE DU CONCEPT DE CONCURRENCE 347
CONCLUSIONS RELATIVES A LA METHODOLOGIE DES ETUDES DE VALIDITE 354

CHAPITRE SIX : ANALYSE DE L'EFFET DES FACTEURS 356

SECTION 1 LA MISE EN FORME DU PROBLEME 358

SS1 ETUDE A PARTIR DES DISTANCES A L'UNIVERS "GLOBAL" 358
SS2 ETUDE A PARTIR DES DISTANCES ENTRE UNIVERS ISSUS DE LA MANIPULATION DES FACTEURS 361
SECTION 2 ANALYSE COMPARATIVE DE L'EFFET DES FACTEURS 364

SS1 COMPARAISON DE L'EFFET DES FACTEURS A PARTIR DE LA SIMILARITE MOYENNE AVEC LA MODALITE "GLOBAL" 364
SS2 COMPARAISON DE L'EFFET DES FACTEURS A PARTIR DES SIMILARITES MOYENNES ENTRE LES UNIVERS 367

SECTION 3 : ANALYSE DE L'EFFET GLOBAL DES FACTEURS SUR CHACUN DES INDICES ETUDE DE LA SENSIBILITE DES INDICES 371
SS1 ETUDE DE L'EFFET GLOBAL DES FACTEURS SUR LES INDICES 371
SS2 ETUDE DE L'EFFET DE CHAQUE FACTEUR SUR CHAQUE INDICE 374

SECTION 4 ANALYSE COMPARATIVE DES EFFETS DE MESURE ET DE STRUCTURE 383

SS1 MISE EN FORME DU PROBLEME 383
SS2 RESULTATS 386
CONCLUSION DU CHAPITRE VI 390

CONCLUSION GENERALE 390

BIBLIOGRAPHIE 402

ANNEXES

ANNEXES A LA PREMIERE PARTIE
ANNEXES A LA DEUXIEME PARTIE

INTRODUCTION GENERALE

L'objectif de la stratégie d'une entreprise est souvent défini comme "l'obtention d'un avantage concurrentiel durable et défendable". Porter (1986) définit l'avantage concurrentiel comme procédant de la valeur qu'une firme peut créer pour ses clients. Cette valeur peut prendre trois formes différentes: un prix inférieur pour une prestation donnée, une prestation unique compensée par un sur-prix, une prestation permettant d'obtenir une utilisation plus économique du produit.

Les moyens d'obtenir un avantage concurrentiel ont été étudiés dans les disciplines de la Stratégie et du Marketing selon des optiques différentes. Dans les approches "production" qui sont le propre de l'analyse stratégique, l'obtention d'un avantage concurrentiel est décrit comme la "détention de compétences distinctives" aux différents niveaux du processus de production (approvisionnements, fabrication, contrôle de la qualité, distribution, ..). Dans les approches centrées sur le marché, qui sont le propre des analyses Marketing, l'obtention d'un avantage concurrentiel est conçu comme la détention d'une supériorité "de position", observée sur le marché.

Dans chacune de ces deux approches, des outils ont été développés pour diagnostiquer et mesurer les différentes formes que peut prendre l'avantage concurrentiel. Dans les approches "production", les outils les plus utilisés sont les concepts de prix relatifs, de parts de marché relatives ainsi que la chaîne de valeur qui constitue, selon Porter, un instrument fondamental de diagnostic de l'avantage concurrentiel. Ainsi, les chaînes de valeur de l'entreprise et des concurrents sont analysées et comparées. D'autres outils existent comme l'analyse en coupe instantanée des courbes d'expérience des différents concurrents ainsi que leur effet d'apprentissage. Les mesures de performance les plus utilisées sont la part de marché et les profits.

Parallèlement, dans les études centrées sur le marché, ont été développés des outils permettant de diagnostiquer des supériorités de position sur le marché. Le chercheur et le praticien disposent ainsi d'un ensemble d'outils dont l'objet principal est l'étude du consommateur. Ces outils permettent d'analyser le comportement du consommateur (modèles de choix, cartes de marché, élasticité des ventes aux actions Marketing, ..). Les mesures de performance les plus utilisées sont la satisfaction du consommateur, la fidélité, la notoriété, la part de marché relative.

Dans un récent article, Day et Windsley (1988) ont montré que ces deux types d'analyse sont individuellement insuffisantes et doivent être menées de façon conjointes. En effet, les approches "Consommateur" ne savent pas comment obtenir une compétence distinctive alors que les approches "Producteur" ne savent pas quelle compétence distinctive sera évaluée sur le marché. Il est donc nécessaire de mener une approche réconciliant ces deux points de vue.

L'analyse de la structure des marchés peut, dans une certaine mesure, aider l'entreprise dans cet objectif de réconciliation. En effet, les deux points de vue (producteur et

consommateur) possèdent en commun la nature de leur approche, qui est essentiellement comparative. Une chaîne de valeur peut être étudiée pour elle-même mais, il y a plus d'avantages à la comparer à celles de ses concurrents, de même pour l'analyse d'une courbe d'expérience ou d'une élasticité, etc..
L'avantage concurrentiel n'a de sens que "relativement" aux concurrents. Un des enjeux majeurs de la recherche d'un avantage concurrentiel est de savoir par rapport à qui il est intéressant d'obtenir un tel avantage.
L'analyse de la structure des marchés ainsi que les méthodologies qui l'accompagnent constituent un outil d'analyse qui aide le responsable dans ses décisions sur la manière :

- d'obtenir des compétences distinctives
- de savoir comment valoriser une compétence distinctive

En tant qu'instrument d'investigation, l'analyse de la structure des marchés permet de mieux comprendre le marché sur lequel l'entreprise se bat, de mieux connaître sa géographie, son découpage, la position des concurrents.
Une telle connaissance permet ensuite de faire des choix stratégiques quant aux marchés sur lesquels il est fondamental d'être présent et sur la manière d'allouer les ressources de l'entreprise au sein des différents sous-marchés. Il est nécessaire, à ce niveau, de réconcilier les approches en termes d'Unités Stratégiques de Base qui constituent souvent l'unité d'analyse en Stratégie et les sous-marchés qui constituent l'unité d'analyse en Stratégie Marketing.
D'autre part, il est important de préciser quelles sont les compétences distinctives spécifiques à mettre en oeuvre sur chacun des sous-marchés et comment les valoriser auprès du consommateur.

En tant qu'instrument de diagnostic, l'analyse de la structure des marchés permet, en mesurant la concurrence entre les marques, de mieux identifier ses principaux concurrents, de connaître la sensibilité de la demande aux actions de l'entreprise mais aussi à celles des concurrents.
En tant qu'instrument de contrôle, l'analyse de la structure des marchés permet de mesurer l'effet des actions qui ont été menées.

Pour toutes ces raisons, nous pensons que l'analyse de la structure des marchés constitue un moyen de mieux harmoniser les approches "producteur" et "consommateur", en cela, elle revêt une importance fondamentale. Depuis la fin des années soixante dix, le nombre d'articles de recherche consacrés à ce domaine est en augmentation permanente, augmentation qui s'est accélérée sur les cinq dernières années. Ce phénomène est une preuve du dynamisme qui caractérise ce pan de la recherche en Marketing.

Dans cette recherche, nous ne considérerons l'analyse de la structure des marchés que selon le seul point de vue de la démarche Marketing, cependant nous pensons que des efforts de recherche devraient être menés dans le sens indiqué précédemment, étant persuadés que les analyses de Day et Wensley sur la complémentarité des approches "producteur" et"consommateur" sont une voie de succès pour l'entreprise dans sa recherche de l'avantage concurrentiel.

L'analyse de la structure des marchés recouvre un ensemble de méthodes et de techniques dont l'objectif est de répondre aux trois questions que doit successivement se poser le responsable d'un marché ou d'un produit :

1 - Quel est le marché auquel appartient mon produit (marque) ? Il s'agit ici de la définition du marché, définition dans le sens où l'on va essayer de :

- lister les produits (marques) appartenant à ce marché
- lister les situations d'usage et de consommation de ces produits.
- connaître quels sont les consommateurs actuels et potentiels

C'est seulement alors qu'il sera possible de calculer des indices permettant d'évaluer les caractéristiques désirées du marché, à savoir :

- son potentiel : taille, croissance
- la pénétration et la vulnérabilité de l'entreprise par rapport à la concurrence
- les taux d'investissement, de profit, le risque caractérisant ce marché

En effet, la définition du marché permet de caractériser les éléments essentiels de l'univers concurrentiel :

- où l'entreprise se bat : son champ de bataille
- contre qui l'entreprise se bat : ses concurrents directs
- comment mener le combat et avec quelles armes : quels types de stratégies et d'actions peuvent être utilisées

Comme le notent A.D. Shocker et al (1987), la concurrence est un concept relatif. En terme de concurrence, une marque est caractérisée par son univers concurrentiel, c'est à dire, l'ensemble des marques avec lesquelles elle est confrontée de manière plus ou moins directe. La définition de cet univers est donc très importante .

2 - Comment le marché (une fois qu'en sont fixées les frontières) se structure-t-il en sous-marchés et selon quelle logique ? Ce découpage pourra être considéré comme le reflet du processus de choix du consommateur concernant les produits (marques) étudiés, mais d'autres logiques sont aussi envisageables . La connaissance de ce processus permettra de découvrir selon quelles caractéristiques du produit se détermine la concurrence. Il n'est donc nul besoin d'insister sur l'intérêt de connaître un tel processus Le principe généralement admis est que la concurrence entre des marques au sein d'un sous-marché particulier est plus forte qu'entre des marques appartenant à différents sous-marchés.

3 - Au sein d'un sous-marché particulier, quelles sont les relations de concurrence qu'entretiennent les produits (marques), est-il possible représenter graphiquement ces relations sous forme de carte ?

IMPORTANCE DE LA DEFINITION ET DE LA STRUCTURE DES MARCHES POUR LE MANAGEMENT

D'une manière générale, l'objectif est de cerner l'univers concurrentiel du produit de la firme et les relations qu'entretiennent ces produits entre eux. L'importance de cette analyse va se retrouver à tous les niveaux de la prise de décision dans l'entreprise.

1 AU NIVEAU STRATEGIQUE

- la définition du champ concurrentiel (champ de bataille : où l'entreprise se bat et contre qui ?)
- l'évaluation de l'avantage concurrentiel possédé par la firme au sein d'un segment stratégique. La détermination de cet avantage passe par la connaissance correcte :

- du segment stratégique auquel appartient le produit. La détermination de ce segment se fait en général à partir de critères orientés "production". Une approche Marketing complémentaire est pourtant instructive.
- de la part de marché relative, comme dans la matrice BCG (Boston Consulting Group) où celle-ci est une des deux dimensions de la matrice.

2 AU NIVEAU DE LA STRATEGIE MARKETING

L'analyse de la structure d'un marché est une étape préalable à la mise en oeuvre d'une stratégie de positionnement du produit (voir figure 1). En effet, ce n'est qu'une fois établis les scénarios d'achat et d'utilisation du produit qu'il sera possible de mener une étude de positionnement. La définition des scénarios de consommation nécessite la mise en oeuvre des étapes principales suivantes :

- définition du marché
- définition des situations d'usage
- segmentation des consommateurs
- représentation de manière fiable de la (ou des) structures existantes, en fonction de différents segments de consommateurs ou différentes situations d'usage

La stratégie pourra concerner :

- le positionnement d'un produit nouveau. En effet, l'analyse fine de la structure du marché permet de découvrir des opportunités de développement.
- le repositionnement d'un produit existant lorsque par exemple une Analyse de la Structure du Marché permet de saisir une modification de la structure qui nécessite une adaptation. Il est très important de noter que l'analyse de la structure des marchés marchés permet de contrôler que le positionnement préconisé est effectivement atteint.
- le positionnement (repositionnement) d'une gamme de produits. Dans la gestion de la gamme de produits, il est important de savoir comment "investir", "couvrir" par différents produits (marques) les différents sous-marchés. Par exemple il est important d'éviter d'avoir plusieurs produits qui se cannibalisent au sein d'un même sous-marché.

3 AU NIVEAU DES POLITIQUES DE MARKETING

Une bonne analyse de la structure d'un marché permet de répondre à des questions aussi courantes que :

- faut-il promouvoir deux marques de manière indépendante ou conjointe ?
- comment fixer nos prix par rapport à ce que font nos concurrents ? Quelles réactions attendre de leur part en réponse à nos propres actions ?
- comment et où placer nos articles en magasin (relation avec les distributeurs et les merchandisers) ?

Nous rajouterons à ce niveau plusieurs considérations relatives aux concepts de:

- Segmentation / Stratégie de Segmentation
- Positionnement / Stratégie de Positionnement
- Analyse de la Structure des Marchés

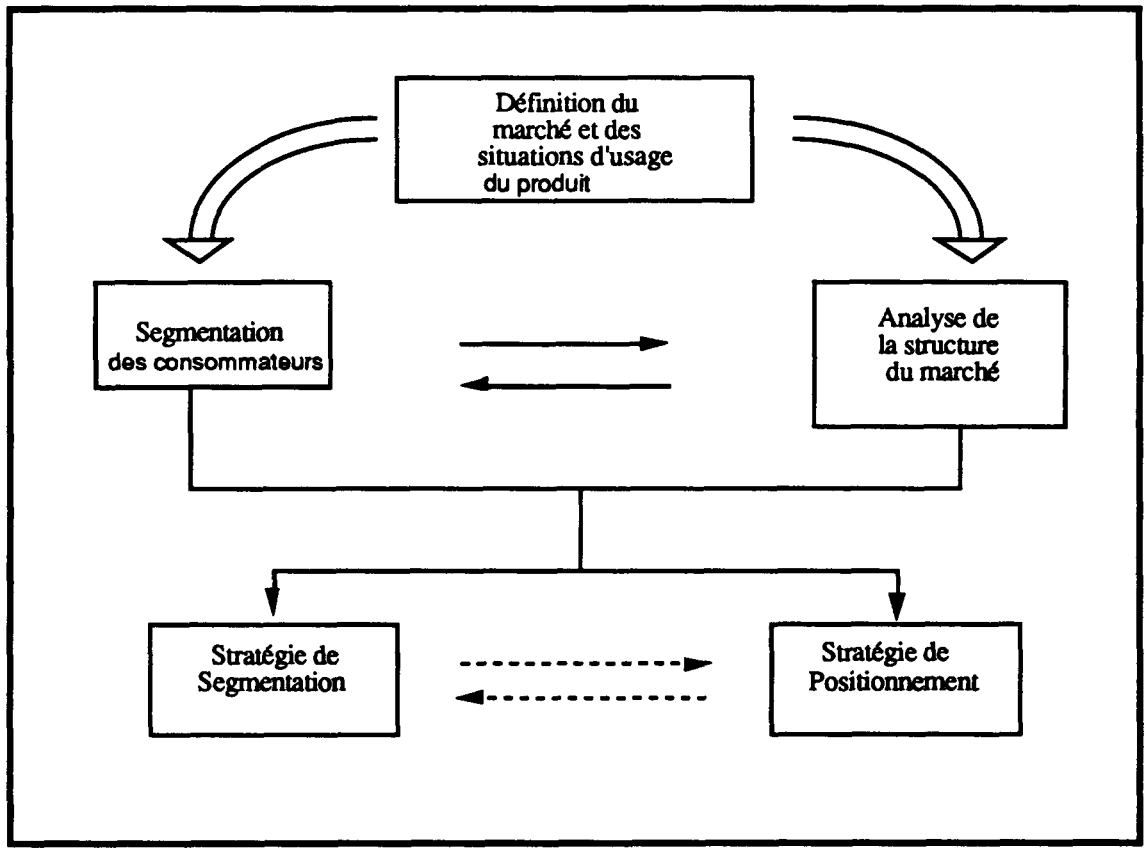
Il peut en effet y avoir des confusions à ce niveau. Le terme "segmentation" peut faire référence à deux concepts bien distincts :

- la Stratégie de Segmentation qui consiste à choisir un type de Marketing (différencié, concentré), une fois découpé le marché en sous-ensembles de consommateurs homogènes répondant de façon identique aux variables d'action Marketing.
- la Segmentation en tant que méthodologie au service de la mise en place d'une Stratégie de Segmentation et qui recouvre un ensemble de techniques (voir Boss 1973, 1974).

De la même manière, le concept de positionnement a donné lieu à bien des controverses (voir Dubois 1979) ; les auteurs n'étant pas d'accord sur la définition à donner à ce concept. Il faut en fait considérer que le même terme recouvre encore ici des significations différentes :

- le Positionnement en tant que stratégie de Marketing et qui peut être défini suivant Dubois (1977 p 87) comme "*une stratégie de marketing consistant à agir sur les croyances saillantes des consommateurs et prospects concernant l'achat du produit, pour qu'il se distingue clairement des concurrents et acquière une position spécifique dans leur esprit, favorisant son achat et sa consommation*".
- le Positionnement en tant que méthodologie d'étude des marchés, ce qui rejoint alors ce que nous appelons "Analyse de la Structure des Marchés". En effet, le succès croissant des stratégies de positionnement au sein des entreprises a conduit à la généralisation de la mise en oeuvre de méthodologies de plus en plus précises (voir Urban et Hauser 1980). A ce développement est liée la multiplication des méthodes et la nécessité du découpage de la méthodologie en plusieurs phases bien distinctes. Sur la figure 1, nous pouvons observer que l'analyse de la structure des marchés correspond à l'une de ces phases.

FIGURE 1 PRINCIPAUX CONCEPTS DES ETUDES PRODUITS



Ainsi, ce que certains auteurs appellent "Positionnement" correspond non pas à une définition de la stratégie de Positionnement mais à une des phases de la méthodologie qu'elle amène à mettre en oeuvre. La littérature marketing nous donne maintes preuves du flou existant autour des trois importants concepts dans les "politiques produits" que sont Segmentation, Positionnement et Analyse de la Structure des Marchés. Par exemple, Myers et Tauber (1977) regroupent sous l'appellation "Market Structure Analysis", des concepts aussi différents que la Segmentation et le Positionnement. Cette confusion est sans doute due à la définition même du concept de "marché", marché en terme de produits (Positionnement) ou de consommateurs (Segmentation). A notre avis, Segmentation et analyse de la structure des marchés sont à la fois :

a - Deux étapes préliminaires à la mise en oeuvre d'une stratégie de Positionnement

b - Deux étapes simultanées et indissociables puisque :

- la première phase dans une méthodologie de segmentation est la "définition" du marché à segmenter (voir Aurier 1989)
- la définition d'un marché peut être spécifique à un ou plusieurs segments de marché particuliers ou encore à une situation d'usage particulière (voir section suivante). Bien souvent, il sera nécessaire avant de segmenter et (ou) définir un marché, d'avoir caractérisé et contrôlé les situations d'usage du produit.

DIFFERENTES APPROCHES POUR L'ANALYSE DE LA STRUCTURE DES MARCHES

Il est possible de structurer les différents courants de littérature concernant le domaine selon plusieurs logiques que nous présentons successivement :

1 A PARTIR DU TYPE DE DEMARCHE UTILISEE

Deux démarches sont en effet possibles :

A - Formuler des hypothèses a priori sur la structure du marché. Ces hypothèses peuvent être fondées sur des études antérieures ou (et) l'expérience du responsable de marché. Une fois les hypothèses formulées, elles seront confrontées aux données du marché ce qui conduira à leur validation ou leur rejet.

Il s'agit là d'une démarche confirmatoire dans laquelle il est nécessaire de formuler des hypothèses et les tester ensuite. Ces hypothèses étant le fruit de l'expérience, elles possèdent une forte validité faciale et externe. Le piège est bien sûr d'être conduit à ne tester que les hypothèses évidentes et de ne pas tester la bonne hypothèse.

B - Collecter des données et ensuite les analyser de manière à caractériser la structure qu'elles sont censées refléter. Il s'agit ici d'une démarche exploratoire (voir Fornell, introduction, 1982 pour une comparaison de l'esprit des deux approches).

Le risque est ici de découvrir une structure ayant une bonne validité interne (expliquant bien les données) mais dont la validité externe pourrait être infirmée par des problèmes de sélection de l'échantillon, de fluctuations d'échantillonnage, d'erreur de mesure, d'utilisation d'un algorithme particulier de traitement.

2 A PARTIR DU TYPE DE DONNEES ANALYSEES

C'est en général le point de vue adopté pour structurer les différentes approches de l'étude de la structure des marchés (voir Shocker et al (1987), Fraser et Bradford (1983), Grover et Srinivasan (1987). Les méthodes sont présentées en fonction du type de données qu'elles utilisent (données de comportement ou données de jugement). Plus qu'une différence sur le type de données utilisé, il apparaît que les points de vue adoptés dans ces deux courants sont sensiblement différents, comme l'ont noté Day, Shocker et Srivastava (1979), Srivastava, Leone et Shocker (1981).

A LES DONNEES DE COMPORTEMENT

Leur caractéristique principale est d'étudier *"ce qui est"* (ce qui se passe sur le marché) par opposition à *"ce qui pourrait être"*. Le parti pris tient plus de la tactique que de la stratégie, les possibilités de prospective sont en effet limitées. En général, les données sont issues d'historiques d'achats des consommateurs, elles peuvent se présenter sous trois formes différentes :

- le "*relevé d'achats de la ménagère*" ("consumer diary panel data") qui est la forme la plus ancienne de recueil L'information porte sur ce que la ménagère a retranscrit de ses achats mais il n'existe aucune information sur ses non-achats.
- les "*données scanner consommateur*" ("consumer scanner data") qui fournissent l'historique exact, à partir du relevé aux caisses à lecture optique, des achats du ménage dans un magasin particulier. Là encore, il n'existe

aucune information sur les non-achats, c'est à dire sur l'offre au point de vente.

- les "*données de scanner magasin*" ("store level scanner data") dans lesquels sont disponibles, en plus des achats du ménage, l'ensemble des variables au point de vente (prix, promotion, présence, linéaire développé, part de marché magasin,...).

La faiblesse des deux derniers types de données est qu'elles ne comptabilisent que les achats du consommateur dans le seul magasin où sa carte est affiliée. C'est pour cela que de grandes sociétés d'études (Nielsen US, Secodip France) ont développé des mesures au niveau de la zone de chalandise: presque tous les achats du ménage dans la zone (on parlera alors de zone fermée) sont alors relevés.

Avec les progrès de la technologie, cette forme de collecte de l'information devrait se banaliser et sera ainsi de moins en moins coûteuse, de plus en plus fiable et donc de plus en plus utilisée. Les données de comportement fournissent une vision de la "réalité du marché", c'est à dire du comportement des consommateurs, une fois que les variables points de vente (disponibilité en rayon, visibilité, promotions...) ont été prises en compte. Les données collectées sont donc largement influencées par les variables de situation d'achat. Par exemple, pour que deux produits puissent être concurrents, il faut déjà qu'ils soient présents ensemble en magasin !

Un certain nombre de limites affectent ce type de données, à savoir (voir Kahn, Kalwani et Morrison 1986) :

- les produits doivent avoir une fréquence d'achat suffisante ainsi qu'un taux de diffusion (part de marché) suffisant, ce qui pose des problèmes pour étudier un produit nouveau. Le produit doit, de plus, être commercialisé par le canal de la grande distribution.
- ces données sont collectées sur une période plus ou moins longue, avec tous les inconvénients que cela comporte: modification des structures de préférences conduisant à une modification des choix .
- ces données se caractérisent souvent par leur "pauvreté" en contenu qualitatif (exemple : achat (1) / non achat (0)), ce qui peut rendre difficile l'interprétation des analyses et donner peu d'éléments pour la mise en place de politiques de Marketing.
- on ne sait pas dans le ménage "qui" a pris la décision d'achat et "qui" consomme. En effet, il n'est pas possible de savoir pour qui dans le ménage sont achetés les produits ni si différentes unités du ménage consomment des produits différents. C'est ainsi que deux produits (marques), achetés pour être consommés par différents individus, dans des conditions différentes, pourront être considérés à tort comme concurrents.
- les situations d'usage prévalant à l'achat n'étant ni connues, ni contrôlables, le raisonnement se fait au niveau agrégé des situations d'usage. Or, certains auteurs ont montré que la situation d'usage explique une part non négligeable de la concurrence entre les produits ou les marques.

Un autre problème lié à l'utilisation des données d'historiques d'achats est la confusion possible entre un effet de recherche de variété et un effet dû à la situation d'usage. Les différents effets peuvent venir se cumuler, rendant le problème très complexe. Comment différencier ces deux effets, sont ils réellement différents? Ceci pose un problème qui n'a pas encore été étudié dans la recherche Marketing. Or, la

différence entre ces deux concepts n'est pas toujours claire, nous verrons pour quelles raisons au chapitre premier.

- il n'est pas possible de cerner les modifications éventuelles des besoins et des attentes pouvant se traduire par une évolution des structures de préférences et donc des structures du marché. C'est, d'un point de vue stratégique, certainement la principale limite de ces données, puisque toute action stratégique (par exemple conception et positionnement d'un produit nouveau) devrait être basée sur une démarche prospective,ce que ne permettent pas de faire les données de comportement.

Pour finir, nous dirons que ces données décrivent la *"réalité du marché"*, par opposition à ce que nous appelons la "*réalité consommateur*" (voir figure 2, plus loin).

B LES DONNEES DE JUGEMENT

Ce sont des données issues d'enquêtes auprès de répondant. Elles peuvent avoir traits aux perceptions, attitudes, préférences, choix, ensembles évoqués, intentions d'achat, choix conditionnels...

Ces données, bien que plus difficiles à collecter (il faut effectuer des études ad hoc qui peuvent être très coûteuses) sont plus riches. Par rapport aux critiques formulées à l'encontre des données de comportement, elles présentent les avantages suivants :

- il y a identité entre répondants et acheteurs / consommateurs
- l'analyse n'est pas limité à des classes particulières de produits de par leur fréquence d'achat, leur taux de consommation ou de diffusion, hormis pour la sélection de l'échantillon qui peut être délicate lorsque le produit est faiblement diffusé.
- la variable situation d'usage peut être, non seulement contrôlée, mais aussi étudiée.
- il est possible de faire référence à un ensemble de variables (par exemple l'importance des attributs du produit, les situations d'usage envisagées) permettant d'apporter des éléments qualitatifs complémentaires lors de la mise en place d'une politique (notamment pour tout ce qui concerne la copy stratégie, le choix des circuits de distribution ..) .

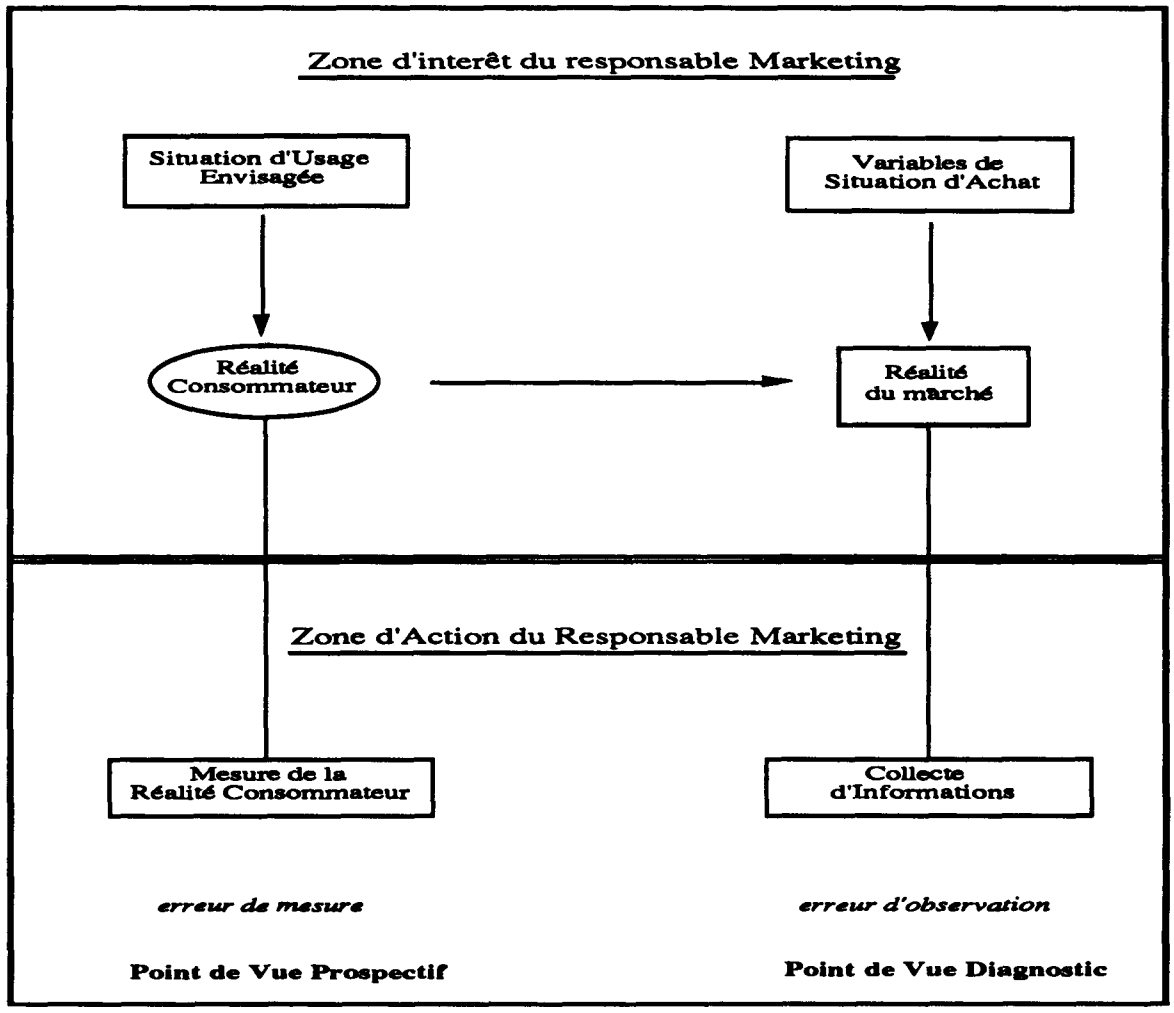
Les obstacles à leur utilisation de façon fiable sont liés :

- aux biais de réponse attachés à toute enquête par questionnaire (surtout lorsque le produit étudié est tel que le répondant ne s'exprime pas librement).
- à la nécessité d'opérationnaliser les variables étudiées qui ne sont que des concepts théoriques non observables et peuvent nécessiter d'effectuer des mesures multiples pour essayer d'éliminer les erreurs de mesure.

Il est souvent reproché à ce type de données de ne pas correspondre aux comportements réellement observés, notamment parce qu'elles ne prennent pas en compte les variables de situation d'achat (disponibilité du produit en magasin, promotions éventuelles, présence en rayon). Ceci est, à notre avis, peu justifié car ce qui est étudie à partir de ces données (sous réserve de l'élimination des erreurs de mesure) est la "*réalité consommateur*" (par opposition à ce que nous avons appelé précédemment "*réalité du marché*"). Or, il est important de noter que pour le chercheur ou le responsable marketing qui étudient le comportement du consommateur, c'est avant tout cette réalité qui est d'intérêt. C'est l'étude de cette réalité consommateur qui permettra d'anticiper les évolutions de la structure d'un marché (extension vers de

nouveaux produits, modification des situations d'usage, modification de la hiérarchie des critères de choix, etc...). La figure 2 suivante peut éclairer la problématique.

FIGURE 2 LA DOUBLE REALITE CONSOMMATEUR / MARCHE



Les concepts non observables sont entourés d'une ligne courbe

La zone d'intérêt du responsable marketing recouvre deux entités : le *"consommateur"* qui est l'unité où se prennent les décisions et le *"marché"* qui est le lieu où se matérialisent les décisions que prennent les consommateurs sur le lieu d'achat, c'est à dire sous l'influence des variables situationnelles. A la différence de la réalité du marché, la réalité consommateur n'est pas directement observable. Elle ne peut être que "mesurée", avec une erreur qu'il s'agit de minimiser et de contrôler. La réalité de marché, elle, peut être directement observée, ce qui n'exclut pas que des erreurs soient commises lors de son observation (erreur sur la retranscription des prix de ventes, des promotions en cours, des linéaires ..). L'observation de l'une ou de l'autre de ces réalités correspond à des objectifs qui peuvent se résumer de la manière suivante.

"*Qu'est-ce qui pourrait être* ?". C'est une optique essentiellement prospective qui nécessite d'étudier la réalité consommateur.

"*Qu'est-ce qui est* ?" . C'est alors une optique de diagnostic et de contrôle des actions passées.

Il n'est pas judicieux de qualifier un type de variable supérieur à l'autre et c'est avant tout le manager qui devra opter pour un point de vue particulier.

3- METHODES AGREGEES / INDIVIDUELLES

Les structures de marché peuvent être étudiées selon un point de vue individuel ou agrégé. En fait, trois types de modèles coexistent selon cette perspective.

A- LES MODELES INDIVIDUELS

Ici, les structures du marché sont estimées au plan individuel et éventuellement agrégées par la suite. Le modèle permet d'obtenir une structure par individu (ou segment) et une structure agrégée. Il faut encore distinguer à ce niveau entre deux catégories de méthodes :

- les méthodes qui estiment des structures individuelles de façon indépendante: une représentation de la structure pour chaque individu.
- les méthodes qui caractérisent les structures à partir de mesures de concurrence faites au plan individuel, mais qui analysent simultanément l'ensemble des données individuelles. Il s'agit donc de méthodes qui sont capables de traiter trois dimensions: l'individu et les deux dimensions liées au croisement des marques entre elles (les modèles d'analyse des similarités, Indscal et Alscal en sont des exemples, ils traitent des matrices stimuli x stimuli x individu).

B- LES MODELES AGREGES

L'information traitée est agrégée :

- soit parce que la forme même de l'indice n'est pas calculable au plan individuel
- soit parce que la méthode utilisée nécessite de traiter des indices agrégés puisqu'elle n'est pas capable de prendre en compte la dimension "*individu*" à cause de la lourdeur de la procédure d'estimation.

4 MODELES DYNAMIQUES / STATIQUES

Il existe peu de méthodologies faisant intervenir le temps dans l'analyse. Le plus souvent, le temps peut être pris en compte en faisant des études successives en coupes instantanées. Peu d'études nous ont par contre montré comment prendre en compte la dimension temps dans l'analyse. Ici aussi, il faudra vraisemblablement utiliser des méthodes permettant d'analyser trois dimensions comme par exemple des matrices produit x attribut x temps. Une rare tentative de modélisation du temps est celle de Cooper (1988) qui prend en compte simultanément l'individu, la marque et la période d'observation (voir chapitre III).

5 METHODES SPATIALES / NON SPATIALES

Comme le notent Shocker et al (1987), les méthodes peuvent être étudiées en fonction du type de résultats qu'elles offrent, à savoir des représentations spatiales ou non spatiales .

A REPRESENTATION SPATIALE

Dans l'hypothèse où la structure d'un marché serait le reflet du processus de décision du consommateur, cette distinction spatiale / non spatiale est alors très importante. En effet, une méthode spatiale où les produits sont représentés sur des continuum (les dimensions engendrant l'espace) peut caractériser :

- un processus de traitement de l'information de type compensatoire
- l'utilisation d'attributs de type continu (par exemple des échelles de différentiel sémantique)

B REPRESENTATION NON SPATIALE

Ce type de représentations caractérise un processus de traitement de l'information de type hiérarchique, disjonctif, dans lequel les produits (marques) sont évalués selon des attributs de type discret.

En fait, de la même manière qu'il existe des modèles de traitement de l'information à deux étapes (voir Bettman 1979, Pras et Summer 1975, Einhorn 1970, Gensh 1987), il est possible de concevoir une démarche analogue dans le domaine de l'analyse de la structure des marchés. La première partie du processus utilisé par le consommateur pour simplifier sa tâche et éviter de se mettre en état de surcharge informationnelle (voir Malhotra 1984, Jacoby 1984) est de type disjonctif. Elle correspond au processus de choix utilisé par le consommateur lorsqu'il envisage le marché de façon large. Une fois isolé un ensemble restreint d'alternatives, un modèle de type compensatoire pourrait être utilisé pour évaluer ces possibilités, ce qui serait alors naturellement représenté par un modèle d'analyse de la structure des marchés de type spatial.Urban, Johnson et Brunwick (1981) ont proposé une telle approche pour l'analyse de la structure des marchés :

- un modèle de type hiérarchique pour caractériser le marché et sa structure en sous-marchés
- un modèle de type spatial pour représenter les relations de concurrence au sein de chaque sous-marché

METHODOLOGIE D'ANALYSE DE LA STRUCTURE DES MARCHES

Nous allons présenter maintenant les phases principales que devrait comporter selon nous une méthodologie d'analyse de la structure d'un marché. Les phases de la méthodologie, notamment les points deux et trois, constituent la base sur laquelle s'articule cette première partie de notre recherche.

Avant de mettre en oeuvre une méthodologie d'analyse de la structure d'un marché il est indispensable d'avoir choisi le niveau d'analyse qui correspond le mieux aux objectifs fixés au préalable. Ceci consiste en fait à choisir entre l'étude de la "réalité

consommateur" ou l'étude de la "réalité du marché". Une fois que ce niveau d'analyse a été fixé, nous proposons une méthodologie comportant les trois phases suivantes :

1 LA DEFINITION DES FRONTIERES DU MARCHE

Le niveau d'analyse retenu précédemment aura beaucoup de conséquences sur la manière de définir le marché. Nous verrons en effet qu'un marché n'existe pas en tant qu'entité unique mais que ses frontières ainsi que la nature de ses unités d'analyse (produits, marques) dépendent du contexte dans lequel s'effectue l'analyse.

2 LE CHOIX D'UN INDICE DE MESURE DE LA CONCURRENCE ENTRE LES MARQUES

Ici encore, le niveau d'analyse retenu joue un rôle fondamental. Nous verrons que pour une analyse se plaçant au niveau de la réalité consommateur, il est nécessaire d'utiliser une mesure de substituabilité. Quand c'est la "réalité de marché" qui est étudiée, il convient alors d'utiliser une mesure de concurrence.

3 ANALYSE DES INDICES OBTENUS DANS LA PHASE PRECEDENTE

Cette phase permettra notamment :

- de caractériser la structure contenue dans les indices de concurrence (substituabilité) c'est à dire la structure du marché et son découpage en sous-marchés.
- de représenter, sous forme de cartes (représentations spatiales) ou de graphes (représentations non spatiales) la physionomie du marché et éventuellement des attentes des consommateurs.

Cette phase se caractérise essentiellement par l'utilisation de techniques d'analyse des données.

PLAN DE LA RECHERCHE

Notre recherche comporte deux grandes parties. La première est une présentation du champ théorique couvert par l'analyse de la structure des marchés et l'étude de la concurrence entre les marques. La deuxième partie est une étude empirique de la validité du concept de concurrence tel qu'il est opérationnalisé par les mesures proposées dans la littérature. Nous y étudions d'autre part l'influence de plusieurs variables (revenu, catégorie socio-professionnelle, taille du foyer, quantité consommée, point de vente, période d'observation). Nous présentons ici le détail de la première partie, celui de la deuxième étant présenté dans l'introduction de la deuxième partie.

Le premier chapitre sera consacré à l'étude du champ théorique couvert par l'Analyse de la Structure des marchés. Dans ce chapitre, nous préciserons et définirons les concepts qui ont, selon nous, une importance capitale pour la compréhension des enjeux théoriques et pratiques de l'analyse de la structure des marchés. Nous étudierons au travers de cinq sections :

(1) Les Variables Situationnelles et la Situation d'Usage
(2) La Recherche de Variété
(3) Les modèles de traitement de l'information et de choix des consommateurs
(4) Les concepts de Concurrence, Substituabilité et Complémentarité

(5) La définition des marchés.

Pourquoi avoir choisi ces cinq domaines particuliers comme soubassement théorique de l'étude de la structure des marchés ? Les points 1 et 2 (Situation d'Usage et Recherche de Variété) ont été choisis parmi d'autres variables importantes appartenant à la panoplie des variables étudiées en comportement du consommateur. Nous avons choisi d'étudier ces variables parce qu'elles constituent à notre avis un puissant modérateur des relations qu'entretiennent des produits ou des marques. La situation d'usage tout d'abord est une variable importante car elle influence fortement la relation du consommateur au produit. Il est difficile d'étudier cette relation sans introduire la dimension situation d'usage. L'idée est que les relations existant entre les produits d'un marché varient selon les situations d'usage envisagées par les consommateurs.

Pour des raisons similaires, nous avons choisi de présenter en détail le concept de Recherche de Variété. En effet, celle-ci peut influencer fortement les relations de concurrence entre les produits (ou les marques) et biaiser ainsi leur mesure. Il est donc important de bien saisir l'impact de cette variable sur les structures du marché étudié.

Nous avons choisi de consacrer la section 3 à la description des principales approches de la modélisation du traitement de l'information et du choix des consommateurs. En effet les modèles d'analyse de la concurrence ont souvent été construits en reprenant le cadre théorique de la modélisation du choix des consommateurs.

Dans la quatrième section, nous présenterons les concepts de base qui permettent d'opérationnaliser la mesure de la relation entre des produits ou des marques. Il s'agit des concepts de Similarité, Substituabilité et Concurrence, concepts qui sont souvent utilisés sans que leur contenu théorique et la manière de les opérationnaliser n'aient été précisés.

La cinquième section sera consacrée à la "Définition des marchés", ceci pour deux raisons essentielles :

- une définition correcte du marché doit être à la base de toute analyse de la structure d'un marché.et constituer une étape préliminaire aux modèles qui pourtant, dans leur grande majorité, considèrent cette étape comme résolue.
- il n'existe pas d'approche formalisée de la définition d'un marché, pas plus qu'il n'existe de définition vraiment opérationnelle d'un marché. A l'inverse, une fois que les frontières d'un marché d'un marché sont fixées, beaucoup de méthodes permettent de découper ce marché en sous-marchés. Il apparait donc plus facile (au plan méthodologique) de découper un marché en sous-marchés que de le définir. Ceci explique pourquoi nous avons choisi de présenter la définition des marchés en tant que concept de base puisqu'il s'agira à la fois de réfléchir sur la signification théorique d'un marché, mais aussi de présenter un certain nombre de techniques d'aide à la définition d'un marché.

Dans le deuxième chapitre, nous présenterons selon une logique méthodologique les principaux modèles d'Analyse de la Structure des Marchés et surtout le type de mesure de la concurrence qu'ils utilisent. Ce domaine jouit depuis quelques années d'un grand succès dans la recherche marketing. Il est donc nécessaire d'offrir un cadre conceptuel permettant de structurer la production foisonnante des dernières années en la matière. Dans la première section, nous présenterons l'utilisation des coefficients d'élasticité croisée comme mesure de la relation entre les marques. Dans la deuxième section, nous présenterons les approches utilisant le concept de "transfert entre marques" comme

mesure de la concurrence. Dans la troisième section, d'autres approches de la mesure de la concurrence à partir d'historiques d'achats seront présentées.

Dans la section quatre, nous exposerons les principales approches de la mesure des relations entre les marques utilisant des données de jugement des consommateurs.

Enfin, dans une dernière section, nous présenterons les approches de la mesure et de la représentation de la concurrence entre les marques qui procèdent à partir de modèles de préférences. Nous verrons que la particularité de ces approches est de ne pas calculer directement les indices de concurrence entre deux marques mais d'utiliser un concept intermédiaire, le point (vecteur) idéal. Nous verrons que ces modèles, de plus en plus, s'accommodent aussi bien de données de jugement (les préférences) que de données de comportement (les préférences révélées par les choix). C'est la raison pour laquelle nous présentons ces deux types de modèles au sein d'une même section.

Il est important de noter que ce deuxième chapitre consacré à l'étude de la mesure des relations entre marques constitue le coeur de notre recherche. La mesure de la concurrence entre les marques constitue en effet l'objet de la deuxième partie de la recherche qui se place, elle, à un niveau empirique.

Le troisième chapitre sera consacré à la présentation des méthodes utilisables pour analyser l'information contenue dans une matrice d'indices de concurrence.

Dans une première section nous opposerons les méthodes de type confirmatoire aux méthodes de type exploratoire. Dans une deuxième section nous évoquerons le problème lié à la non-symétrie des indices de concurrence (substituabilité), puisque cette non symétrie constitue selon nous une caractéristique fondamentale du phénomène concurrentiel. Dans une troisième section, nous présenterons quelques méthodes statistiques permettant d'élargir le cadre actuel d'analyse de la concurrence. Nous évoquerons notamment l'utilisation de méthodes permettant d'analyser des données "ternaires", comme c'est le cas quand pour un ensemble de matrices d'indices de concurrence calculés pour un ensemble de situations d'usage du produit.

Dans une quatrième section, nous présenterons de récentes méthodes d'analyse typologique permettant d'obtenir des partitions de recouvrement ou bien des partitions floues. Ce type de méthode présente un intérêt évident pour l'étude de la structure d'un marché qui n'a vraisemblablement rien d'une structure de partition forte dans laquelle chaque marque appartiendrait à un sous-marché et un seul. Nous évoquerons aussi l'intérêt qu'il y a à utiliser des méthodes permettant d'analyser directement des données de type binaire ou nominal (par exemple: achat / non-achat d'une marque).

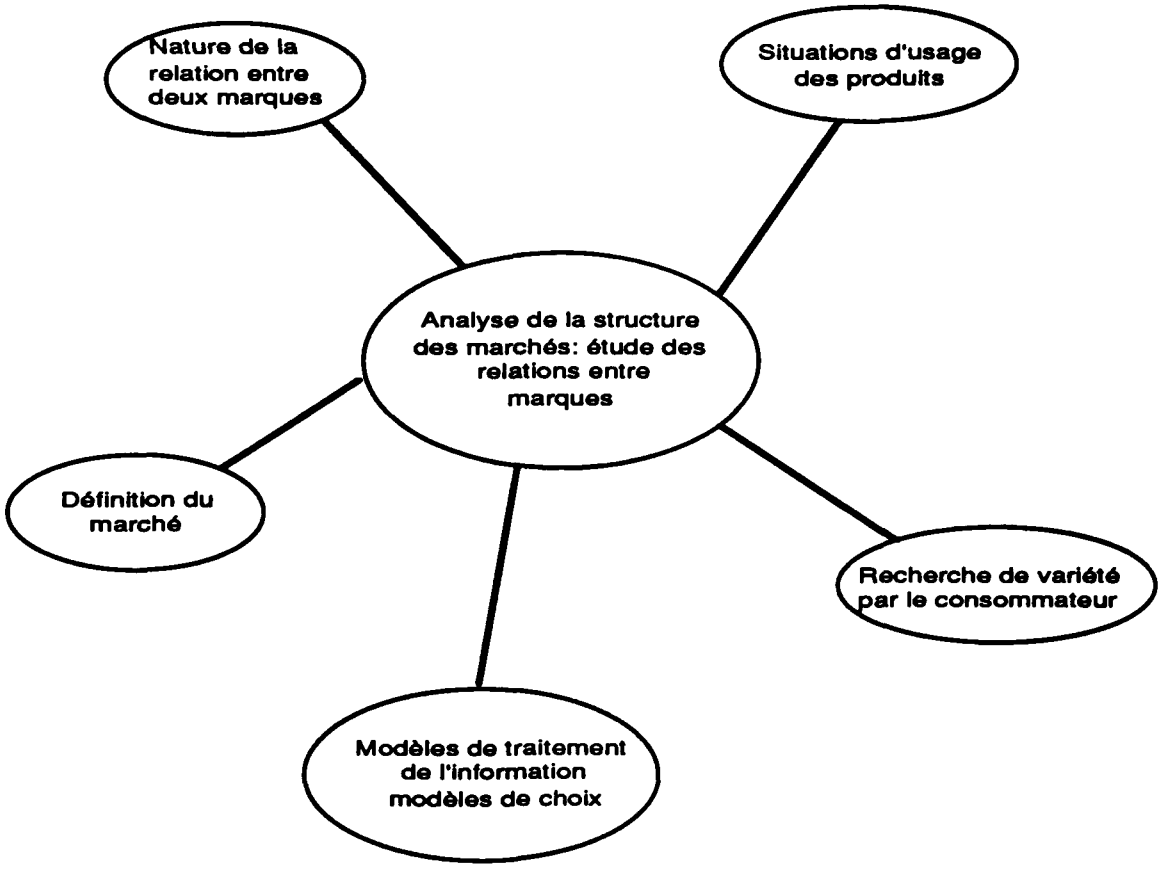
PREMIERE PARTIE

ANALYSE DE LA STRUCTURE DES MARCHES : LE POINT SUR LES ENJEUX THEORIQUES

CHAPITRE PREMIER PRESENTATION DES CONCEPTS FONDAMENTAUX

Nous avons choisi de présenter dans ce chapitre les concepts qu'il nous semble indispensable de replacer par rapport au champ théorique couvert par l'analyse de la structure des marchés. Comme nous l'avons dit dans l'introduction, ces concepts ont une importance capitale pour la compréhension des enjeux théoriques et pratiques de l'analyse de la structure des marchés. Ces concepts appartiennent à l'univers théorique proche des modèles de concurrence, nous considèrerons qu'ils en constituent les satellites. Notre objectif, dans ce chapitre, ne sera pas de faire une revue exhaustive de la littérature relative aux concepts présentés, mais de présenter les éléments nécessaires à l'étude approfondie des modèles d'analyse de la relation entre les marques.

FIGURE 3: CONCEPTS FONDAMENTAUX POUR L'ANALYSE DE LA STRUCTURE DES MARCHES



SECTION 1 LA SITUATION D'USAGE

La situation d'usage, dans l'étude de la concurrence entre les marques, revêt une grande importance. Elle n'a fait l'objet de recherches systématiques dans le domaine du comportement du consommateur et dans l'étude de la perception des marchés que depuis peu, avec les travaux de Belk (1974, 1975), Srivastava et al (1978), Srivastava (1980, 1981). Peu d'ouvrages traitant du comportement du consommateur y font explicitement référence. Belk en 1974, donnait déjà une explication à ce manque d'intérêt des chercheurs au travers des motifs suivants :

- la situation est un concept vague
- il n'y a pas de procédure systématique pour la caractériser
- les implications stratégiques liées à la connaissance des situations d'usage du produit ne sont pas toujours évidentes

A l'heure actuelle, ces motifs ne tiennent plus grâce aux travaux qui ont été faits par quelques chercheurs (très peu nombreux en fait). Il est donc injustifiable de passer cette variable sous silence lors de l'étude du jugement ou du comportement des consommateurs. Un des premiers auteurs à intégrer la situation d'usage dans une méthode de définition des marchés est Stefflre (1971 et 1979). Quelques années plus tard, Day, Shocker et Srivastava (1979) ont proposé d'utiliser la substituabilité des produits dans les situations d'usage comme indice de mesure de la relation entre les marques.

Nous aborderons successivement dans cette section:

- la place qu'occupe la situation d'usage parmi les autres variables dites situationnelles.
- à quels niveaux intervient la situation d'usage lors de la définition et l'analyse de la structure d'un marché (notamment pour la mesure de la relation entre les produits ou les marques).
- comment elle a été prise en compte dans les diverses approches.

SS1 LES VARIABLES SITUATIONNELLES DANS LA LITTERATURE MARKETING

L'étude en Marketing de Variables Situationnelles découle de la constatation empiriquement observée dans diverses recherches d'un écart entre :

1- La prévision des préférences ou des comportements à partir de variables explicatives qui peuvent être selon les cas, des perceptions, des préférences ou des intentions . Par exemple on peut essayer de :

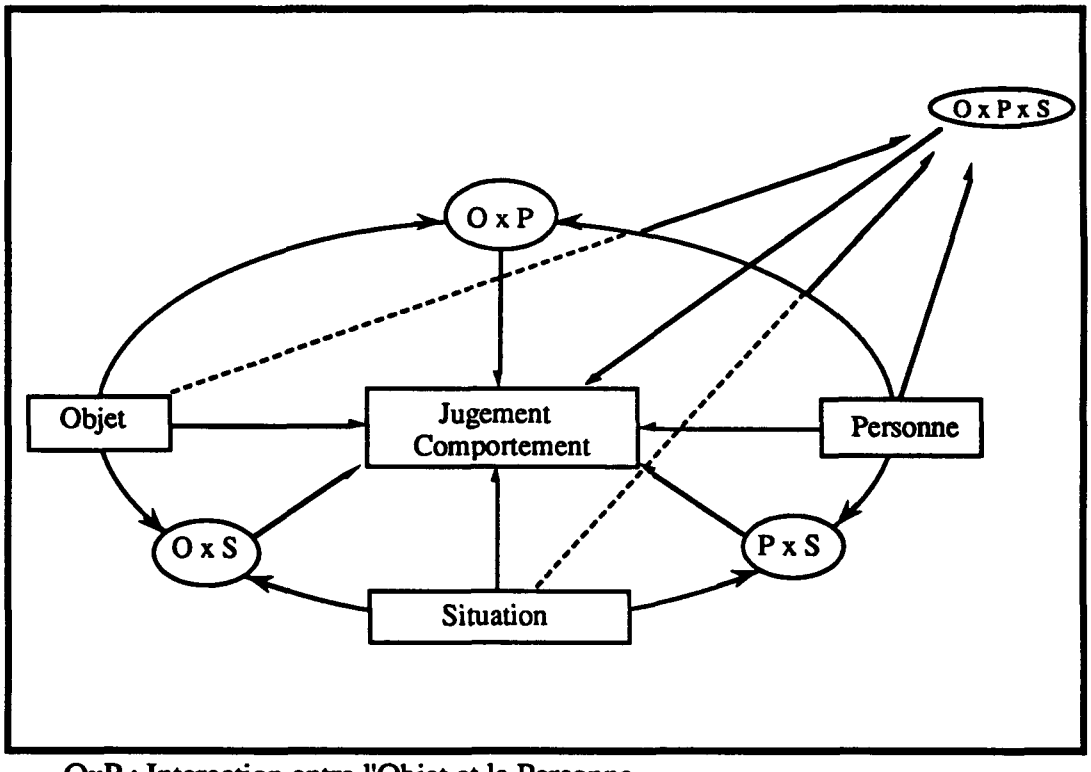
- prévoir les préférences à partir des perceptions. C'est ce qui est fait couramment dans les modèles multi-attributs.
- prévoir le comportement à partir des préférences ou des intentions, comme c'est le cas dans les modèles de choix du consommateur.

2- Les variables expliquées, qui sont soit des préférences, soit des comportements.

Statistiquement, le problème se pose en terme de "faiblesse du pourcentage de variance expliquée du comportement (ou des intentions ou des préférences) par les variables explicatives mesurées". La sophistication croissante des modèles explicatifs du comportement n'a malheureusement pas permis de résoudre ce problème.

Un des premiers auteurs à avoir montré la nécessité de prendre en compte la situation est Belk (1974) qui a observé dans une étude que 50% de la variance des attitudes pour des produits alimentaires était expliquée par l'effet principal de la situation et de son interaction avec d'autres variables. Belk (1974) définit la situation comme "l'ensemble des facteurs observés, particuliers à un instant ou un endroit, qui ne découlent pas de la connaissance de la personne ou du stimulus (alternatives de choix) et qui ont un effet systématique et démontrable sur le comportement". Cette définition est fallacieuse puisque la situation y est définie de façon "complémentaire" à l'individu et à l'objet. Tout ce qui n'est dû ni à l'un ni à l'autre serait dû à une troisième variable non directement mesurable mais qui serait le complément des deux premières. Avec cette définition, Belk considère en fait qu'il y a l'objet et ses caractéristiques (stables), l'individu et ses caractéristiques (instables) et puis le reste. Malgré ses faiblesses, cette définition a posé les bases du paradigme POS, Personne, Objet, Situation (voir Leigh et Martin 1981) qui est représenté sur la figure 4.

FIGURE 4 LE PARADIGME PERSONNE x OBJET x SITUATION



OxP : Interaction entre l'Objet et la Personne
OxS : Interaction entre l'Objet et la Situation
PxS : Interaction entre la Personne et la Situation
OxPxS : Interaction entre Objet, Personne et Situation

Selon ce paradigme, le comportement (ou le jugement) est défini comme étant le résultat de l'action de trois facteurs, la personne, l'objet et la situation mais aussi des interactions entre ces trois variables.

Gross et Niman (1975, voir aussi Dickson 1982), dans une revue de littérature, notent que les variables qui contribuent à la différence entre attitude (nous regroupons sous une même appellation d'attitude, les perceptions, préférences et intentions de comportement) et comportement sont les facteurs "individu", "situation" et "méthodologie" (de recherche ou d'enquête). De même, Wong et Sheth (1985) notent que quatre éléments participent à la différence observée entre l'intention de comportement (mesurée par des échelles d'attitude) et le comportement effectif :

1- Les événements inattendus. Ils regroupent en particulier:

- les caractéristiques intra-magasin
- les fluctuations de prix, les promotions
- la disponibilité du produit

Il s'agit donc d'un ensemble de variables couramment appelées variables de situation d'achat .

2 Les caractéristiques de l'individu et notamment sa capacité à agir par lui-même. Plus celle ci serait forte et plus l'individu aurait tendance à se comporter de manière différente par rapport à ses intentions.

3 L'environnement social. Il recouvre les "croyances normatives", ce que l'individu pense à propos de ce que des personnes importantes pour lui attendent de son comportement et les "situations anticipées" que l'individu envisage en exprimant son intention de comportement (voir Belk 1974 et 1979).

4 L'implication, plus celle-ci serait forte et plus l'individu, ayant fourni un effort de recherche d'information et percevant un fort risque (social, financier, psychologique ..) à l'achat aurait tendance à se comporter comme prévu. Nous noterons que l'implication correspond dans la figure 4 à l'interaction entre la personne et l'objet .

Plusieurs études montrent que parmi l'ensemble de ces variables, ce sont les facteurs situationnels qui expliquent le mieux le comportement. Notamment, la situation expliquerait un pourcentage de variance du comportement supérieur à celui expliqué par les différences individuelles (voir Wong et Sheth 1985 pour une revue de ces études). Le comportement prévu à partir de la connaissance de la situation serait donc plus fiable qu'à partir de la connaissance de l'individu. Miller et Ginter (1979), dans la même lignée de recherche, ont validé les hypothèses suivantes :

- le niveau d'achat des marques varie selon la Situation d'Usage et cette influence joue de façon plus ou moins sensible selon les marques envisagées. Certaines marques seraient finalement considérées indifféremment, quelle que soit la Situation d'Usage, d'autres au contraire seraient fortement associées, concernant leur perception, aux Situations d'Usage envisagées.
- la perception des marques varie selon les situations d'usage.

- l'importance accordée aux attributs des produits (attributs qui interviennent dans le processus de choix), est elle-même liée aux situations d'usage envisagées par le consommateur.

Ces deux hypothèses sont très importantes, elles nous montrent que le consommateur peut utiliser des modèles d'évaluation des marques dont les paramètres varient en fonction des situations d'usage envisagées (que ces modèles soient compensatoires ou non). Dans une série d'articles sur l'influence de la situation d'usage sur la perception des couples produits x marchés, Srivastava (1980, 1981) note que la situation d'usage envisagée influence les poids accordés aux attributs, les points idéaux (c'est à dire les préférences) mais contrairement aux résultats de Miller et Ginter, l'évaluation des marques selon les attributs resterait finalement assez stable d'une situation d'usage à l'autre. Ceci signifie que les comportements peuvent varier par le biais de la modification des préférences qui dépend elle même de la variation des poids accordés aux attributs, en fonction des situations d'usage envisagées. La perception quant à elle resterait inchangée.

SS2 LE CHAMP COUVERT PAR LES VARIABLES DE SITUATION

Il est possible de scinder, pour les présenter, ces variables en deux groupes, selon qu'elles ont un effet direct ou indirect sur le comportement d'achat ou d'usage (voir Figure 5 suivante).

SS2-1 LES VARIABLES DE SITUATION AYANT UN EFFET DIRECT SUR LE COMPORTEMENT: LA SITUATION D'ACHAT

Il s'agit ici essentiellement d'événements non attendus par le consommateur et qui modifient son comportement au moment de l'achat (voir Wong et Sheth 1985). Ce sont des variables *"intra-magasin"* telles que la publicité sur le lieu de vente (PLV), la promotion, la disponibilité, l'ambiance et qui sont généralement regroupées sous l'appellation de "Variables de Situation d'Achat". Elles affectent essentiellement ce que nous avons appelé plus bas, dans la figure 5, la *"réalité de marché"*.

SS2-2 LES VARIABLES DE SITUATION AYANT UN EFFET INDIRECT SUR LE COMPORTEMENT

Ces variables affectent ce que nous appelons la "réalité du consommateur" par le biais de ses attitudes qui affectent ensuite le comportement. C'est pourquoi nous donnons à ces variables le qualificatif d'effet "indirect". Il s'agit essentiellement de la variable Situation d'Usage.
Nous retiendrons dans nos analyses ultérieures que les variables situationnelles recouvrent principalement deux types :

- les variables de situation d'achat
- les variables de situation d'usage

Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

Ces deux types de variables, par leurs effets respectifs, peuvent expliquer en grande partie les différences entre comportement observé et comportement prévu. Nous retiendrons ici certaines considérations qui font que l'effet de la situation d'usage sur le comportement est difficile à isoler. En effet, l'attitude mesurée dans un questionnaire et l'attitude conduisant au comportement peuvent différer :

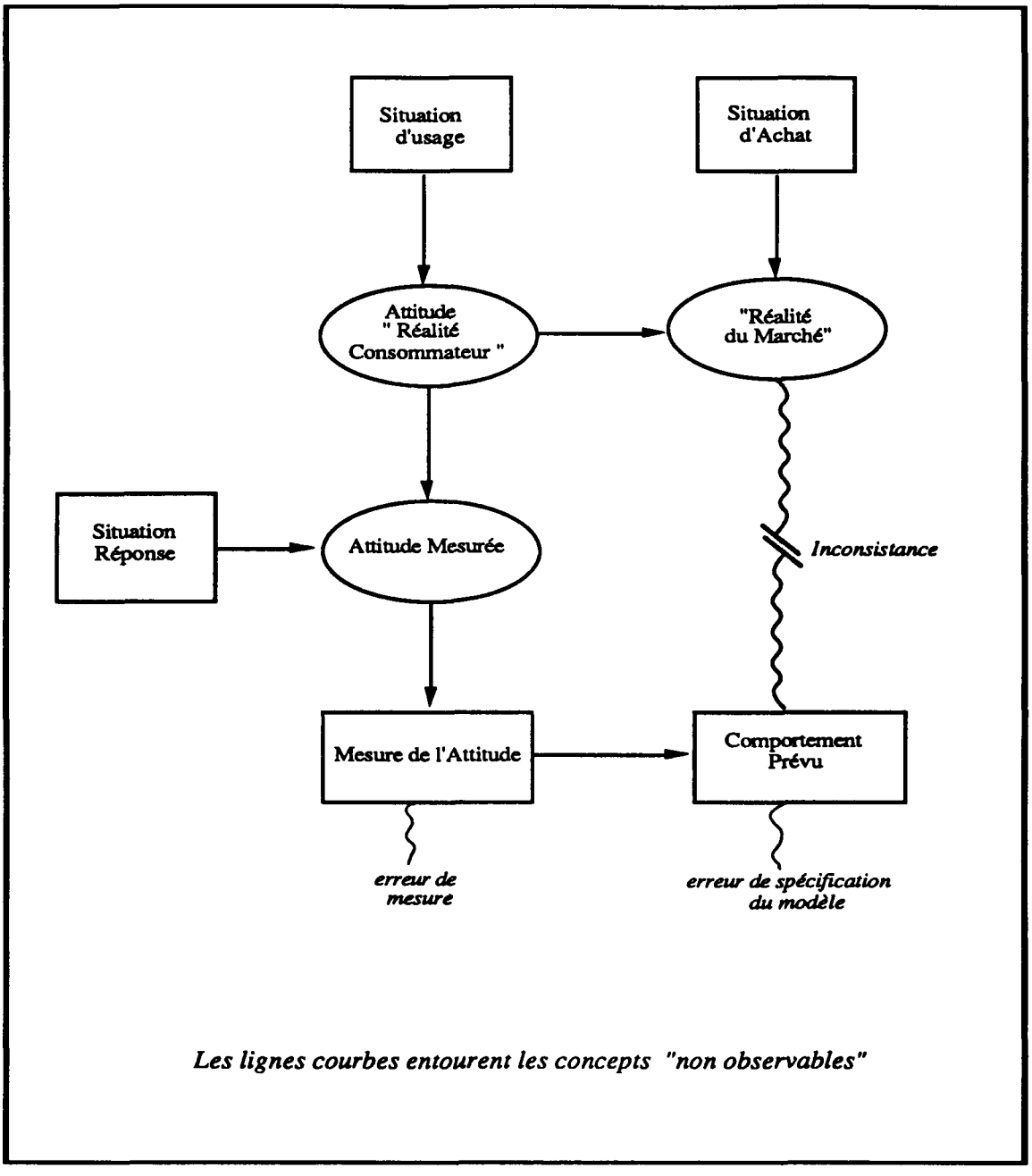
- parce qu'il est difficile de bien mesurer cette attitude
- parce que les conditions de formation de ces attitudes diffèrent d'un contexte à l'autre (c'est à dire entre la mise en situation du répondant lors de l'administration d'un questionnaire et la situation de comportement).

a - PROBLEME DE MESURE

L'attitude est une variable latente dont on ne mesure que le reflet dans des questionnaires (Bagozzi chapitre III 1980). La façon dont l'attitude se manifeste, donc se mesure, est fonction des pressions dues à la situation de réponse à ce questionnaire. La mesure des attitudes intègre au moins deux types de perturbations :

- l'erreur de mesure, due à la nature de l'échelle de mesure (le répondant reporte mal son attitude sur l'échelle).
- l'influence de la situation de réponse lors du questionnaire (personnalité de l'enquêteur, prescriptions normatives que se donne le répondant, etc...). Ainsi, l'attitude mesurée ne correspond pas à l'attitude réelle de l'individu (variable latente non observable). Ceci explique en partie la différence entre comportement observé et attitude mesurée (voir figure 5).

FIGURE 5 ETUDE DE LA REALITE CONSOMMATEUR / MARCHE



Les lignes courbes entourent les concepts "non observables"

b CONDITIONS DE FORMATION DES ATTITUDES

Beaucoup d'auteurs ont insisté sur le fait qu'il peut y avoir une différence entre attitude et comportement par le simple fait de la différence de spécificité du stimulus dans les deux cas. Dans les questionnaires, le stimulus auquel fait référence l'attitude telle qu'elle est mesurée est beaucoup plus général que le stimulus conduisant à l'expression d'un comportement réel (voir Gross et Niman, 1975, Leigh et Martin 1981). Or, la situation d'usage du produit est un des éléments importants qu'il faut prendre en compte pour éviter que le niveau de généralité conduisant au comportement de réponse à un questionnaire soit trop différent de celui conduisant au comportement de réponse sur le marché.

Lors de la mesure des attitudes, afin d'expliquer et prévoir des comportements de marché, l'erreur de mesure, l'effet de la situation de réponse au questionnaire, les différences de niveaux de spécificité entre l'objet de l'attitude mesurée et le stimulus à l'origine du comportement effectif sont autant d'éléments qui vont contribuer à un écart observé entre la mesure effectuée et le comportement observé. De plus, lors de la confrontation de la réalité consommateur, telle qu'elle aura été mesurée, avec le comportement, l'action des variables de situation d'achat viendra encore nuire à la qualité des modélisations (explication, prévision) qui auront été faites (voir figure 5). Il est à noter que nous ne tenons pas compte ici des autres variables explicatives des différences entre attitudes et comportements que sont les caractéristiques de l'individu, l'implication
A l'inverse, lors de l'observation des comportements tels qu'ils se manifestent (réalité de marché), il n'est pas possible d'en déduire pour autant les attitudes caractérisant la réalité consommateur puisque cette mesure prend en compte l'influence des variables de situation d'achat qui ne doivent pas être considérées lors de l'étude de la réalité consommateur. Ces variables de situation d'achat constituent le "pont" de passage entre les deux réalités.

La spécification correcte de la situation d'usage est donc importante puisqu'elle permet d'effectuer une mesure des attitudes plus juste en précisant l'objet de ces attitudes et une minimisation.de l'écart entre le comportement prévu et réalisé.

SS3 SITUATION D'USAGE ET ANALYSE DE LA STRUCTURE DES MARCHES

Puisque la situation d'usage influence perceptions et attitudes, il devrait donc être nécessaire de la prendre en compte pour définir ou structurer un marché. Il y a deux manières de tenir compte de la situation d'usage :

- analyser des situations objectives et significatives pour le consommateur, ces situations seront spécifiques à chaque marché étudié. L'unité analysée est ici, comme le note Fennell (1978), "la situation d'usage telle qu'elle est perçue par le consommateur". L'auteur préconise de déterminer ces situations d'usage en termes de *"conditions pour lesquelles le consommateur achète un produit"* (voir tableau suivant).

TABLEAU 1 MOTIVATIONS ET SITUATION D'USAGE

Types de motivations du consommateur

Perception de la situation d'usage du produit	Objet de la recherche de marque
1 Problème actuel	Résoudre le problème
2 Problème potentiel	Prévenir le problème
3 Diminution normale	Maintenir un état stable
4 Opportunité	Explorer
5 Opportunité de plaisir sensoriel	Recherche de plaisir
6 Problème relié au produit	Résoudre un conflit
7 Satisfaction-Frustation	Restructurer la situation

"Adapté de Fennell (1978) p 41"

Selon cette approche, la situation d'usage est concrète, donc bien adaptée en marketing bien que la diversité des situations d'usage puisse poser des problèmes lors de sa prise en compte.
Il existe une autre façon d'envisager la situation d'usage.

- classer les situations d'usage en fonction de la nature de l'état psychologique avec lequel elles sont reliées. Lutz et Kakkar (1975) ont effectué une classification des situations d'usage en termes de plaisir, éveil et dominance. Il est aisé de constater que cette approche n'est pas satisfaisante en Marketing, car trop générale.

Dans notre cadre d'analyse, nous considèrerons que la situation d'usage intervient à deux niveaux, sur la mesure de la concurrence et sur la hiérarchie du marché.

SS3-1 EFFET DE LA SITUATION D'USAGE SUR LA MESURE DE LA CONCURRENCE ENTRE LES MARQUES

Il paraît naturel de considérer que deux marques peuvent être concurrentes pour une situation d'usage donnée sans l'être pour une autre. Le consommateur recherche les bénéfices fournis par le produit plutôt que le produit lui-même, (Shocker et Srivastava 1979). Or, c'est l'usage qui requiert du produit des bénéfices particuliers. Nous en déduire alors les hypothèses suivantes :

- dans une perspective de segmentation des consommateurs, pour une situation d'usage particulière, deux individus seront identiques s'ils ont le même type de bénéfices recherchés. C'est le principe qui est retenu dans la pratique la segmentation par les bénéfices (Haley 1968).
- dans une perspective de mesure de la concurrence, pour une situation d'usage particulière, deux produits seront substituables s'ils offrent le même pattern de

bénéfices attendu par un groupe particulier d'individus, pour cette situation d'usage.
- inversement, (voir Shocker et Srivastava 79), deux situations d'usage seront équivalentes si elles exigent de la part des produits (pour un groupe particulier d'individus) les mêmes patterns de bénéfices attendus par les consommateurs.

Deux produits jugés substituables dans une situation d'usage donnée peuvent ne pas l'être dans une autre, ce qui signifie qu'ils appartiennent au même sous-marché dans un cas et pas dans l'autre. Il est donc nécessaire de "contrôler" la situation d'usage si on pense au départ qu'elle a un impact sur les attitudes et les comportements.

SS3-2 EFFET DE LA SITUATION D'USAGE SUR LA STRUCTURE DES MARCHES

Comme l'ont noté Shocker et al (1987), la hiérarchie d'un marché peut être considérée comme un indicateur plus ou moins valide du processus de choix utilisé par les consommateurs pour effectuer leurs choix. La prise en compte de la situation d'usage, dans l'analyse de la structure des marchés, peut se faire selon deux voies différentes :

- soit considérer qu'à chaque situation d'usage ou groupe de situations d'usage correspond une structure de marché particulière. C'est ce que nous avons noté précédemment quand nous avons vu qu'en fonction de la situation d'usage envisagée, le consommateur pouvait traiter différemment l'information, soit parce qu'il fait varier les poids accordés aux attributs, soit parce qu'il perçoit différemment les produits selon les situations d'usage.
- soit considérer que la situation d'usage peut être un critère de hiérarchisation du marché. C'est le cas quand le marché se partitionne en fonction des différentes situations d'usage ou groupes de situations d'usage (par exemple : ski de fond, ski de randonnée, ski alpin). Selon la terminologie du modèle de Hendry (voir au chapitre II), un marché est structuré selon la "forme" quand il se subdivise sur la base d'un attribut (ou plusieurs) des produits de ce marché. Or, chaque situation d'usage requiert une combinaison particulière d'attributs des produits ("un paquet de bénéfice attendus"), il serait ainsi plus simple pour l'individu de structurer le marché en termes de situations d'usage que de le faire attribut par attribut (c'est à dire bénéfice par bénéfice).

SS4 PRISE EN COMPTE DE LA SITUATION D'USAGE DANS LES MODELES D'ANALYSE DE LA STRUCTURE DES MARCHES

Selon le type de données utilisées (jugement, comportement), le type de modèle, la situation d'usage peut être totalement laissée pour compte ou au contraire être au centre de l'analyse. Nous pensons qu'il existe deux types de prise en compte de la situation d'usage :

- une prise en compte active, où la situation d'usage est la base sur laquelle repose la construction des indices de mesure de la concurrence.
- une prise en compte passive où il ne s'agit que de contrôler l'effet de la situation d'usage sans vraiment l'intégrer dans la construction du modèle.

SS4-1 - PRISE EN COMPTE DIRECTE DE LA SITUATION D'USAGE

Ici, la capacité du produit (ou de la marque) à satisfaire une situation d'usage particulière sert de base à l'étude des relations de concurrence entre les produits. Cette approche est essentiellement le fait d'un petit nombre de chercheurs ayant beaucoup étudié l'influence de la situation d'usage sur la perception des couples produits x marchés (voir Srivastava, Shocker et Day 1978, Srivastava 1980 et 1981, Belk 1979).

Il apparaît à travers ces travaux que la perception des produits par les répondants est très homogène lorsque l'effet de la situation d'usage est contrôlé, ce qui signifie que la situation d'usage explique une part non négligeable de la concurrence entre les produits. Des produits possédant les mêmes patterns d'usages sont vraisemblablement substituables. L'inconvénient de la prise en compte de la situation d'usage pour l'étude de la perception des couples produits x marchés (donc de la mise en oeuvre d'un modèle d'analyse de la structure des marchés avec des données de jugement), réside dans l'augmentation considérable de l'information à collecter puisqu'il faut connaître perceptions, attitudes et préférences au niveau de chacune des situations d'usage . Cependant nous noterons :

- qu'il n'est pas forcément nécessaire de s'intéresser à toutes les situations d'usage
- que, comme l'ont montré Srivastava Shocker et Day (1979), Srivastava (1980 et 1981), l'hétérogénéité apparente des consommateurs diminue beaucoup lorsque la situation d'usage est contrôlée Ceci est une conséquence directe de l'effet de spécificité du stimulus de l'attitude noté par Fishbein (1966) et Dollard (1949). Ainsi, dans les études, la différence entre préférence et choix est souvent attribuée à la variance de l'erreur, les différences de préférences entre individus à l'hétérogénéité des consommateurs. En fait, les effets de ces deux types de variance sont en partie éliminées avec le contrôle de la situation d'usage. Les consommateurs étant alors moins hétérogènes, la taille d'échantillon requise devient moindre. Ceci peut compenser le surcroît d'information à collecter occasionné par la prise en compte de la situation d'usage.

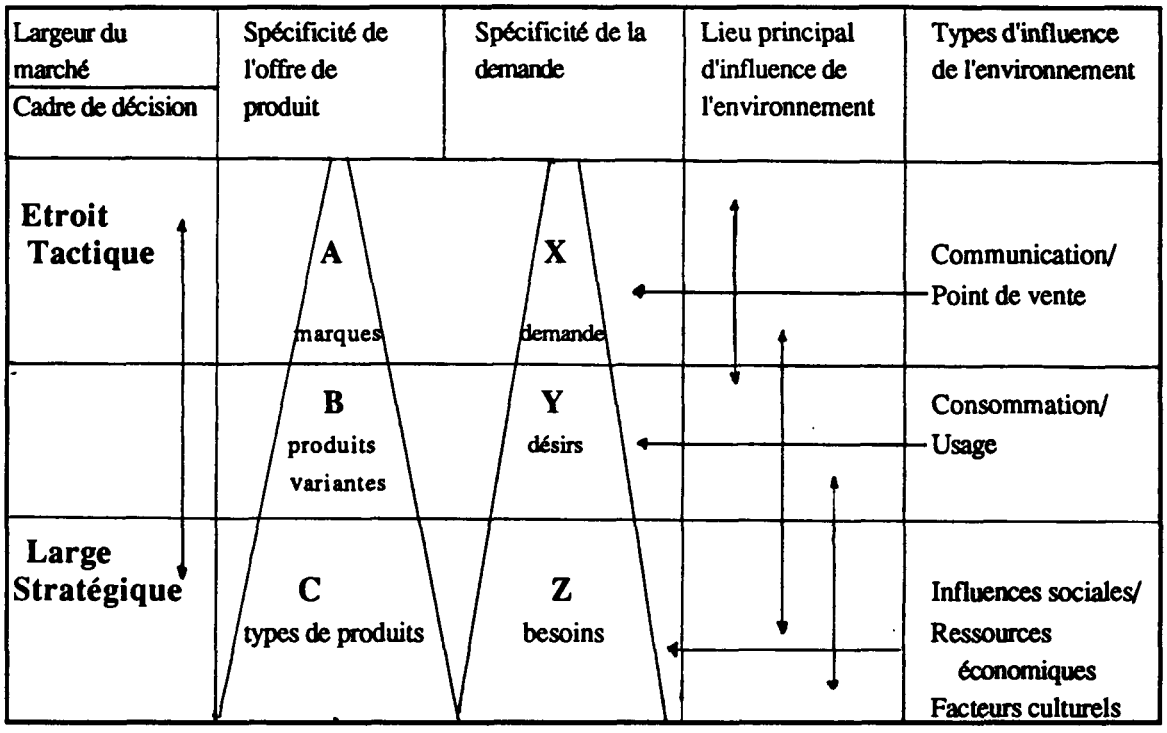
La mesure de la substituabilité à partir de la situation d'usage doit être considérée dans un contexte d'analyse large et prospectif. L'intérêt ne se porte pas ici sur la concurrence directe entre les produits (à fortiori entre marques), telle qu'elle s'exprime sur le marché, mais à sur une concurrence plus large, à plus long terme (c'est ce que Srivastava, Leone et Shocker appellent "la concurrence potentielle).

L'intérêt d'une telle mesure de la substituabilité n'est donc pas d'analyser les relations entre des marques aux caractéristiques identiques appartenant à un sous marché défini au sens étroit, relations qui sont fortement affectées par les variables situation d'achat et communication". La mesure intéresse ici des "produits" (ou variantes de produits) qui "peuvent" entrer en concurrence sous certaines conditions et en particulier la situation

d'usage (voir figure 6) et non pas des "marques" pour lesquelles l'univers concurrentiel est en général défini de façon étroite .

FIGURE 6 LES NIVEAUX D'ANALYSE SUR UN MARCHE

Cadre d'analyse des relations de marché



"Adapté de Srivastava, Léone, Shocker 1981, p40"

La figure 6 suggère que l'analyse de la concurrence (et donc de la structure d'un marché) peut s'effectuer, selon le point de vue adopté par le manager, à différents niveaux de marché (voir les auteurs cités pour plus de détails):

- relations entre marques qui sont essentiellement affectées par la communication et les variables de situation d'achat. Ce type de relation s'exprime au sein de ce que nous avons appelé la "réalité de marché", c'est à dire l'adaptation de la réalité consommateur aux contraintes réelles du marché (comme par exemple les variables situationnelles à effet direct). Ce type de relations peut être étudié à partir de données de comportement (historiques d'achats).
- relations entre variantes de produits qui, elles, s'inscrivent plutôt au niveau de la "réalité consommateur" et seraient fortement influencées par la situation d'usage. Les données de comportement ne représentent plus alors qu'un pâle reflet de cette réalité.
- relations entre types de produits, ce qui place l'analyse au niveau des besoins des consommateurs. Ces besoins sont influencés par un ensemble de variables appartenant à un univers beaucoup plus large (facteurs culturels, sociaux, économiques). C'est en se plaçant à ce niveau qu'il sera possible de prendre en

compte les mutations de l'environnement et donc d'anticiper les évolutions futures. Au contraire, plus nous remontons dans la figure vers des niveaux où le type d'influence est spécifique et plus grand est le risque de ne pas voir arriver le changement même si les relations au sein d'un univers étroit sont analysées plus finement.

La mesure de la substituabilité par les situations d'usage est donc mieux adaptée à l'étude des développements possibles d'un marché (potentialités de développement de nouveaux produits, diversification de la firme vers d'autres types de produits) et à l'évolution de sa structure.

SS4-2 PRISE EN COMPTE INDIRECTE DE LA SITUATION D'USAGE

Dans ces approches, la situation d'usage n'est pas réellement intégrée dans un modèle d'analyse de la concurrence. On lui reconnaît avoir cependant un effet sur la mesure de la concurrence entre les produits et le type de hiérarchie structurant le marché. La situation d'usage peut être alors intégrée soit explicitement, soit implicitement.

* INTEGRATION EXPLICITE (CONTROLE DE LA SITUATION D'USAGE)

Dans ce cas, il est nécessaire de connaître les situations d'usage relatives à la catégorie de produits étudiée, ce qui n'est généralement pas possible avec l'utilisation de données historiques d'achats des consommateurs. Il est alors possible de "contrôler" la variable situation d'usage en testant (dans les modèles confirmatoires) l'adéquation aux données, de différentes structures du marché correspondant aux différentes situations d'usages envisagées. C'est ce qui est fait dans le modèle PRODEGY (Urban, Johnson, Hauser 1984, voir chapitre II).

* INTEGRATION IMPLICITE

Sans disposer d'information concernant l'adéquation des produits aux différentes situations d'usage, ce type d'approche permet néanmoins d'en tenir compte pour étudier la structure du marché .Le modèle de Fraser et Bradford (1983), voir chapitre II section 3) permet de prendre en compte dans une certaine mesure, la situation d'usage.

Nous avons pu rendre compte, dans cette section, de l'importance que revêt la situation d'usage lors de l'étude des relations entre produits (ou marques) appartenant à un marché. Nous avons noté la place importante qu'occupe la situation d'usage parmi les variables situationnelles, quel est son impact sur les perceptions, attitudes et comportements, en quoi elle peut expliquer une partie des relations de concurrence entre les marques. Nous allons maintenant étudier une autre variable, la "recherche de Variété", qu'il est aussi nécessaire de considérer lors de l'étude des relations existant entre des marques.

SECTION 2 LA RECHERCHE DE VARIETE

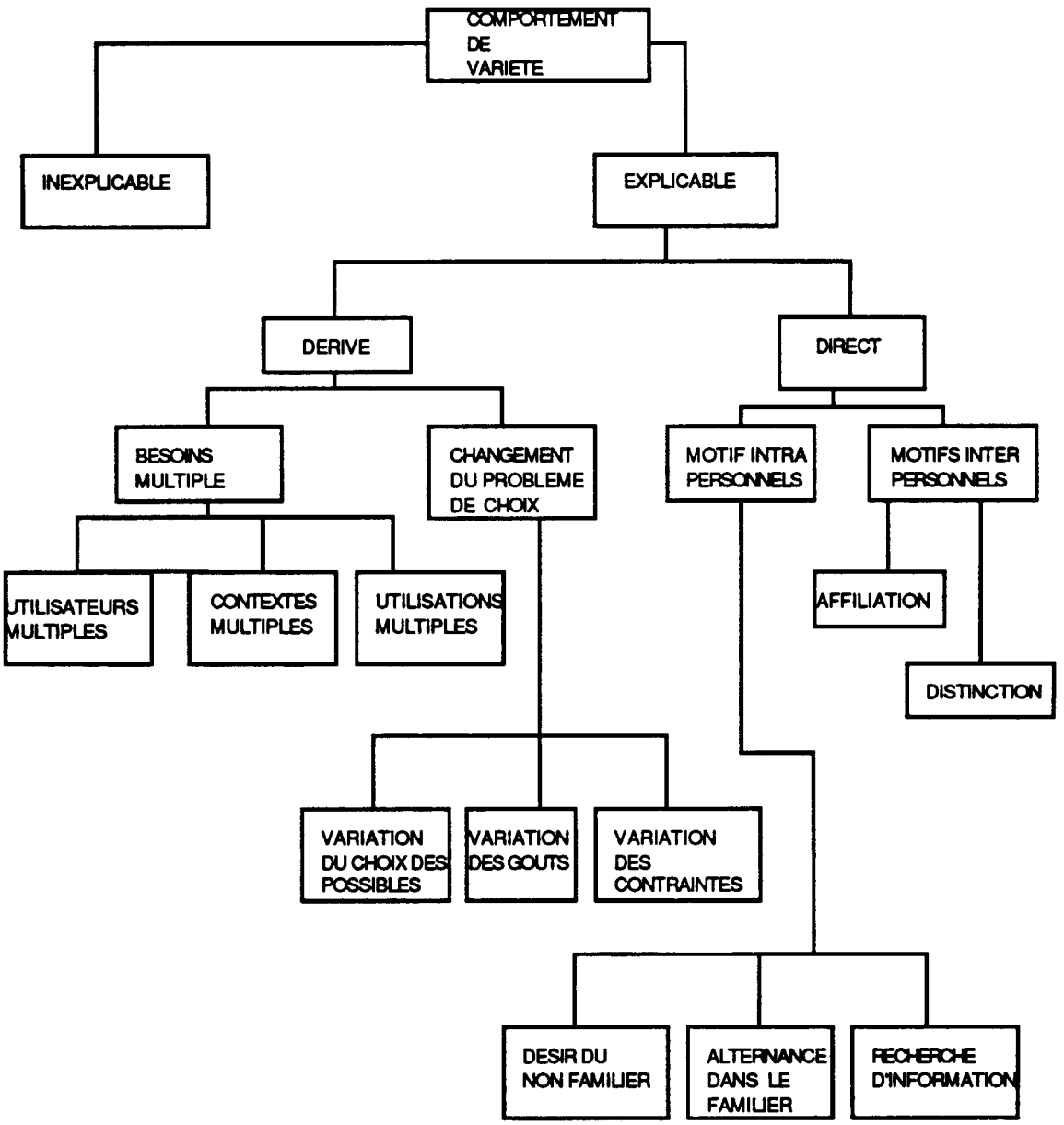
La recherche de variété (Variety Seeking) fait partie des concepts que nous jugeons centraux dans notre recherche sur l'analyse de la structure des marchés.
Nous proposons ici de présenter le concept de recherche de variété tel qu'il a été développé dans la théorie marketing. Dans une première partie, nous effectuerons une revue de la littérature portant sur le sujet. Dans une deuxième partie, nous aborderons le problème de la mesure de la recherche de variété. Enfin, dans une troisième partie, nous présenterons les implications de l'existence de recherche de variété pour l'étude de la structure d'un marché et l'analyse des relations de concurrence entre marques.

SS1 RECHERCHE DE VARIETE: PRESENTATION ET DEFINITIONS

Le concept recherche de variété est jeune dans la littérature Marketing comparé au concept de fidélité qui a fait l'objet de nombreux articles (voir Jacoby et Kyner 1973, pour une revue de littérature).

Reprenant le schéma présenté par McAlister et Pessemier (1982), lui-même issu de l'analyse de Laurent (1978), les divers mécanismes conduisant les individus à la recherche de variété peuvent être représentés de la manière suivante :

FIGURE 7 TAXONOMIE DES THEORIES DU COMPORTEMENT DE RECHERCHE DE VARIETE



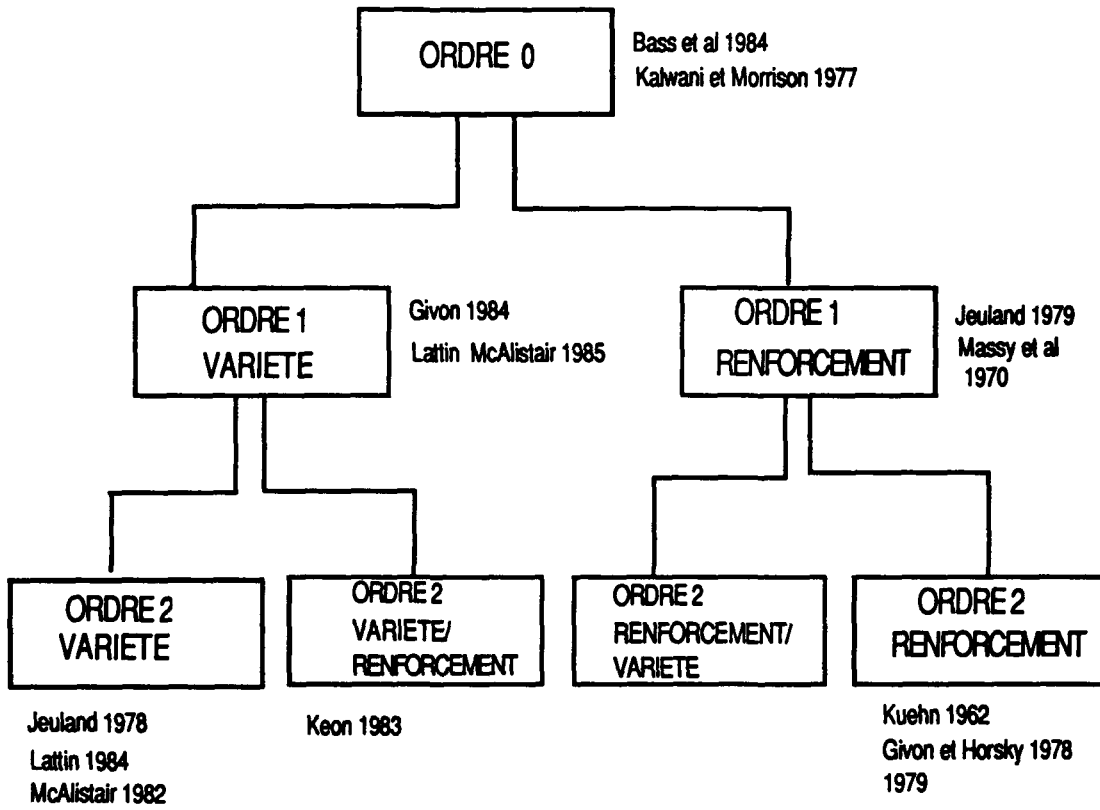
"Adapté de Laurent (1978) "

Nous commenterons ce schéma brièvement et renvoyons le lecteur aux références précédentes pour plus de détails.
Le "comportement de variété" d'un individu peut être considéré comme inexplicable ou explicable.

SS1-1 APPROCHE STOCHASTIQUE (comportement inexplicable)

Le principe général de ces approches (Voir Bass 1974, Pessemier 1978, Bass et Pilon 1980, Jeuland 1979) est de considérer que les préférences (Bass 1974) ou les points idéaux d'un consommateur (Huber et Reibstein 1978 cité dans McAlister et Pessemier 1982) dépendent d'un processus stochastique. Selon le type d'hypothèses sur lesquelles sont fondés ces modèles, la taxonomie suivante a été proposée par Kahn, Kalwani et Morrison (1986).

FIGURE 8 TAXONOMIE DES MODELES DE RECHERCHE DE VARIETE ET DE RENFORCEMENT



"Adapté de Kahn et al (1986) p93 "

Nous allons successivement commenter ces différents types de modèles.

SS1-1-1 LES MODELES D'ORDRE 0:

Ce modèle caractérise l'absence de comportement de recherche de variété de la part du consommateur. Soit p la probabilité marginale de choix d'une marque x. La probabilité d'achat de cette marque est constante, quelque soit le choix effectué lors de l'occasion précédente.

P(1/1) = P(1/0) = p

P(1/1) est la probabilité d'acheter la marque à l'occasion t, sachant qu'elle a été achetée à l'occasion précédente, t-1

P(1/0) est la probabilité d'acheter la marque à l'occasion t, sachant qu'elle n'a pas été achetée lors de la dernière occasion d'achat, t-1

SS1-1-2 LES MODELES D'ORDRE 1

Ces modèles permettent de prendre en compte l'effet d'un achat sur l'achat immédiatement suivant. Deux types de comportements sont possibles.

1. Modèle d'ordre 1, avec recherche de variété:

Ce modèle permet de caractériser un effet de recherche de variété par rapport à l'achat précédent. Si une marque a été choisie lors de l'achat précédent (en t-1), la probabilité qu'elle soit rachetée (en t) sera d'autant plus diminuée que la recherche de variété sera forte.

P(1/1) = p - Vp

V est le coefficient de recherche de variété caractérisant le marché. Si V=0, la recherche de variété est nulle, ce qui correspond au cas d'un modèle d'ordre 0. A l'inverse, si V=1, la probabilité de réachat d'une marque est nulle.

2 Modèle d'ordre 1, avec renforcement

Ce modèle permet de caractériser un comportement de renforcement du consommateur (fidélisation) par rapport à l'achat précédent. La probabilité de réachat d'une marque sera alors d'autant plus forte que la tendance au renforcement est élevée.

P(1/1) = p + (1-p)R

(1-p) est la probabilité de non achat de la marque x

R est le coefficient de renforcement. Il est compris entre 0 (absence de renforcement) et 1 (renforcement très fort). Dans le cas où la valeur du coefficient de renforcement est forte (proche de 1), la probabilité de réachat d'une marque, lorsqu'elle a été choisie à l'occasion précédente, tend vers 1.

SS1-1-3 LES MODELES D'ORDRE 2

Ces modèles permettent de prendre en compte les deux derniers achats du consommateur. Quatre types de comportements sont alors identifiables:

1 Recherche de variété d'ordre 2:

Dans ce cas, la probabilité d'achat d'une marque sera d'autant plus faible qu'elle aura été achetée lors des deux dernières occasions: l'effet de lassitude par rapport à la marque s'accentue, avec une expérience prolongée sur deux occasions.

P(1/11) < P(1/01)

P(1/11) est la probabilité d'achat de la marque x quand elle a été achetée lors des deux dernières occasions (t-1 et t-2)

P(1/01) est la probabilité d'achat de la marque x quand elle n'a été achetée que lors de la dernière occasion d'achat (t-1)

P(1/11) = p - Vp

$$P(1/01) = p - V^2p$$

Comme V^2 est inférieur ou égal à V , P(1/11) est lui même inférieur à P(1/01)

2 Renforcement d'ordre 2

Dans ce cas, la probabilité de réachat d'une marque va encore augmenter si elle a été achetée lors des deux dernières occasions d'achat.

$$P(1/11) = p + R(1-p)$$
$$P(1/11) = p + R^2(1-p)$$
$$P(1/11) > p(1/01)$$

Deux autres modèles intermédiaires existent encore.

3 Modèle de variété / renforcement

Si une marque a été achetée à la seule occasion précédente, il y a tendance a la recherche de variété.

$$P(1/10) = p - Vp$$

Par contre, si une marque a été achetée lors des deux occasions précédentes, il y a tendance au renforcement.

$$P(1/11) = p + R(1-p)$$

4 Modèle de renforcement /Variété

Si une marque a été achetée seulement lors de la dernière occasion, il y a alors tendance au renforcement.

$$P(1/01) = p + R(1-p)$$

Si, par contre, la marque a été achetée lors des deux dernières occasions, il y a tendance à la recherche de variété.

$$P(1/11) = p - Vp$$

Dans tous ces modèles, les probabilités de transitions entre marques (probabilités conditionnelles) sont estimées à partir des historiques d'achats. Leur limite essentielle tient à ce qu'ils n'expliquent en rien le pourquoi ni le quand du comportement de Recherche de Variété, pas plus qu'ils ne permettent de savoir exactement ce que recouvre ce comportement. Est-il dû à une lassitude vis à vis d'un attribut du produit, à un changement de la situation d'usage du produit, à la coexistence de goûts hétérogènes au sein du même foyer?

SS1-2 APPROCHE DETERMINISTE (comportement explicable)

Le comportement de variété peut alors être dû à deux grands types de facteurs ;

SS1-2-1 LES FACTEURS DERIVES

Dans ce cas de figure, le consommateur observe un comportement de variété pour des raisons autres que la recherche de variété, raisons qui sont de deux ordres:

1 L'existence de besoins multiples

Ces besoins sont liés à l'existence d'une variable très importante en marketing: la situation d'usage. Le changement ne se fait pas pour lui même, mais pour une cause indirecte : le changement de la situation d'usage du produit. A chaque situation d'usage correspond une combinaison particulière d'attributs du produit attendue par le

consommateur. Selon la situation d'usage envisagée, l'importance des attributs de la marque mais aussi l'évaluation des marques selon ces attributs peuvent varier. Ces changements, replacés dans le contexte des modèles de choix multi-attributs expliquent alors les changements opérés en matière de choix des marques.

2 Un changement dans la structure des préférences du consommateur

Celle-ci peut être due soit :

- à une variation des produits disponibles (suppression ou lancement de produit, disponibilité du produit en magasin...)
- au changement des goûts des consommateurs résultant d'une évolution de son environnement (publicité par exemple), ou de lui-même.
- à un changement dans les contraintes du choix (dû par exemple à un changement de profession)

SS1-2-2 FACTEURS DIRECTS

Dans ce cas, le consommateur va changer de marque ou de variante de produit (voir Laurent 1978 page 1 pour une définition) parce qu'il recherche effectivement la variété dans sa consommation. Cette recherche peut être due à des motifs intra-personnel (propres à l'individu) ou inter-personnels (son entourage).

1 Motivations intra-personnelles

Les recherches en psychologie et en comportement du consommateur ont formulé un ensemble de théories synthétisées par Driver et Streufert (1964, cité dans McAlister et Pessemier 1982) selon lesquelles l'individu est supposé avoir un niveau optimal de stimulation. En dessous de ce niveau, le consommateur entrerait dans un processus visant à ramener cette stimulation à un niveau acceptable :

- soit par le désir (l'achat ou la consommation) de produits non familiers, ce désir se manifestant généralement par l'achat de produits nouveaux.
- soit par le désir de changement parmi des produits familiers, stratégie visant plutôt à minimiser le risque (voir Brickman et d'Amato 1975).

Ces comportements d'alternance ont été modélisés par Farqhar et Rao (1976), Jeuland (1978), McAlister (1979), McAlister (1982). Dans ces modélisations, les auteurs considèrent que l'attribut d'un produit influence la formation des préférences à deux niveaux :

- en fonction de son importance (ou de sa déterminance), c'est l'approche classique des modèles de préférences.
- en fonction de la fréquence avec laquelle il a été consommé auparavant, étant entendu que plus il aura été consommé et plus le consommateur en quête de variété aura tendance à faire de la possession de cet attribut par un produit un facteur de non-achat.

A ce sujet, nous préciserons que Farqhar et Rao (1976) définissent quatre types d'attributs :

-des attributs d'équilibre (équi-balancing) pour lesquels le consommateur recherche une homogénéité des scores des différentes marques qu'il envisage d'acheter .

-des attributs de déséquilibre (counter-balancing) qui sont ceux pour lesquels le consommateur recherche une large diversité des scores des marques.

-des attributs désirables
-des attributs indésirables

Ces modèles sont construits selon deux optiques :

- une optique instantanée, ce qui place l'analyse dans le cas d'une acquisition concurrente de différents produits (voir McAlister 1979).
- une optique dynamique (McAlister 1982, voir annexe 2), où les produits sont acquis (consommés) en différentes périodes. Dans ce cas, il faut se garder de confondre l'effet de la recherche de variété avec celui qui est dû à la modification (définitive) des goûts et des préférences, auquel cas il s'agira d'un facteur de type dérivé.

Simonson (1990) a montré que dans la deuxième optique, la recherche de variété est moins forte que dans la première, celle où les produits sont acquis de manière instantanée.
Une autre explication de la recherche de variété peut être trouvée dans le désir pour le consommateur d'acquérir de l'information. Les produits sont alors achetés afin de mieux les connaître (acquisition d'information) ou de se les rappeler (rappel d'informations oubliées) .

2 Motivations inter-personnelles

Dans ce cas, le comportement de recherche de variété est la manifestation de motivations ayant trait principalement :

- aux besoins de manifestation de l'appartenance à un groupe dont le consommateur imite les changements de comportement.
- au besoin de manifestation de l'identité individuelle conduisant à l'adoption d'un style de vie particulier (par exemple, une forte propension à l'achat de produits nouveaux).

Les deux approches de la recherche de variété que nous venons de présenter (stochastique et déterministe) recouvrent finalement un champ très large. Devons-nous pour autant en déduire que la recherche de variété correspond à l'ensemble des situations que nous avons évoquées? Les définitions qui ont été données à la recherche de variété recouvrent en fait des domaines d'étendue variable. Kahn, Kalwani et Morrison (1986) définissent la recherche de variété comme *"une tendance délibérée à changer de la marque achetée lors de la dernière occasion d'achat pour une autre"*. L'objet de leur approche est avant tout de caractériser chaque marque ou chaque sous-marché par un indice de recherche de variété et donc par un type particulier de modèle de choix, d'ordre zéro, un ou deux. Cette définition est générale, puisque la recherche de variété y est définie en termes de changement de marque, sans référence aux raisons de ce changement. C'est en général le point de vue adopté dans les modèles stochastiques de choix dans lesquels les seules informations disponibles sont les changements entre marques. Ces modèles procèdent en effet à partir de données de comportement effectif (données de panels consommateurs) et l'information disponible se limite à la connaissance des achats effectués en diverses occasions d'achats. Rien n'est connu sur les motifs de ces achats: situation d'usage du produit, hétérogénéité des goûts dans le ménage, recherche effective de variété, etc.

A notre avis, une définition plus restrictive doit être donnée au concept de recherche de variété, puisque le comportement de "variété dérivé" doit être considéré comme une résultante de l'effet de l'environnement qui peut être contrôlé, (par exemple la situation d'usage) et non comme un processus psychologique caractérisant le consommateur. Ainsi, certains auteurs adoptent une définition étroite de la recherche de variété. McAlister (1982) ne considère comme recherche de variété que *"le changement pour des motivations directes parmi un ensemble de marques connues. Dans le cas où les marques sont inconnues, il s'agirait de recherche d'information"*.
Une définition précise de la recherche de variété ainsi que la volonté d'en expliquer les motifs nécessitent l'utilisation de données de jugements (perceptions, attitudes, préférences et choix).

SS2 - MESURE DE LA RECHERCHE DE VARIETE

Il est important, soit pour caractériser le modèle de choix d'un individu, soit pour caractériser et comprendre les transferts entre marques sur un marché, de pouvoir mesurer le niveau de la recherche de variété. Nous retrouvons alors la dichotomie qui oppose les modèles stochastiques de choix, fonctionnant à partir de données de comportement (transferts entre marques) aux modèles déterministes qui utilisent en général des données de jugement (perceptions, préférences et choix).

SS2-1 APPROCHES DE TYPE STOCHASTIQUE

Le comportement de variété à l'égard d'une marque au sein d'un sous marché, ou du sous-marché lui-même, est caractérisé en déterminant quel modèle de choix est le plus compatible avec les probabilités calculées à partir des historiques d'achats des marques. Parmi les différents modèles de la figure 8, afin de savoir quel type de modèle de choix utiliser, Kahn, Kalwani et Morrison (1986) ont proposé des tests de signe selon trois critères C1, C2 et C3 :

C1 = PA(1/10) - PA(1/01)
C2 = PA(1/1100) - PA(1/0011)
C3 = C2 - C1

où :

- PA(1/10) est la probabilité conditionnelle empirique (agrégée sur l'ensemble des consommateurs) de choisir la marque x avec l'historique d'achat 10 (achat de x en t-2, non-achat de x en t-1)
- PA(1/01) est la probabilité conditionnelle empirique (agrégée sur l'ensemble des consommateurs) de choisir la marque x , avec l'historique d'achat 01 (non-achat de x en t-2, achat de x en t-1)
- PA(1/1100) est la probabilité conditionnelle empirique (agrégée sur l'ensemble des consommateurs) de choisir la marque x avec l'historique d'achat 1100 (achat de x en t-4 et t-3, non-achat de x en t-2 et t-1)
- PA(1/0011) est la probabilité conditionnelle empirique (agrégée sur l'ensemble des consommateurs) de choisir la marque x avec l'historique d'achat 0011 (non-achat de x en t-4 et t-3, achat de x en t-2 et t-1)

Les valeurs des indices de recherche de variété ou de renforcement sont estimées à partir de ces probabilités empiriques. Par exemple, pour un modèle d'ordre 1 avec recherche de variété :

$$V = PA(1/10) - PA(1/01)$$

V est l'indice de recherche de variété, V appartient à l'intervalle (0,1)

Ces trois tests permettent de savoir quel est l'ordre du modèle correspondant à la situation étudiée, parmi les différents modèles utilisables, en comparant le signe obtenu avec ceux qui figurent dans le tableau proposé par les auteurs (voir tableau suivant).

TABLEAU 2 SIGNIFICATION DES TESTS DE SIGNES

Modèle	C1	C2	C3
Ordre 0	0	0	0
Ordre 1, variété	+	+	0
Ordre 2, variété	+	+	+
Ordre 2, var./renf.	+	-	-
Ordre 1, renforcement	-	-	0
Ordre 2, renforcement	-	-	-
Ordre 2, renf./var.	-	+	+

Givon (1984) a proposé un modèle qui permet de prendre en compte simultanément n marques au plan individuel. Dans ce modèle, la probabilité conditionnelle du choix de la marque j sachant que la marque i a été achetée à la période précédente, P(j/i), est une combinaison linéaire de deux types d'utilité: l'utilité associée à la marque pour laquelle le consommateur change et l'utilité associée au changement lui même.

$P(j/i) = V[\ 1/(n-1)] + (1-V)\pi_j$ pour tout j différent de i

π_j est la préférence du consommateur pour la marque j (la probabilité marginale de choix de la marque j).

n est le nombre de marques dans l'ensemble évoqué de ce consommateur c'est à dire les marques qu'il a achetées sur la période étudiée.

V est le niveau d'utilité associée au changement. Si V est proche de 1, le consommateur est fortement chercheur de variété, à l'inverse, si V est proche de 0, le consommateur ne trouve aucune utilité dans le changement.

Remarquons que lorsque la tendance à la recherche de variété du consommateur est très forte (V proche de 1), la probabilité de changement pour une autre marque j n'est fonction que de V et du nombre de marques dans l'ensemble évoqué moins une (la marque i dont le consommateur change). P(j/i) s'écrit alors :

$$P(j/i) = V/(n-1) \approx 1/(n-1)$$

Toutes les marques, exceptée i, ont ainsi la même probabilité d'être choisies et ceci quelle que soit la marque i achetée à la période précédente. La recherche de variété aboutit à l'uniformisation des préférences, c'est à dire la diminution de la préférence pour les marques dont la préférence était à la base forte et inversement l'augmentation de la préférence pour les marques dont la préférence était à la base faible. Cet effet est d'autant plus fort que le nombre de marques dans l'ensemble évoqué du consommateur est faible. Quand la recherche de variété est faible (V proche de 0), le choix d'une marque est uniquement fonction de la préférence du consommateur, quelle que soit la marque achetée à la période précédente.

V est estimé par la méthode du maximum de vraisemblance, à partir des probabilités marginales de choix des marques et des transferts effectués entre les marques.

Dans les études empiriques qu'il a menées, Givon (1984 et 1985) a caractérisé les points suivants :

1 - La recherche de variété est un concept relié à l'attribut du produit et non au produit lui même. Pour satisfaire son besoin de variété relativement à un attribut définissant une partition particulière, le consommateur est amené à effectuer son choix dans une autre partition ce qui montre bien que le produit est désiré non pour lui même mais pour l'ensemble des bénéfices qu'il apporte au consommateur. Les indices V varient beaucoup, pour le même consommateur, d'un attribut à l'autre.

2 - Les coefficients V de recherche de variété varient beaucoup d'un individu à un autre et d'une catégorie de produit à une autre. Cette constatation amène l'auteur à proposer une segmentation des consommateurs sur la base de leurs coefficients de recherche de variété selon chaque attribut. Deux individus ayant les mêmes indices auraient les mêmes hiérarchies de choix et devraient ainsi répondre de façon équivalente à une action de lancement de nouveau produit ou de repositionnement d'une marque sur le marché. Cette approche est conceptuellement très satisfaisante parce qu'elle permet en fait de segmenter et de déterminer la structure du marché de façon conjointe.

SS2-2 APPROCHES DE TYPE DETERMINISTE

Il y a alors deux voies principales pour mesurer la recherche de variété :

SS2-2-1 MESURE DIRECTE

Certains auteurs ont proposé et développé des échelles de mesure de la recherche de variété telle qu'elle est reportée par l'individu lui-même, lors d'enquêtes. Zuckerman (1979) a développé une échelle de mesure de la recherche de sensations (sensation Seeking Scale), mesurant la tendance au comportement varié. Elle comporte quatre dimensions :

- la recherche d'aventure
- la recherche d'expérience
- la recherche du défendu
- la recherche du non routinier

En fait, cette mesure caractérise l'individu lui-même et non son comportement par rapport à une catégorie de produits (ou une marque). Ainsi, elle mesure plus la tendance

générale au comportement varié que la recherche de variété au sens où on l'entend en marketing, c'est à dire relativement à un produit (ou une catégorie de produit) particulier .

S2-2-2 MESURE HYBRIDE

Ce type de mesure fait appel à deux types d'informations :

- informations sur les perceptions, attitudes et préférences des consommateurs
- informations relatives aux comportements (les historiques d'achats)

Deux grandes approches coexistent :

1 - Une approche que nous qualifierons de statique (la notion de satiété inter-temporelle n'apparaît pas explicitement). La recherche de variété est alors estimée:

-à partir de l'historique des achats effectifs (comportements)

-à partir des distances perceptuelles entre des marques appartenant au même sous-marché (il faut alors considérer que le sous-marché est défini sans ambiguïté).

Pessemier et Handelsman (1984) ont développé une telle approche avec une mesure appelée Index of Temporal Variety (voir en Annexe 1, voir aussi Handelsman 1982, cité dans Pessemier et Handelsman 1984).

Ce type d'approche convient lorsque le comportement de recherche de variété s'exerce à un moment donné du temps. Deux cas de figure peuvent alors se présenter en fonction des conditions du choix (voir McAlister 1979).

a) Plusieurs items sont choisis et consommés. C'est l'exemple du consommateur faisant un choix de boissons aux fruits aux parfums différents.
b) Plusieurs items sont choisis mais un seul sera consommé. C'est le cas d'un étudiant effectuant le choix entre plusieurs écoles de commerce. Parmi celles qu'il aura choisies et auxquelles il intègrera, il retiendra en fin de compte celle qui lui apporte le plus d'utilité.

Dans ces deux cas, l'hypothèse de base des modèles de préférence classiques, à savoir qu'un choix particulier est indépendant des précédents, ne tient plus. Les sélections successives d'items sont dépendantes les unes des autres.

Le premier cas (a), fait appel à un modèle de satiété des attributs pour prévoir les préférences. La préférence pour un groupe de produits g, particulier, est définie comme la somme, pondérée par le poids des attributs, des différentiels des inventaires d'attributs par rapport à l'état idéal souhaité.

$$AS(g) = \sum_{k=1}^{K} w_k (x_{gk} - x_k)^2$$

AS(g) est l'utilité assignée par le modèle au groupe de produits g
w_k est le poids de l'attribut k dans l'évaluation du groupe de produits g
x_{gk} est la somme des valeurs de l'attribut k obtenue par les produits du groupe g
x_k est le niveau idéal attendu sur l'attribut k

Le deuxième cas (b) où l'individu ne doit finalement consommer qu'un seul produit, fait appel à un modèle de "loterie" qui calcule l'espérance de l'utilité d'un groupe particulier d'éventualités. Avec ce type de modèles, le consommateur ne choisit pas les produits (marques) individuellement, par ordre d'utilité décroissante. Il choisit en fait des sous-groupes de marques maximisant son utilité, sachant que l'espérance de l'utilité de chaque choix n'est pas indépendante des autres choix possibles. C'est ainsi qu'un étudiant en gestion pourra associer dans son choix une école de haut niveau d'utilité pour lui mais pour laquelle il anticipe une faible probabilité de réussite et une école de faible utilité à laquelle il associe par contre une forte probabilité de réussite.

2 - Une approche dynamique dans laquelle apparait explicitement la notion de satiété par rapport à un comportement (Jeuland 1978) ou par rapport à l'inventaire accumulé dans le temps d'un attribut (voir le modèle proposé par McAlister 1982, en Annexe).

En résumé, nous dirons que le concept de recherche de variété peut être étudié et mesuré de manières très différentes selon :

- la définition qui lui est donnée: étendue ou étroite
- les hypothèses formulées quant au comportement, qui peut être vu comme stochastique ou déterministe.
- le type de données utilisées: comportement / jugement / mélange des deux
- l'optique, qui peut être dynamique (différents produits consommés au cours du temps: variété temporelle) ou statique (un ensemble de produits différents à un instant donné)
- le niveau d'analyse, individuel ou agrégé. Les modèles proposés par Lattin et McAlistair procèdent au plan individuel. Leur apport essentiel est qu'ils permettent de comprendre le comportement de recherche de variété d'un consommateur placé dans une situation de choix. Dans les modèles comme ceux qui ont été proposés par Kahn et al (1986), Givon et Horsky (1979), l'objectif est moins de comprendre et expliquer le comportement individuel que de déterminer le type de modèle de choix prévalant sur un marché et de caractériser ce marché par un niveau particulier de recherche de variété.
- l'objectif de l'étude de la recherche de variété qui peut être sa mesure, l'explication de sa formation et ses conséquences sur les préférences et le choix (approche explicative) ou simplement son intégration en tant que paramètre permettant de spécifier la forme d'un modèle (approche stochastique).

SS3 IMPLICATIONS POUR L'ANALYSE DE LA STRUCTURE DES MARCHES

Après avoir présenté les principales approches de la mesure de la recherche de variété, il est maintenant intéressant d'en prendre en compte les effets pour l'étude de la structure d'un marché. Nous avons déjà vu que l'analyse de la structure d'un marché peut être considérée, d'un point de vue méthodologique, comme étant constituée de trois phases:

- la définition du marché étudié, c'est à dire l'univers concurrentiel
- l'étude et la mesure des niveaux de relations entre les marques. Cette relation peut prendre principalement deux formes (voir Aurier 1990) qui sont la substituabilité et la concurrence. La substituabilité se définit au niveau de l'individu, à partir de ses jugements. La concurrence se définit au niveau du marché, à partir des comportements des acheteurs d'une catégorie de produits. Une des mesures de la concurrence les plus classiques en marketing est le transfert entre deux marques, d'une occasion d'achat à une autre.(voir Lehmann 1972, Kalwani et Morrison 1977, Rao et Sabalava 1981, Elrod 1988, Aurier 1990.
- l'analyse de la structure contenue dans les mesures effectuées lors de la phase précédente et qui se matérialise sous la forme de cartes (représentations spatiales) ou de graphes (représentations non-spatiales)

La recherche de variété doit être prise en compte dans la deuxième phase, quand est mesuré le niveau des relations existant entre les marques. En effet, le concept de recherche de variété nous paraît important pour au moins trois raisons, dans le cadre de l'étude du phénomène concurrentiel. :

1 - L'analyse de la structure des marchés à l'aide des historiques d'achats des consommateurs (données de comportement) fait intervenir des modèles stochastiques dont la forme dépend de l'existence et du type de recherche de variété caractérisant un marché ou un sous-marché particulier. Par exemple, en présence de comportement de recherche de variété (recherche de variété non nulle), l'utilisation d'un modèle stationnaire d'ordre 0 semble peu adaptée. Ainsi, la présence d'un type particulier de recherche de variété dictera le type de modèle à utiliser par la suite pour déterminer (ou tester) la structure du marché étudié. Par exemple, dans le modèle de nouveau produit PRODEGY (Urban, Johnson et Hauser 1984), le modèle ACRM (Aggregate Constant Ratio Model) fait référence à un processus stationnaire d'ordre O et donc, même si ce n'est pas explicité par les auteurs, à l'absence (ou à la non prise en compte) de la recherche de variété.

2 -La recherche de variété doit être prise en compte pour mesurer le lien de substituabilité existant entre deux marques. En effet, lors de l'étude du lien entre deux marques à partir des comportements (c'est à dire la concurrence), le fait que deux marques soient achetées successivement ne signifie pas pour autant que ces deux marques sont substituables. Au contraire, si le consommateur les achète pour satisfaire son besoin de recherche de variété, nous pourrons en conclure qu'elles sont complémentaires.

Nous retiendrons de cette analyse que deux marques peuvent être effectivement en concurrence sur un marché sans pour autant être substituables, dans la mesure ou elles ont été achetées pour satisfaire des besoins complémentaires liés à la recherche de variété. Pour une relation de concurrence donnée, mesurée au niveau du marché (à partir des comportements), peuvent ainsi, en fonction de la présence ou de la non-présence de recherche de variété, correspondre respectivement deux types de relations dans l'esprit du consommateur: de la complémentarité ou de la substituabilité (voir tableau 3 suivant).

TABLEAU 3 EFFET DE LA RECHERCHE DE VARIETE SUR LA NATURE DE LA RELATION ENTRE DEUX MARQUES

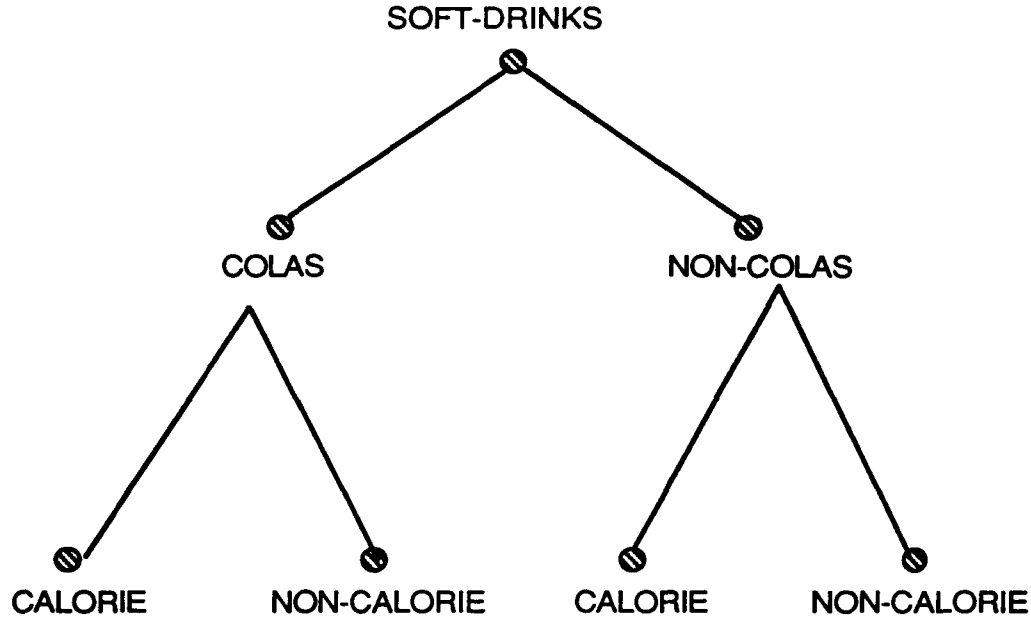
	MARCHE (Comportements)	CONSOMMATEUR (Jugements)	
		Présence de Recherche de Variété	
		OUI	NON
Nature de la Relation entre deux Marques	CONCURRENCE	COMPLEMENTARITE	SUBSTITUABILITE

Givon (1984), Lattin et McAlister (1985), ont proposé deux modèles d'analyse de la concurrence prenant en compte la recherche de variété qui caractérisent ce phénomène. Ces modèles sont tous les deux estimés à partir des historiques d'achats des consommateurs.

Dans le modèle de Givon, pour chaque consommateur, un niveau de recherche de variété caractérise chaque attribut de la hiérarchie de choix. Les attributs étudiés, pour une catégorie de produits particulière, sont de type *"physique et discret"* (par exemple gazeux/non-gazeux pour le marché des softs drink). La hiérarchie de choix d'un individu peut être représentée sous la forme d'un arbre de décision (voir la figure 9 suivante). Le marché est successivement partagé en sous-marchés sur la base des attributs considérés. Par exemple, sur la figure 9, le marché est d'abord partagé en deux sous-marchés sur la base du caractère cola ou non-cola des produits. Ces deux sous-marchés sont ensuite partagés sur la base du caractère calorique ou non-calorique des produits.

Pour l'estimation des paramètres, Givon (1985) a utilisé la modélisation de la recherche de variété que nous avons présentée dans la section précédente (Givon 1984).et qui est développée à la section SS3.2.1.

FIGURE 9 ARBRE INDIVIDUEL DE DECISION ET STRUCTURE D'UN MARCHE



Dans le modèle proposé par Lattin et McAlistair (1985), la recherche de variété est considérée globalement, par rapport à l'ensemble des bénéfices apportés par une marque, en comparaison de ce qu'a apporté la marque choisie lors de l'occasion d'achat précédente. La recherche de variété caractérise un l'individu relativement à un marché particulier et non pas cet individu attribut par attribut, comme dans le modèle de Givon. L'achat d'une marque i, à l'occasion t-1, diminue la probabilité d'achat d'une marque similaire à i, à l'occasion t. La diminution de la probabilité est d'autant plus forte que les deux marques sont similaires (nous dirons substituables) dans l'esprit du consommateur. Ce modèle sera développé en détail à la section SS3.2.2.

Le résultat essentiel que nous pouvons retenir de ces deux approches est le lien qui existe entre la recherche de variété et la mesure de la concurrence. Beaucoup d'auteurs ont proposé de mesurer la concurrence entre deux marques à partir des transferts entre marques (transferts observés sur des historiques d'achats des consommateurs). Si le transfert entre deux marques est important, il en est déduit généralement que le lien concurrentiel entre ces deux marques est fort. Ceci ne signifie pas pour autant que les deux marques sont substituables dans l'esprit du consommateur. En effet:

- si l'individu manifeste un niveau de recherche de variété important, les deux marques sont achetées parce qu'elles apportent des bénéfices différents au consommateur, donc qu'il les juge complémentaires.
- si l'individu ne manifeste pas de recherche de variété, nous pourrons dire que ces deux marques sont substituables.

3 - La situation d'usage du produit est une variable dont l'influence sur les comportements d'achats et la structure des marchés est très importante. Elle a donné lieu à un courant de littérature analysant la structure des marchés à partir de la situation d'usage

des produits (voir Shocker et Srivastava 1979, Srivastava, Leone et Shocker 1981, Srivastava,Alpert et Shocker 1984, Srivastava 1980 et 1981). Il apparaît (voir McAlister et Pessemier 1982) que la situation d'usage serait en fait une cause indirecte de recherche de variété, un consommateur achetant différentes marques pour satisfaire des besoins différents nés de situations d'usage différentes. Il est important de faire la part entre ce qui est dû à la recherche de variété de ce qui est dû à la situation d'usage. Suivant McAlister (1982), nous dirons qu'il n'est pas exact de considérer qu'il y a Recherche de Variété lorsque celle-ci n'est que l'expression indirecte d'une autre motivation née par exemple de la situation d'usage du produit (comportement de variété dérivée).

SECTION 3 THEORIE DES CHOIX ET ANALYSE DE LA STRUCTURE DES MARCHES

La théorie du choix des consommateurs tient beaucoup de place dans la recherche concernant la mesure de la concurrence entre les marques et l'analyse de la structure des marchés. Selon certains auteurs (voir Rao et Sabavala 1981, Givon 1985, Grover et Srinivasan 1987, Shocker et al 1987), la structure d'un marché serait, ou pourrait être, l'expression du processus de choix suivi par le consommateur pour effectuer ses choix. D'autre part, le calcul et l'estimation d'indices de la concurrence sont souvent fonction du type d'hypothèses qui sont formulées quand on modélise le choix du consommateur. Pour ces raisons, nous pensons qu'il était indispensable, dans un chapitre consacré aux concepts de base, de présenter les principaux courants de pensées relatifs au comportement du consommateur.

Comme c'est le cas dans d'autre domaines (notamment celui de la statistique multivariée), le Marketing a trouvé ses bases théoriques au sein de deux disciplines, l'Economie et la Psychologie.

Selon l'approche économique, le consommateur est considéré comme rationnel. Il effectue ses choix parmi un ensemble d'alternatives décrites en termes d'attributs physiques. Ce type de choix ont été modélisés par Lancaster (1966).

Dans l'approche psychologique, le consommateur est considéré comme étant irrationnel. En tant qu'"individu", il est l'objet de l'investigation. Sa caractéristique principale est de percevoir les objets qui l'entourent en fonction de ses propres caractéristiques mais aussi de son environnement (culture, famille, médias, voir Baudrillard 1968).

Bass (1974), a posé le problème de savoir si la préférence (telle qu'elle se révèle à partir des comportements) devait être considérée comme stochastique ou déterministe. Il aboutit à la conclusion que le comportement est en réalité déterministe, mais fonction d'un tel nombre de variables, qui de plus se manifestent avec une fréquence imprévisible, qu'il est plus réaliste de considérer le comportement comme stochastique. La conclusion de l'auteur est la suivante :

- si l'objectif est d'expliquer un comportement, alors l'approche doit être déterministe.
- si l'objectif est de prévoir le comportement, alors l'approche doit être stochastique.

Ce point de vue est aussi celui de Pessemier (1978). Ainsi coexistent en marketing, deux grands types d'approches :

- une approche déterministe, où le comportement du consommateur est vu en terme de "processus" dont on essai de déterminer les phases principales. Ces approches sont généralement individuelles puisque l'objet de l'étude est plus la compréhension de l'individu que son comportement (approche du psychologue).
- une approche dite "stochastique" dont l'objectif essentiel est la prévision des parts de marché des marques, des probabilités de choix et des probabilités de transferts d'une marque vers une autre. Les approches stochastiques sont en général de type agrégé car

l'aléa qui réside au plan individuel tend à disparaître (par compensation) au plan agrégé (voir Bass 1974).

En fait, il est important selon nous, de faire la distinction entre deux familles de modèles souvent regroupées sous la même dénomination, ce qui est la source de confusions :

- les modèles stochastiques
- les modèles probabilistes

Dans les premiers, le comportement est modélisé comme étant en soi probabiliste. La réalisation du comportement est caractérisée par un ensemble de probabilités.

Dans les seconds, l'aléa statistique représente une différence entre les prévisions et les réalisations ou une erreur de la mesure. L'introduction de l'aléa correspond plus, à notre avis, à une mise en forme statistique du problème permettant son traitement à l'aide des méthodes statistiques disponibles (par exemple modèle linéaire général, modèle multinomial probit). L'aléa est plus utilisé comme un aveu d'incapacité à prévoir correctement le comportement que pour signifier que le comportement est stochastique.

Nous présenterons donc ici trois grandes familles de modèles de choix en montrant notamment comment elles évoluent et tendent à converger les unes vers les autres :

- les modèles stochastiques de choix
- les modèles probabilistes de choix
- les modèles déterministes de choix

SS1 MODELES STOCHASTIQUES DE CHOIX

Ce type de modèle remonte au milieu des années cinquante. Ils ont été construits pour satisfaire les besoins du gestionnaire, à savoir essentiellement la prévision des achats et des parts de marché. Cette prévision s'effectue manifestement sans intérêt pour le comportement du consommateur. Dans cette approche, les modèles ont pour objectif de reproduire au mieux la réalité et non de l'expliquer. Comme le notent Massy, Montgommery et Morrison (1970) dans l'introduction de leur ouvrage consacré aux modèles stochastiques, "on observe le comportement sans se poser de questions sur les causes de ce comportement".

La force de ce type d'approches est qu'elles prévoient bien. Cependant, lorsque l'environnement se modifie, ces méthodes ne sont ni capables d'analyser les causes du changement, ni capables de prévoir correctement. D'autre part, les chercheurs et praticiens ont observé que prévoir ne suffit pas toujours en gestion, il faut aussi savoir comprendre et expliquer. C'est le cas par exemple lorsqu'une entreprise souhaite lancer un nouveau produit.

En terme de comportement de choix du consommateur, l'hypothèse de base de ce type de modèle est de considérer le comportement comme stochastique. L'argument majeur invoqué pour défendre ce point de vue (voir par exemple Jeuland 1979) est que les consommateurs ont des préférences révélées (les parts de marché des marques qu'ils achètent) qui sont stables, mais ils changent cependant de marque. Notons au passage qu'un tel comportement de changement apparent pourrait s'expliquer autrement, par la

diversité des goûts, des individus qui composent le ménage, par la diversité des situations d'usage du produit, par la recherche de variété du consommateur ..

En simplifiant, nous pouvons dire qu'il existe trois grands types de modèles stochastiques de choix :

- les modèles d'ordre 0
- les modèles à deux états
- les modèles d'apprentissage

SS1-1 MODELES D'ORDRE ZERO

Dans un tel modèle, le consommateur a une probabilité constante d'acheter la marque i, à chaque occasion d'achat. La caractéristique principale d'un tel modèle est de considérer qu'il y a indépendance entre le comportement présent et le comportement passé. Au niveau individuel ou agrégé, les préférences des consommateurs sont décrites par un vecteur de probabilités de choix des marques (préférences révélées) : P = {p1, p2, ...pn}
Lorsque le choix se limite à deux alternatives : achat de i ou non achat de i, il s'agit d'un processus de Bernoulli à deux états. Lorsque plus de deux états sont pris en compte, il s'agit alors d'un processus multinomial. L'agrégation des choix individuels en choix agrégés, pour prendre en compte l'hétérogénéité des consommateurs, doit se faire en donnant une loi statistique au paramètre individuel p(i). A l'origine, la loi utilisée était la loi béta. Dans ce modèle, les probabilités conditionnelles de choix sont égales aux probabilités marginales.

$$P(i/j) = P(ij)/P(j)$$
$$= P(i).P(j)/P(j) = P(i)$$

P(i/j) est la probabilité de choix de la marque i à la période t sachant que la marque j a été achetée à la période t-1 .
P(ij) est la probabilité de choix successif des marques i et j
P(i), P(j) sont les probabilités marginales de choix des marques i et j

Les probabilités P(i) sont estimées à partir de données de panels. L'estimateur de maximum de vraisemblance de P(i): P*(i) est : P*(i)= ni/n

ni est le nombre d'achats de la marque i sur la période
n est le nombre total d'achats sur la période, concernant l'ensemble de choix étudié

SS1-2 MODELES DE MARKOV A DEUX ETATS

Le choix d'une marque à la période t est ici fonction du choix effectué à la (ou les) périodes précédentes. Dans un modèle de markov d'ordre un, l'achat en t est fonction des choix effectués en t-1. Dans un modèle de markov d'ordre deux, l'achat en t est fonction des choix effectués en t-2 et t-1, etc ...

A partir du vecteur de préférences (stationnaire) P {p1, p2 ...pn}, il est possible de calculer les probabilités conditionnelles d'achat d'une marque i sachant que la marque j a été achetée à la période précédente. L'estimation du maximum de vraisemblance de p(i/j) est donnée par : $P^*(i/j) = nij/ni$

nij est le nombre d'achats de la marque j à la période t-1, précédent un achat de la marque i à la période t
ni est le nombre d'achats de la marque i en t

Il faut remarquer que P(i/j) constitue une mesure intéressante de la concurrence qui s'exerce entre les marques i et j. Elle a notamment l'avantage de ne pas pouvoir être négative. Comme précédemment, l'agrégation sur les individus peut se faire en donnant une loi béta aux paramètres individuels.

SS1-3 MODELE D'APPRENTISSAGE

Le modèle d'ordre 1 constitue un système dans lequel l'individu fait un apprentissage très rapide de ses achats antérieurs puisqu'il ne tient compte, lors de son choix à la période t, que du choix précédent. Le modèle d'apprentissage linéaire (Linear Learning Model) a été proposé par Kuehn (1958, voir dans Massy et al 1970, Givon et Horsky 1979). Il peut s'exprimer de la manière suivante :

$$Pr[i_{t+1}/X_{it}] = a + b.X_{it} + c.P_{it}$$

i_{t+1} est l'événement "acheter la marque i à l'occasion t+1"
$P_{it} = Pr(it)$ est la probabilité d'achat de la marque i à la période t
$X_{it} = 1$ si la marque i a été achetée en t
$X_{it} = 0$ autrement
b caractérise la dernière expérience avec la marque
c est un coefficient d'apprentissage
$a+b+c \geq 0$ et $a+b+c \leq 1$

La probabilité qu'un consommateur achète une marque est fonction des expériences favorables et défavorables passées avec cette marque. Avec le modèle d'apprentissage, le comportement est considéré comme "*adaptatif*", ce qui correspond au cas le plus général. Les modèles d'ordre 0 et d'ordre 1 sont en fait des cas particuliers de ce modèle.

Si l'individu n'apprend rien, a = b = 0 et c = 1, on retrouve alors un modèle d'ordre zéro.

S'il apprend très vite, c = 1, on retrouve le modèle de Markov d'ordre un.

Malgré son intérêt conceptuel, ce modèle a été peu étudié. Nous citerons toutefois une application de Carman (1966) dans laquelle l'auteur définit un opérateur de gain en part de marché et un opérateur de pertes en part de marché ainsi que des probabilités d'achats conditionnelles aux périodes d'achats des marques.

Givon et Horsky (1978, 1979) ont présenté une utilisation du modèle d'apprentissage et montré qu'il est mieux adapté que les modèles d'ordre zéro et d'ordre un pour de

nombreuses catégories de produits. Il est à regretter que de tels travaux n'aient pas été poursuivis.

Les modèles d'ordre zéro ou d'ordre un ont été les plus utilisés, notamment dans le cadre de la prévision des parts de marché (modèles de prévision de parts de marché).

Les distributions statistiques permettant de gérer la loi du paramètre Pi (probabilité d'achat de la marque i) pour un individu ainsi que la distribution de ces valeurs au sein de la population toute entière ont été affinées au cours des années. Nous citerons la loi binomiale négative (NBD) d'Ehrenberg (1972), la loi Multinomiale Dirichlet (voir Jeuland, Bass et Wright 1980, voir aussi Schmittlein, Bemmaor et Morrison 1985 pour une présentation de ces lois et de leurs conditions d'utilisation). Un tel effort pour la recherche de lois permettant de décrire le comportement d'achat d'un individu aléatoire peut s'expliquer par l'objectif essentiel de ces modèles qui est d'effectuer des prévisions au plan agrégé. De la pertinence de la loi utilisée, dépendent la qualité des agrégations prenant en compte l'hétérogénéité individuelle et donc en fin de compte la qualité des prévisions de parts de marché.

Ces modèles permettent ainsi d'effectuer des prévisions précises des parts de marché des marques, au plan agrégé, ceci à partir des comportements passés. Cependant, nous citerons deux points qui constituent leur faiblesse :

1 - Ces modèles ne sont en rien explicatifs du comportement des consommateurs. De plus ils ne font pas de lien et ne permettent pas d'expliquer l'effet des variables du mix, ce qui limite ainsi leur qualité en tant qu'outil de diagnostic des actions menées. Nous avons relevé cette critique dans un article de Blattberg (1981) consacré à l'évaluation des modèles stochastiques de parts de marché. Nous citerons cependant quelques tentatives qui ont été faites dans ce sens. Jeuland (1979) a proposé un modèle qui permet de prendre en compte la disponibilité de la marque (distribution). Bass et Pilon (1980) ont proposé un modèle qui prend en compte les variables du mix. Ils ont notamment montré que les transferts entre marques dépendent de ces variables.

Dans le même objectif d'élargir la portée des modèles stochastiques de choix, Marley (1981) a proposé un modèle stochastique de choix qui est compatible avec le modèle d'éliminations par les aspects (EBA) de Tversky et prend en compte les caractéristiques des produits.

2 - Ils utilisent généralement l'hypothèse de stationnarité des préférences (le vecteur P {p1 ..pn}) sur la période d'étude. Cette hypothèse, pour rester tenable, nécessite de n'étudier que des marques dont la fréquence d'achat est suffisamment courte (voir Bass 1974). Cette hypothèse est une conséquence de la modélisation "boîte noire" du consommateur, à partir de ses seuls comportements passés. Puisque le modèle n'explique pas, il est obligé de se placer dans un contexte particulièrement stable.

Or, le lancement d'une nouvelle marque, une nouvelle campagne de communication sont des actions qui peuvent modifier sensiblement l'environnement, rendant alors les modèles inopérants. Ces modèles sont surtout efficaces pour effecteur des prévisions au niveau agrégé, lorsque le marché est en équilibre (stationnaire).

Le choix de l'ordre du modèle (zéro, un ou plus) a donné lieu à de nombreuses discussions. Massy et al (1970, chapitre 3) reprenant Frank (1962), ont montré qu'un modèle d'ordre zéro peut passer pour un modèle d'ordre un pour au moins deux raisons:

- parce que l'hypothèse de stabilité des préférences ne tient pas

- parce des individus qui suivent un modèle d'ordre zéro mais avec des probabilités d'achat des marques différentes ont été agrégés (voir aussi Givon et Horsky 1978).

Bass, Givon, Kalwani et al (1984) ont montré que dans une grande majorité de cas, les comportements sont de type stationnaire. Lorsque les processus sont stationnaires, les processus de choix individuels sont essentiellement d'ordre zéro. Ceci explique (selon les auteurs) pourquoi les modèles d'ordre zéro fonctionnent assez bien au plan agrégé.

Givon et Horsky (1978 et 1979) ont montré que le modèle d'ordre un ne fonctionne sur aucun marché, que le modèle d'ordre zéro fonctionne bien pour le marché des médicaments et que le modèle d'apprentissage fonctionne bien sur plusieurs marchés. Ils ont proposé un modèle permettant de prendre en compte (lors de l'agrégation) la possibilité que les individus suivent des processus d'ordre différents (zéro, un ou apprentissage). Ce modèle fonctionne bien pour une majorité de biens de grande consommation. Jones et Zufryden (1982, voir dans Bass, Givon et al 1984) ont montré que l'ordre d'un modèle peut être influencé par les variables du mix marketing, notamment la publicité.

Enfin, un article de Kahn, Morrison et Wright (1986) a fait apparaître une des faiblesses des données de panels dans lesquels les différents individus composant le ménage sont agrégés *"de fait"* . Les auteurs ont montré (à l'inverse de Franck 1960), que l'agrégation d'individus qui ont des modèles de choix d'ordre 1 tend vers un modèle d'ordre 0 . Ceci serait d'autant plus vrai que la taille du ménage est importante.

Ces résultats constituent-ils un frein à l'utilisation des modèles d'ordre zéro ? Tout dépend en fait du niveau d'analyse fixé. Si nous considérons que le ménage représente une unité d'observation à part entière qui mérite d'être étudiée en tant que tel, alors la revue de littérature précédente justifie dans la majorité des cas des biens de consommation courante l'utilisation des modèles d'ordre zéro.

Les modèles stochastiques de choix et les hypothèses sur l'ordre de ces modèles, ont une place privilégiée dans les modèles d'analyse de la concurrence. Beaucoup de méthodes reposent sur les théorèmes développés dans les modèles stochastiques de parts de marché. Notamment, le modèle de Hendry (voir Kalwani et Morrison 1977) repose sur les hypothèses d'un modèle d'ordre zéro. Les modèles qui prennent en compte la recherche de variété (voir section 2 et voir Kahn Kalwani et Morrison 1986) sont des modèles d'ordre un ou plus.

Les résultats précédents tendent à indiquer l'utilisation de données de transferts "agrégées" lors de l'utilisation d'un modèle d'ordre zéro.

Nous allons présenter maintenant l'approche économique "classique" du comportement de choix du consommateur, approche que nous intitulons "modèles probabilistes de choix".

SS2 MODELES DE CHOIX PROBABILISTES

Les économistes ont longtemps considéré le produit et le consommateur (homo aeconomicus) comme des unités homogènes non différentes et non différenciées. Le consommateur est notamment considéré comme une boîte noire qui intègre de l'information "objective et complète" en provenance des marchés (prix, caractéristiques des produits, contraintes de budget ..). Il est conçu comme rationnel, homogène puisqu'il se trouve dans une situation d'information parfaite (tous les consommateurs disposent de la même information). Il effectue ses choix de manière à optimiser son utilité, par rapport à un ensemble d'alternatives et sous la contrainte de la rareté de ses ressources (voir Micallef, 1982). Dans l'introduction de son article de 1980 intitulé *"The choice theory approach to market research"*, Daniel McFadden résume bien ce point en qualifiant le consommateur vu par les économistes de *"boite noire optimisante"*. Un des premiers auteurs à avoir remis en cause cette vision et proposé une autre formalisation du comportement du consommateur est Thurstone (1927), qui d'ailleurs n'était pas économiste mais psychomètre. Dans le modèle de Thurstone, l'utilité que procure une alternative à un individu est fonction de ses caractéristiques mais aussi d'un élément aléatoire qui prend en compte les biais perceptuels de cet individu, son information imparfaite, l'effet des variables présentes à la situation de son choix ..

Ainsi l'utilité U_j pour l'alternative j peut s'écrire : $U_j = V_j + \varepsilon_j$

V_j est la composante déterministe de l'utilité

ε_j est la composante stochastique de l'utilité

Repris ensuite par les travaux de Luce (1959) dans son célèbre axiome et Marshak (1960, cité dans McFadden 1986), le modèle d'utilité du consommateur peut alors être exprimé de la manière suivante :

Soient :

E_j = 1 à J, l'ensemble des alternatives de choix

Chaque alternative j est caractérisée par son vecteur d'attributs Z_j et son utilité aléatoire $U(Z_j)$

La probabilité de choix de l'alternative i : P(i/Z) s'écrit:

$$P(i/Z) = P[\, u(Z_i) \geq u(Z_j)\,] \text{ pour tout j appartenant à Ej}$$

$$P(i/Z,\pi) = V(Z_i,\pi) / \sum_{j=1}^{J} V(Z_j,\pi)$$

$V(Z_i,\pi)$ est la valeur scalée de l'utilité procurée par l'alternative i, au facteur d'échelle π près.

En estimant après transformations par les logarithmes des $V(Z_i,\pi)$ et en posant que les termes εi suivent tous une loi "exponentielle négative double", Mc Fadden à formulé

(1965, voir dans McFadden 1986) le modèle Multinomial Logit (MNL) qui a connu depuis un succès très important.

Une des limites essentielles de ce modèle tient à ce que, de par la forme de la fonction d'utilité, l'utilité de chaque alternative est fonction de ses seules caractéristiques et ne tient pas compte des caractéristiques des J-1 autres. Cette hypothèse est appelée "Indépendance des Alternatives non Pertinentes" (IIA, Independence of irrelevant alternatives). Elle peut être formulée de la manière suivante:

soit E_j l'ensemble des alternatives du choix
soit S_j un sous ensemble de E_j

$$\frac{PS(i)}{PS(j)} = \frac{PE(i)}{PE(j)}$$

PE est la probabilité de choix dans l'ensemble E_j
PS est la probabilité de choix dans l'ensemble S_j

Cette propriété signifie que le rapport des probabilités de choix de deux alternatives ne dépend pas de l'ensemble dans lequel s'inscrivent ces deux alternatives.

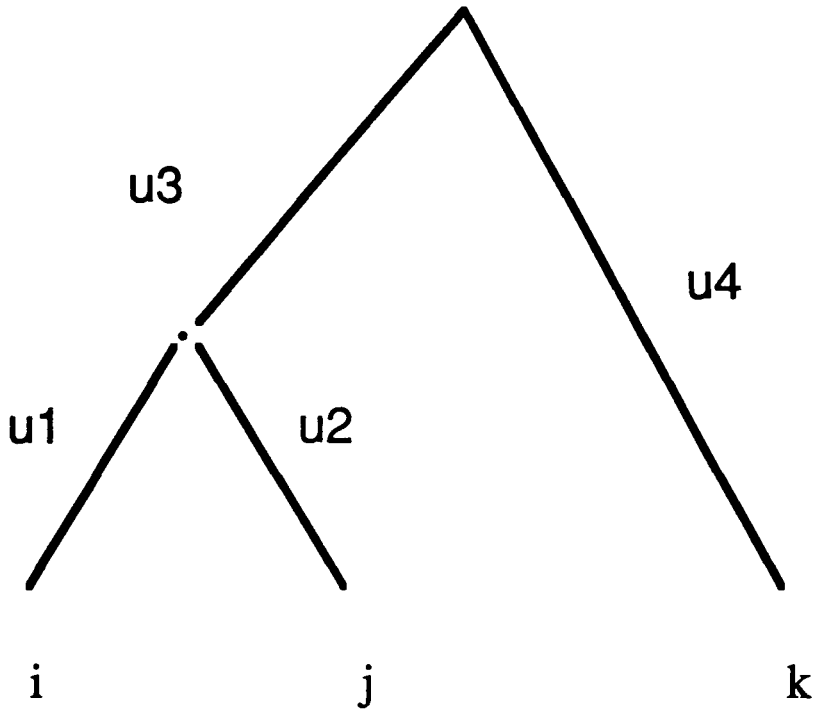
Cette hypothèse à des conséquences extrêmement importantes dans le cadre de la théorie des choix et de la mesure de la concurrence entre les marques. Elle signifie en effet qu'au sein d'un marché, lorsqu'une marque est introduite, elle constitue sa part de marché proportionnellement aux parts des marques déjà présentes sur le marché. Soient deux marques A et B aux parts de marché de 90% et 10% . La part de marché d'une troisième marque C introduite sur le marché sera constituée à 90% d'anciens acheteurs de A et à 10% d'anciens acheteurs de B. Cette hypothèse n'est tenable que lorsque les trois marques A, B et C présentent des niveaux de substituabilité comparables. Si tel n'est pas le cas, la marque C fortement substituable à B prendrait plus de 10% de sa part de marché sur les acheteurs de B.

Tversky (1972), et Tversky et Sattath (1979) ont proposé un modèle alternatif qui ne repose pas sur l'hypothèse des IIA. Il s'agit du modèle d'*élimination par les aspects* (EBA :Elimination By Aspects, ou encore HEBA : Hierarchical Elimination By Aspects).

Ce modèle a connu en fait, en marketing, essentiellement un succès d'estime. Son utilisation est restée très limitée, ceci pour la raison essentielle qu'il nécessite une procédure d'estimation complexe. Dans ce modèle, le consommateur effectue ses choix selon une procédure hiérarchique. Chaque alternative de choix est décrite selon ses caractéristiques ou aspects. Pour aboutir à un choix, le consommateur sélectionne un aspect (caractéristique) au hasard. Il élimine les alternatives qui ne possèdent pas cet aspect. Il sélectionne ensuite un autre aspect, toujours au hasard et élimine parmi les alternatives restantes, celles qui ne possèdent pas la caractéristique. Le processus est reconduit jusqu'à l'obtention d'un choix unique.

Ce processus peut être représenté sous la forme d'un arbre hiérarchique de choix dans lequel les "*noeuds*" représentent les étapes possibles du choix et la longueur des branches, l'utilité relative du choix. Avec trois alternatives i, j et k on peut obtenir l'arbre suivant :

EXEMPLE: ARBRE HIERARCHIQUE DE CHOIX



Sur cet exemple, les marques i et j sont plus ressemblantes entre elles qu'elles ne le sont avec la marque k. L'utilité u3 liée au segment commun (caractéristiques communes) aux deux marques i et j permet de prendre en compte la similarité entre ces deux marques. Dans cet exemple, plus u3 est faible et plus l'hypothèse des IIA est tenable. A l'inverse, plus u3 est forte et plus i et j possèdent en commun une valeur d'utilité que ne possède pas k . Dans ce cas, l'hypothèse des IIA n'est plus tenable.

McFadden (1980) a proposé plusieurs alternatives permettant de prendre en compte cette limite sérieuse à son modèle MNL. L'auteur remarque toutefois que les conditions d'application du modèle MNL correspondent à la majeure partie des cas réels, ce qui nous parait contestable. L'auteur propose essentiellement deux alternatives pour se passer de l'hypothèse des IIA . La première solution consiste à découper le marché en sous ensembles homogènes de marques (sur quelle base ?) et utiliser des modèles Logit "nichés" au sein de chacun des groupes ainsi formés. L'autre solution consiste à utiliser le modèle à "*Valeurs Extrêmes*" (EVM) qu'il a proposé. Ce modèle nécessite l'estimation de très nombreux paramètres, ce qui limite le nombre de marques qu'il est possible d'analyser simultanément. Pour une présentation détaillée des modèles logit et de leur extension nous renvoyons aux ouvrages de Manski et McFadden (1981), Maddala (1983) et aux articles d'Amémiya (1981) et Hausman (1980).

L'hypothèse des IIA tient une place privilégiée dans les modèles de choix du consommateur et dans la mesure de la concurrence. Il n'est cependant pas dans l'objet de cette section de faire un état précis de ce débat. Nous renvoyons le lecteur à la section 5 du chapitre II ainsi qu'à Currim (1982) et Kamakura et Srivastava (1984 et 1986) pour un débat appliqué au domaine de la mesure de la concurrence entre les marques.

Les caractéristiques principales de ces modèles de choix du consommateur sont finalement les suivantes :

- ils considèrent en général le consommateur à son niveau agrégé. L'hétérogénéité des consommateurs est prise en compte par l'utilisation de distributions statistiques qui décrivent la loi des paramètres individuels sur l'ensemble de la population. Même si le modèle procède au plan individuel, l'objectif n'est pas d'obtenir des prévisions au plan individuel mais au plan agrégé (segment de consommateur ou marché total).

- ils utilisent des variables qui sont le plus souvent directement observables: caractéristiques objectives des individus (socio-démographiques), attributs physiques des produits etc.. Les techniques statistiques utilisées reflètent largement ce fait. Elle considèrent toutes que les variables du modèle sont exemptes d'erreur de mesure (MNL, Modèle linéaire général).

- leur objectif essentiel est la prédiction. Cependant, à la différence des modèles stochastiques de parts de marché ces modèles ont une ambition explicative nettement plus prononcée.

SS3 MODELES DE CHOIX DETERMINISTES

Ces modèles correspondent à l'approche la plus répandue en "Etudes Marketing", celle qui s'est centrée sur l'étude du comportement du consommateur. Le comportement est présenté comme le résultat d'un processus déterministe qui prend en compte les caractéristiques du consommateur, son environnement mais aussi ses perceptions, son mode de traitement de l'information, la formation de ses préférences et de ses choix.

La vision du comportement du consommateur est à l'opposé de celle qui est retenue dans les modèle stochastiques de choix, puisque celui-ci est considéré comme foncièrement déterministe et explicable. Avoir une approche déterministe du comportement c'est avant toutes choses vouloir comprendre et expliquer les raisons de ce comportement.

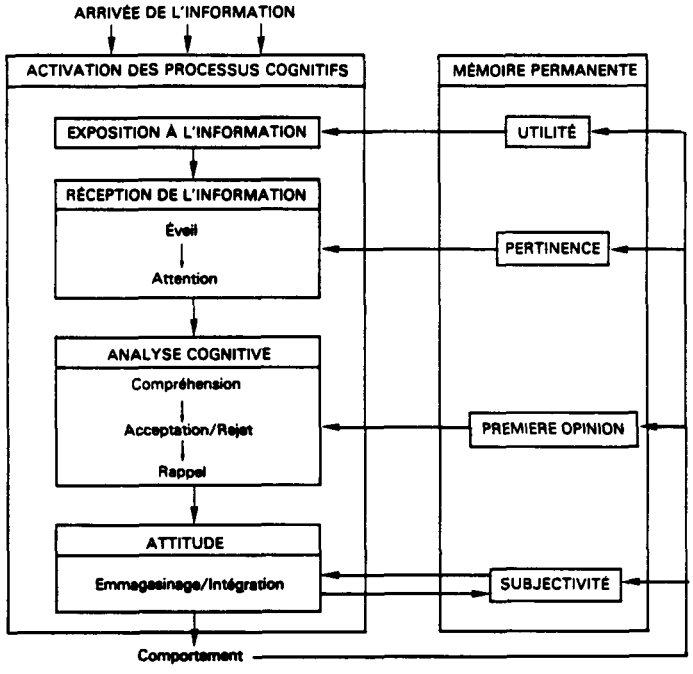
Les recherches en psychologie et leurs applications au comportement du consommateur ont montré qu'il s'agit d'un processus très complexe. De cette complexité découle deux faits :

- les approches déterministes prévoient assez mal le comportement du consommateur
- les approches déterministes ont beaucoup de peine à raisonner au plan agrégé et donc à pouvoir faire des prévisions à ce niveau (prévisions qui sont pourtant indispensables au manager). L'agrégation se fait à partir des moyennes des paramètres estimés au plan individuel, ce qui constitue une méthode impropre à prendre en compte l'hétérogénéité des consommateurs. Ceci est certainement une des raisons pour lesquelles le concept de segmentation a obtenu autant de succès en recherche marketing. L'agrégation au niveau de segments de consommateurs qui constituent en fait autant de consommateurs "types" (chaque segment étant constitué par un seul consommateur type) justifie des agrégations aussi réductrices que la moyenne. Nous pensons, pour notre part, que dans beaucoup de recherches, la constitution de segments est plus liée à cette insuffisance des modèles à prendre en compte l'hétérogénéité des consommateurs qu'à un réel intérêt managérial à segmenter le marché.

Un exemple qui reflète bien la vision en terme de "processus" de l'approche déterministe est le modèle de "protocole" développé par Bettman (1971 et 1979, voir section 5) qui présente une analyse des processus de choix des consommateurs très poussée. Cette analyse tente de rendre compte de tous les facteurs qui conduisent au choix. Par contre, de par sa complexité, elle est peu apte à effectuer correctement l'agrégation des individus.

Une autre approche de l'étude de la formation des préférences et des choix est issue de la psychologie cognitive où le consommateur est essentiellement conçu comme une "unité de traitement de l'information". Le modèle classique suivant, extrait de Dussart (1983, p144) illustre bien cette approche en terme de processus.

FIGURE 10 MODELE DE TRAITEMENT DE L'INFORMATION



Adapté de: Stemthal, B., Craig, S., *Consumer Behavior: An Information Processing Perspective*, Prentice-Hall, 1982, p. 68. Reproduit avec l'autorisation de Prentice-Hall Inc. Copyright © 1982.

Le processus de traitement de l'information.

"Extrait de Dussart 1983 p144"

La caractéristique commune à toute cette famille de modèles de traitement de l'information est qu'ils sont multi-attributs. L'individu utilise plusieurs attributs, successivement ou simultanément, afin de former son attitude qui elle-même détermine ses préférences.

De nombreux modèles de traitement de l'information "multi-attributs" ont été proposés (voir Pras 1977 pour une revue). Le plus connu de ces modèles est celui qui a été

proposé par Fishbein et adapté en Marketing par Wilkie et Pessemier (1973). L'attitude par rapport à une alternative a: A_a, s'exprime de la manière suivante :

$$A_a = \sum_{i=1}^{n} b_i.e_i$$

b_i est la force de la croyance que la marque a possède le ième attribut saillant.
e_i représente le degré de désir pour le ième attribut saillant

La particularité de ce modèle est qu'il considère la formation des attitudes et par extension des choix, comme un processus de type continu (compensatoire). L'attitude est une fonction linéaire des croyances relatives aux attributs perçus du produit. Une telle représentation de l'attitude est le résultat d'un processus d'agrégation "continu" des perceptions. Il est ainsi tout à fait compatible avec les méthodes statistiques qui ont été proposées pour analyser les perceptions et les préférences (Analyses Factorielles et Discriminantes, Techniques de Scaling).

Le type de modèle de choix qui résulte de la théorie du traitement de l'information est donc essentiellement "*spatial*", continu (voir Lynch 1985, Holbrook 1981, Dillon et al 1985). Le processus de choix (de préférence) est "*déterministe*", les marques choisies sont celles qui sont les plus proches des points idéaux de l'individu (modèle de Coombs 1964). Cette modélisation du comportement de choix s'oppose à celle qui est développée dans le modèle EBA de Tversky, modèle de type "*hiérarchique*", donc "*non continu et non spatial*"

Plus récemment, certain auteurs ont proposé une modélisation du traitement de l'information et des choix selon un processus à deux étapes (voir Gensch, 1987, pour une revue de littérature à ce sujet). Pras et Summers (1975) ont montré que lorsque toutes les alternatives de choix sont prises en compte, le modèle conjonctif fonctionne mieux que le modèle compensatoire. Par contre, avec les seules alternatives acceptables, c'est le modèle compensatoire qui fonctionne le mieux. Payne (1976), à l'aide de l'analyse des protocoles, a montré que le consommateur utilise un modèle de type EBA (hiérarchique) dans la première phase de son choix afin de le simplifier. Il utilise ensuite un modèle de type compensatoire pour aboutir à un choix final (voir Gensch 1987 pour une revue de littérature complète sur les modèles à deux étapes). Gensch a montré la supériorité empirique d'un modèle à deux étapes par rapport au modèle logit à une étape (les conditions de mise en oeuvre de ce modèle ne nous paraissent pourtant pas très fiables étant donné la nature du produit, des moyens de transports maritimes et le faible nombre d'alternatives).

Comme le note Blattberg (1981 p 202), le modèle de Hendry repose sur l'hypothèse d'un modèle de choix à deux étapes. La première (hiérarchique) consiste au choix d'une partition qui peut être effectué sur la base de la "*forme du produit*" (attribut physique) ou bien de la marque. La deuxième étape consiste au choix final, sur la base de la marque si dans l'étape précédente le consommateur a utilisé la forme, ou de la forme dans le cas inverse.

Certains auteurs, notamment Shocker et al 1987 ont émis l'hypothèse du lien qu'il pourrait y avoir entre "*processus de choix en deux étapes*" et "*analyse de la structure du marché*".

La première phase du processus de choix, serait à mettre en parallèle avec une structure du marché de type hiérarchique dans lequel le marché se découpe en sous-marchés. Cette structure serait essentiellement le fait d'attributs de type "*physique*" et à "*niveaux discrets*". Les modèles Hendry (Kalwani et al 1977) et Prodegy (Urban et al 1984), les approches proposées par Rao et Sabavala (1981), Grover et Dillon (1985), Givon (1985), Kumar et Sashi (1989) correspondent à ce type de processus. Le point commun à toutes ces approches est d'utiliser des mesures de transferts entre marques (ou des changements forcés, pour le modèle Prodegy) pour mesurer la concurrence entre les marques. Nous présentons ces approches au chapitre II. Nous pensons que ces approches ont essentiellement pour objet de caractériser la structure du marché (en termes d'attributs physiques), plus que de mesurer la concurrence entre les marques au sein d'un sous-marché.

La deuxième phase du processus de choix serait à mettre en parallèle avec une représentation spatiale de la concurrence. Nous trouvons alors toutes les approches classiques proposant des représentations de la concurrence à partir de techniques de scaling appliquées sur des données perceptuelles (voir Dillon et al 1985) mais aussi tous les modèles de préférence et de choix qui ont été proposés dans la littérature (voir chapitre II, section 5 pour une présentation de ces modèles). L'objet de ces approches est essentiellement de comprendre des relations de concurrence entre des marques au sein d'un champ concurrentiel défini au sens étroit. Les dimensions permettant de décrire la concurrence sont, selon les cas, de type "*perceptuel*" ou "*physique*". Holbrook (1981) a développé un modèle cohérent avec cette optique. Le choix y est le résultat d'un processus de traitement de l'information en deux étapes : une étape dans laquelle les attributs physiques sont convertis en attributs perçus, suivie d'une étape ou les attributs perçus sont combinés de manière à former un affect.

Johnson (1984) a montré que plus les produits sont différents, plus le consommateur aurait tendance à utiliser une stratégie inter-attributs du type EBA, ce qui correspond à un processus de choix de type hiérarchique. Lorsque les produits sont assez similaires (donc après la sélection hiérarchique) le consommateur utiliserait plutôt une stratégie de type intra-attribut, c'est à dire compensatoire et continue. Cependant, aucune validation empirique n'a, à notre connaissance, établi le lien entre les deux niveaux d'analyse que sont processus individuel de choix (réalité consommateur) et la structure du marché (réalité de marché).

Nous avons, dans cette section, présenté trois approches de l'étude du comportement de choix :

1- Une approche qualifiée de "*stochastique*" qui considère le comportement individuel comme essentiellement stochastique. L'accent, dans cette approche, est surtout porté sur la qualité de l'agrégation des individus et sur la prise en compte de leur hétérogénéité. Ceci présente l'avantage d'obtenir des estimateurs qui, au plan agrégé, sont très fiables. L'objectif essentiel est en effet de faire des prévisions. Il n'est pas d'expliquer les processus de choix ou bien l'influence des variables d'actions sur les comportements. Dans cette approche, seuls les comportements passés sont observés.

2- Une approche "*probabiliste*" dans laquelle l'introduction de variables aléatoires correspond plus au souci de prendre en compte l'erreur de mesure et de prévision qu'a la reconnaissance que le comportement est stochastique. Cette approche repose sur la

théorie de l'utilité, dont le modèle de Luce est un exemple bien connu, ainsi que des modèles économiques comme le modèle de Lancaster (Lancaster 1966). Ces modèles sont estimés à partir de l'observation des comportements passés. Ils essaient d'expliquer ce comportement à partir des caractéristiques (la plus part du temps objectives) des produits et éventuellement des individus. Ces modèles allient donc deux types d'objectifs : la prédiction et l'explication.

3 Une approche "*déterministe*" qui est centrée sur le comportement individuel. Elle procède en terme de processus dont l'objet principal est de comprendre et expliquer le traitement de l'information conduisant à une prise de décision au plan individuel. La caractéristique de cette approche, par opposition aux deux précédentes, est d'utiliser des mesures de concepts "non observables" que sont perceptions, préférences et intentions.

Nous pensons pourtant, que ces trois approches du comportement de choix seront de moins en moins opposables dans les années à venir et ne doivent, par conséquent, plus être considérées comme des domaines de recherche cloisonnés. Les approches "Gestion", "Economie" et "Psychologie" convergent et vont de plus en plus converger vers des modèles uniques. Ceci est à mettre sur le compte de l'interpénétration croissante des disciplines (notamment grâce aux revues et colloques scientifiques) mais aussi aux progrès de l'informatique qui rendent opérationnelles des méthodes statistiques qui étaient inutilisables auparavant car trop coûteuses en temps de traitement. Ces progrès permettent d'intégrer de plus en plus de paramètres dans les modèles et donc de prendre en compte de plus en plus de variables. Les progrès concernant la qualité des données disponibles expliquent aussi cette tendance à l'uniformisation des modèles. Nous citerons en particulier l'enrichissement considérable des données de panels qui se transforment en "Base de Données à Source Unique" (voir deuxième partie chapitre I). Cet enrichissement permet de prendre en compte les comportements mais aussi les variables explicatives de ce comportement, descripteurs du produit, du ménage, actions marketing.

Enfin, les gestionnaires sont de plus en plus conscients que prévoir est important mais ne suffit pas, surtout lorsque l'entreprise se trouve dans un environnement en perpétuel changement. Il est alors nécessaire de comprendre, expliquer, diagnostiquer.

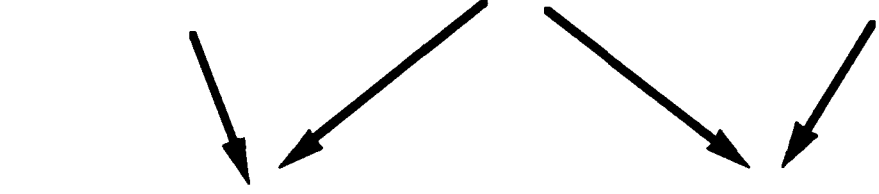
Dans le sens de l'unification des différentes approches, nous pouvons citer les travaux de McFadden (1980 et 1986) dont un des objets est l'introduction de données perceptuelles dans les modèles économétriques de choix. De même, Wierenga (1983 et 1984) a ajouté la dimension "perceptuelle" au modèle de Lancaster. Ces tentatives reflètent une tendance à la fusion des modèles de choix probabilistes et déterministes. Nous pouvons encore citer Hauser et Simmie (1981) qui ont traité le problème de l'hétérogénéité des perceptions dans les modèles de choix du type "modèle de Lancaster". Ils ont notamment montré que les notions de frontière efficiente se transforment en zone efficiente lorsque cette hétérogénéité est prise en compte. Ceci explique pourquoi une marque qui est en deçà de la frontière efficiente a tout de même une part de marché non-nulle.

D'autre part, Bass et Pilon (1980), Jeuland (1979) ont introduit des variables d'action marketing dans leurs modèles stochastiques de parts de marché. Carpenter et Lehmann (1985) ont proposé un modèle explicatif des transferts entre marques à partir des variables d'action du mix et des attributs du produit et de l'interaction des deux. Blattberg (1981) présente des exemples d'intégration de variables du mix (notamment publicité et

promotion) dans les modèles stochastiques de parts de marché et insiste sur la nécessité pour ces modèles de devenir plus explicatifs. Les modèles stochastiques tendent ainsi à se rapprocher des modèles probabilistes en acquérant une dimension explicative.

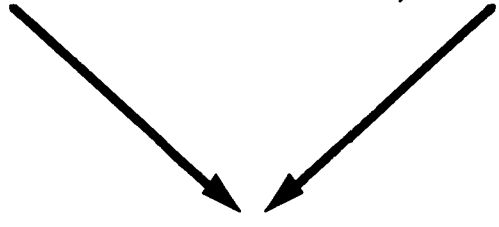
TABLEAU 4 EVOLUTION DES MODELES DE CHOIX

MODELE STOCHASTIQUE DE CHOIX	MODELE PROBABILISTE DE CHOIX	MODELE DETERMINISTE DE CHOIX
MODELE DE PART DE MARCHE	MODELE DE CHOIX	MODELE DE PREFERENCE
DONNEES DE COMPORTEMENT PASSE	-DONNEES DE COMPORTEMENT -CARACTERISTIQUES PHYSIQUES PRODUIT ET CONSOMMATEUR	DONNEES PERCEPT-UELLES ET NON OBSERVABLES
DISTRIBUTION STATISTIQUE AU NIVEAU INDIVIDUEL PUIS AGREGE	-MODELISATION AU PLAN INDIVIDUEL -MODELISATION PUIS AGREGATION	-NIVEAU INDIVIDUEL AGREGATION EVENTUELLE AU NIVEAU DU SEGMENT
- OBJECTIF = PREVISION	-OBJECTIF = EXPLIQUER ET PREVOIR	- OBJECTIF = COMPRENDRE ET EXPLIQUER



MODELE STOCHASTIQUE DE CHOIX PRENANT EN COMPTE LES VARIABLES D'ACTION (ex: BASS et PILON 1980 JEULAND 1979)

MODELE DE PREFERENCE PROBABILISTE (ex: COOPER et NAKANISHI 1983 KAMAKURA et SRIVASTAVA 1986)



"MODELE DE CHOIX DU CONSOMMATEUR"

SECTION 4 SIMILARITE SUBSTITUABILITE CONCURRENCE

Lors de l'étude des liens respectifs existant entre deux marques (ou deux produits), trois concepts sont souvent simultanément ou alternativement utilisés. Il s'agit des concepts de similarité, substituabilité et concurrence.
En particulier, ces trois concepts ont été largement utilisés dans la recherche marketing en matière d'études de positionnement des produits et d'analyse de la structure des marchés. Ce fait est générateur de confusion.

Dans cette section, nous essaierons de préciser la nature de ces concepts à partir des définitions qui en ont été données. Nous préciserons aussi le niveau d'analyse auquel ils se placent. Pour finir, nous les comparerons successivement deux à deux.

SS1 LES CONCEPTS DE SIMILARITE SUBSTITUABILITE CONCURRENCE DANS LA LITTERATURE

SS1-1 SIMILARITE

La similarité perçue est un concept qui a été très largement utilisé en Marketing, notamment dans les études de positionnement (voir Dubois 1979). Son utilisation est allée de pair avec celle de l'analyse des similarités qui a connu un très grand succès dans les années soixante dix. Etonnamment, l'utilisation intensive de ce concept, en Marketing, n'est pas allée de pair avec des recherches visant à en préciser la nature théorique et à permettre son opérationnalisation correcte. Les manuels traitant de l'analyse des similarités (voir le numéro spécial de la revue Applied Psychological Measurement de 1983, Davies et Coxon 1981, Schiffman et al 1981, Shepard et al 1972) restent très réservés sur ce sujet. Dubois et Evrard (1976), dans une réflexion sur l'utilisation de l'analyse des similarités, ont insisté sur ce point. Les auteurs citent notamment Green qui d'une part définit le concept en terme de "relation de dominance ou de consonance" et qui d'autre part l'opérationnalise en terme de "substitution d'un objet à un autre objet pris en référence".

De nombreux psychologues ont remis en cause l'hypothèse selon laquelle, les similarités seraient symétriques. Le jugement de similarité est, selon Tversky (1977), directionnel :"*y is like x*". Selon Krumhansl (1978), la similarité d'un objet avec lui-même dépendrait du contexte dans lequel est envisagé cet objet. C'est ainsi que la similarité de l'objet x avec lui même pourrait être différente de celle de l'objet y avec lui même. Toujours selon cet auteur, l'assymétrie de la relation pourrait être due au fait que les jugements de similarité sont fonction de "la facilité avec laquelle un objet peut être transformé pour entrer en correspondance avec un autre". C'est une définition de la similarité qui, à notre avis, est singulièrement proche de celle de la substituabilité. Tversky (1977) a proposé un modèle théorique de la similarité. Il y définit la similarité comme une combinaison linéaire des caractéristiques "*communes et distinctes des objets*". La similarité Sxy entre deux objets x et y est définie de manière directionnelle c'est à dire en terme de similarité de x vers y comme :

$$S_{xy} = \alpha.f(x*y) - ß.f(x-y) - \gamma.f(y-x)$$

α, ß et γ sont des coefficients positifs
f est la fonction de similarité
x*y représente le nombre de caractéristiques communes aux deux objets
x-y représente le nombre de caractéristiques possédées uniquement par x
y-x représente le nombre de caractéristiques possédées uniquement par y

Dans une application empirique, l'auteur opérationnalise la fonction f par la valeur N qui est le nombre total de caractéristiques possédées par les deux objets x et y. D'autre part, il retient une définition symétrique du concept en posant ß=γ (ce qui constitue selon l'auteur un cas très particulier).
La similarité s'exprime alors comme :

$$S_{xy} = c + \alpha.N(x*y) - ß.[N(x-y) + N(y-x)]$$
c est une constante

La similarité ainsi représentée est une relation de type symétrique entre les deux objets. C'est une combinaison linéaire de leurs ressemblances et de leurs différences.

Krumhansl (1978) a présenté un modèle dans lequel la similarité est modélisée de la façon suivante.
S_{xy} , la similarité perçue entre x et y, est une fonction monotone décroissante de d''_{xy}, la distance "modifiée" entre x et y.

$$d''_{xy} = d_{xy} + a.\pi(x) + b.\pi(y)$$

d_{xy} est la distance euclidienne entre x et y
$\pi(x)$ et $\pi(y)$ sont les densités des points x et y relativement au contexte dans lequel ils sont considérés.
$\pi(x)$ et $\pi(y)$ sont deux fonctions exprimant la densité des zones correspondant à x et à y .

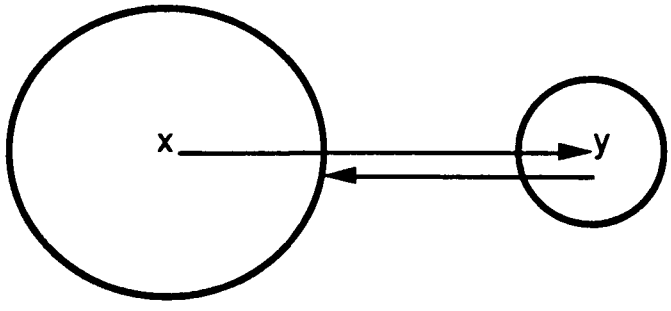
Une opérationnalisation proposée (voir DeSarbo et Manrai 1988 pour plus de détails) de la fonction π est la suivante :

$$\pi(x) = \sum_{i \in R} n_{ix} \quad \text{pour i différent de x}$$

R est l'ensemble des objets environnant x
$n_{ix} = 1$ si $d_{ix} \leq r$ (d_{ix} est la distance entre i et x)
$n_{ix} = 0$ si $d_{ix} > r$
r est un rayon constant donné
$\pi(x)$ est donc une mesure de la densité du voisinage autour de x
a et b sont deux paramètres non négatifs.

Un tel modèle, permet de concevoir la non-symétrie des relations de similarité perçue, comme sur l'exemple suivant.

Exemple de relation non-symétrique:



Sur cet exemple, x est moins similaire à y que y n'est similaire à x. Nous pouvons illustrer cet exemple par les déplacements entre une grande ville et un petit village. La distance du centre de la ville à l'entrée du village est plus grande que la distance du centre du village à l'entrée de la grande ville.

SS1-2 SUBSTITUABILITE

Le concept de "Substituabilité" est issu de l'analyse économique. Traditionnellement, les économistes définissent deux biens comme substituables quand leurs élasticités croisées au prix sont positives. Il y a une relation bi-univoque entre les deux concepts. Le choix entre deux biens étant conçu en terme d'allocation de ressources rares pour la maximisation de l'utilité du consommateur, la substituabilité est vue comme la possibilité d'allouer un budget à différents types de produits sans pour autant définir le contexte dans lequel l'utilité doit être maximisée.

En marketing, la substituabilité fait référence non pas à un problème d'allocation de ressources mais à la satisfaction des bénéfices attendus pour une situation d'usage (ou d'achat) particulière du produit: c'est la *"substitution dans l'usage"*. Ainsi, en Marketing, substituabilité et élasticité croisée au prix ne sont pas aussi intimement liées qu'en économie. D'autres variables que le prix ou les quantités sont accessibles pour influencer le marché, ce qui signifie que d'autres types d'élasticité sont envisageables (par exemple, l'élasticité croisée à la publicité).

Shocker et al (1987) proposent une définition de la substituabilité qui est relative à la notion de situation d'usage: *"deux produits sont substituables s'ils offrent le même éventail de bénéfices attendus pour une situation d'usage particulière"*.

Fraser et Bradford (1983) donnent la définition suivante de la substituabilité : *"Deux produits sont dits parfaitement substituables s'ils sont utilisés de façon interchangeable pour une situation d'usage identique"*. Ce qui fait dire à ces auteurs qu'une élasticité croisée au prix fortement positive est une condition suffisante mais non nécessaire de substituabilité. Deux produits peuvent être substitués pour d'autres raisons que le lien de leurs fonctions de demande respectivement à leurs prix relatifs.

Ce qu'il est important de noter dans cette définition est l'effet de la situation d'usage . En effet, comme nous l'avons déjà souligné dans la section 1, deux produits peuvent être

substituts pour une situation d'usage particulière sans l'être pour une autre. Il est donc nécessaire de préciser la situation d'usage du produit lors de la mesure la substituabilité.

Ainsi, la substituabilité s'inscrit dans un cadre de référence spécifique : "*deux biens sont substituables parce qu'ils amènent une combinaison de bénéfices qui les rend acceptables pour une situation d'usage particulière*". C'est donc la confrontation à une situation qui permet d'évaluer le degré de substituabilité entre les produits.

SS1-3 CONCURRENCE

Le concept de concurrence a fait l'objet, lui aussi, de trop peu de recherches sur son contenu théorique. Nous verrons dans le chapitre suivant qu'un grand nombre de mesures de la concurrence ont été proposées. Cependant, peu d'auteurs ont essayé, avant de proposer leur mesure, de préciser le contenu et les contours du concept. Il semble qu'il y ait une confusion entre le concept et son opérationnalisation.

Porter, dans son ouvrage "*L'avantage Concurrentiel*" (1986) définit l'avantage concurrentiel et les moyens de l'obtenir, il ne précise cependant pas la nature même de la concurrence. Shocker est un auteur qui a beaucoup réfléchi sur la nature des concepts de substituabilité et de concurrence. Selon lui, la caractéristique essentielle de la concurrence est d'être relative. Elle se définit "*par rapport à ..*". Il définit la concurrence (1986) en terme d'impact Marketing : "*deux produits sont concurrents si leurs performances sur le marché (les ventes) peuvent être substantiellement affectées par les actions marketing (prix, promotion ..) de l'un ou de l'autre*".

C'est une définition qui correspond à l'acception classique du terme de concurrence. La nature de la relation de concurrence est une relation de force qui s'exprime sur le marché en termes de résultats. Nous ferons remarquer que le concept d'élasticité croisée des économistes est tout à fait compatible avec une telle définition de la concurrence, à la différence de la substituabilité au sens marketing du terme.

SS2 NIVEAU D'ANALYSE DES CONCEPTS DE SIMILARITE SUBSTITUABILITE CONCURRENCE

Pour mieux comprendre les concepts de similarité, substituabilité, concurrence, il est utile de préciser leur niveau d'analyse.

Nous avons, dans l'introduction à ce chapitre, distingué deux types de réalité : la "réalité consommateur" et la "réalité du marché". Ces deux réalités constituent deux niveaux d'analyse différents par rapport auxquels il est bon de replacer les trois concepts étudiés. Nous avons signifié que la réalité consommateur se définit au niveau de l'individu, dans des enquêtes sur ses perceptions et ses attitudes. La réalité de marché se place, elle, exclusivement au niveau des comportements qui sont, en particulier, influencés par les variables de situation d'achat.

En se conformant strictement à ce schéma d'analyse, nous constatons que:

- les concepts de similarité et de substituabilité relèvent de la réalité du consommateur et prennent naissance au niveau de l'individu.
- le concept de concurrence relève de la réalité du marché et se définit au plan agrégé.

Le croisement des niveaux d'agrégation et du type de réalité permet de replacer les concepts dans le tableau suivant.

TABLEAU 5A NIVEAUX D'ANALYSE ET RELATIONS ENTRE MARQUES

	REALITE DU CONSOMMATEUR	REALITE DU MARCHE
NIVEAU INDIVIDUEL	SIMILARITE SUBSTITUABILITE	
NIVEAU AGREGE		CONCURRENCE

Les deux cases non remplies (2 et 3) peuvent en fait l'être, en prenant certaines précautions.

La case 3 correspond à une réalité de consommateur. En effet, la réalité consommateur prend naissance au niveau de l'individu, mais rien n'empêche ensuite d'en effectuer l'agrégation. C'est ce qui est le plus souvent fait dans les études marketing. Cependant, son niveau agrégé ne lui procure pas pour autant "valeur de marché".

La case 2 correspond à un comportement individuel observé au niveau du marché. Nous y trouvons les concepts de "*préférence révélée*" tels que des auteurs comme Bass (1974), Jeuland (1978) les définissent, c'est à dire à partir des parts d'achats des marques dans les achats totaux. Nous y trouvons aussi le concept de "*substitution révélée*" proposé par Fraser et Bradford (1983). Nous y trouvons enfin la notion de "*concurrence révélée*" , définie par la répartition relative des achats dans le panier de la ménagère, tel que le propose Merunka (1988). Dans ce dernier cas, le terme de concurrence devrait être, selon nous, remplacé par celui de substitution puisque (toujours selon nous) la concurrence n'a de signification qu'au seul plan agrégé.

Nous pouvons alors remplir les cases vides du tableau précédent.

TABLEAU 5B NIVEAUX D'ANALYSE ET RELATIONS ENTRE MARQUES

	REALITE DU CONSOMMATEUR	REALITE DU MARCHE
NIVEAU INDIVIDUEL	SIMILARITE SUBSTITUABILITE	SUBSTITUABILITE REVELEE
NIVEAU AGREGE	SIMILARITE OU SUBSTITUABILITE AGREGE	CONCURRENCE

A partir de ces constatations, nous pouvons tirer les conclusions suivantes.

Si c'est la "réalité consommateur" qui est étudiée, il est nécessaire de procéder à des mesures au niveau individuel, à partir de données de jugement, quitte ensuite à agréger ces mesures.

A l'inverse, si c'est la concurrence sur le marché qui est étudiée, il nous parait plus judicieux de mener directement des analyses au plan agrégé, sans passer par des estimations au plan individuel. Par contre, il n'est pas exclu d'étudier le concept de "substituabilité révélée", auquel cas il est nécessaire de procéder à des estimations individuelles à partir de comportements d'achats.

SS3 COMPARAISONS ENTRE LES CONCEPTS

SS3-1 SIMILARITE ET CONCURRENCE

La comparaison qui a suscité le plus grand nombre d'interrogations dans la littérature est celle qui est relative à la comparaison entre similarité perçue et concurrence. En effet, de nombreux chercheurs ont fait remarquer que les cartes obtenues à partir de données de similarité perçue ne constituent pas une représentation du phénomène concurrentiel. En effet, de telles cartes accentuent plutôt les dimensions qui permettent au répondant de différencier les différentes alternatives qui lui sont proposé. Il n'y a aucune raison pour que ces dimensions de discrimination perceptuelle correspondent aux dimensions qui expliquent la concurrence.

Nous citerons Lehmann (1972), qui a montré que les cartes obtenues à partir de similarités perçues ou de transferts entre marques ne convergent pas. De même Dubois (1977), Myers et Tauber (1977 p 8) ont aussi critiqué l'assimilation qui est faite entre similarité et concurrence. Fraser et Bradford (1983) font remarquer qu'il peut n'y avoir aucun transfert entre deux produits similaires qui sont alors considérés comme étant non-concurrents.

Nous retrouvons le même type de critique dans Chandon et Strazziéri (1986) selon lesquels la similarité perçue est trop éloignée (en terme de processus) de l'acte d'achat pour constituer une mesure de concurrence. Ils ont proposé une grille d'analyse intéressante qui permet de faire la différence entre les deux concepts (voir tableau suivant).

TABLEAU 6 LIEN ENTRE SIMILARITE ET CONCURRENCE

		INTENSITE DE CONCURRENCE	
		FORTE	FAIBLE
SIMILARITE PERCEPTUELLE	FORTE	CONCURRENCE FRONTALE OU ASSIMILATION	?
	FAIBLE	CONCURRENCE DE REMPLACEMENT OU DE SUBSTITUTION	NON-CONCURRENCE ET NICHES ELOIGNEES

"Adapté de Chandon et Strazzieri 1986 p 35"

Ce tableau nous montre que deux produits peuvent être concurrents sans pour autant être similaires.

Un exemple simple peut illustrer ce point. Soient deux marques A et B dont les positionnements sont bien différents. La marque A est positionnée "haut de gamme" et apporte des bénéfices fortement appréciés par le consommateur. La marque B est positionnée "bas de gamme" et pratique une politique de prix qui lui permet de prendre des parts de marché sur la marque A. Une carte de positionnement perceptuel opposerait ces deux marques, alors qu'une carte reposant sur les relations concurrentielles les placera à proximité l'une de l'autre.

La différence entre similarité perceptuelle et concurrence est d'ailleurs à la base des stratégies de Positionnement. Pour se positionner, une marque va essayer d'être perçue différemment de ses concurrentes. Si ce positionnement est réussi (apprécié par le consommateur) la marque fera de la concurrence aux autres marques dont elle a su se différencier d'un point de vue perceptuel.

A partir de ces constatations, il apparait tout à fait insuffisant de baser des études de positionnement préalables à un lancement de nouveau produit ou destinées à contrôler le positionnement, à partir de cartes construites selon des données de comportement, notamment quand l'imaginaire associé au produit constitue un facteur déterminant du succès commercial.

Si une marque A concurrence plus une marque B, cela signifie que les actions de la marque A ont plus d'impact sur les résultats de la marque B que les actions de la marque B n'ont d'influence sur les résultats de la marque A. Barnett (1976), dans un article relatif au modèle d'attraction de Bell, Barney et Little (théorème des parts de marché) a remis en cause cette hypothèse et généralisé les hypothèses du modèle ou cas ou cette symétrie n'est plus vérifiée.

Ainsi, selon nous, toute méthode représentant des marques à partir de leurs indices de concurrence devrait être à même de prendre en compte cette non-symétrie.

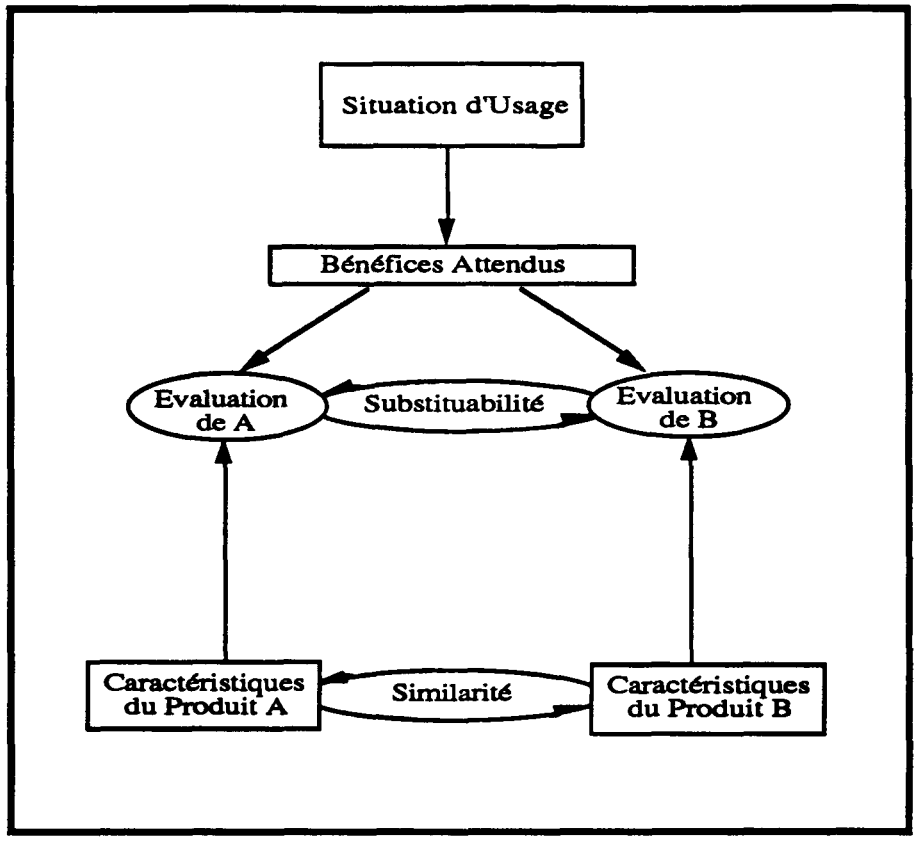
SS3-2 SIMILARITE ET SUBSTITUABILITE

Dans un récent article, Ratneshwar et Shocker (1988) ont étudié le lien entre la substituabilité perçue dans l'usage et la similarité perçue. Les auteurs montrent notamment que le nombre commun de situations d'usage dans lesquelles les produits sont adéquats (la substituabilité dans l'usage) prévoit mal la similarité. Par exemple les fruits et les yaourts sont substituables pour certaines situations d'usage mais, par contre, ne sont pas perçus comme similaires. Ce résultat apparait troublant aux auteurs, à tel point qu'ils l'attribuent à des causes méthodologiques. Au vu de ce que nous avons dit précédemment, ce résultat n'a en fait rien de troublant.

La similarité représente le lien entre deux objets dans des conditions très générales, en quelque sorte de manière absolue. La substituabilité perçue est d'une tout autre nature. Elle se définit par rapport à un contexte particulier, par rapport à la satisfaction du besoin tel qu'il se manifeste par rapport à une situation d'usage particulière. La capacité d'un produit à satisfaire une situation d'usage particulière peut être obtenue par différentes combinaisons de caractéristiques du produit. C'est ainsi que deux produits différents peuvent satisfaire à l'identique une même situation d'usage. En fait l'adéquation à une situation d'usage nécessite que le produit satisfasse les bénéfices "liés" à cette situation d'usage.

La similarité, elle, fait référence aux "caractéristiques" du produit. Deux produits sont perçus similaires si ils présentent un nombre important de caractéristiques communes. De cette différence entre bénéfices liés à une situation et caractéristiques du produit, naît la différence entre similarité et substituabilité. C'est ainsi que deux produits différents peuvent satisfaire à l'identique, une même situation. La figure suivante permettra d'éclairer notre raisonnement.

FIGURE 11 DIFFERENCE ENTRE SUBSTITUABILITE ET SIMILARITE



(Dans ce graphique, nous faisons l'hypothèse que les consommateurs sont homogènes et attendent les même bénéfices pour une même situation d'usage.)
Nous pouvons constater que la similarité est une relation directe entre les deux produits A et B.
A l'inverse, la substituabilité naît de l'interaction entre l'évaluation des deux produits pour une situation d'usage particulière. Celle-ci joue un rôle de médiation entre les deux produits. Si la substituabilité se définit par rapport à la "capacité à satisfaire une situation d'usage particulière", il faut alors la concevoir en tant que concept non-symétrique. En effet, si le produit A est évalué comme plus adéquat à une situation que le produit B, il se substitue plus facilement à B que B ne se substitue à A, pour cette même occasion d'usage.

Pour conclure cette partie, nous ferons les deux remarques suivantes :

- la similarité perçue n'est pas une condition nécessaire à la substituabilité. Nous l'avons dit, différentes combinaisons d'attributs peuvent satisfaire la même situation d'usage.
- l'existence d'une similarité perçue n'est pas une condition suffisante de substituabilité. Il y a dans la substituabilité une dimension affective qui fait que deux produits, avant d'être substituables, doivent être acceptables. Il est alors possible de faire le lien entre substituabilité et préférence ou appartenance à l'ensemble évoqué.

Pour une situation d'usage donnée, il est raisonnable de penser qu'un consommateur considère deux marques substituables :

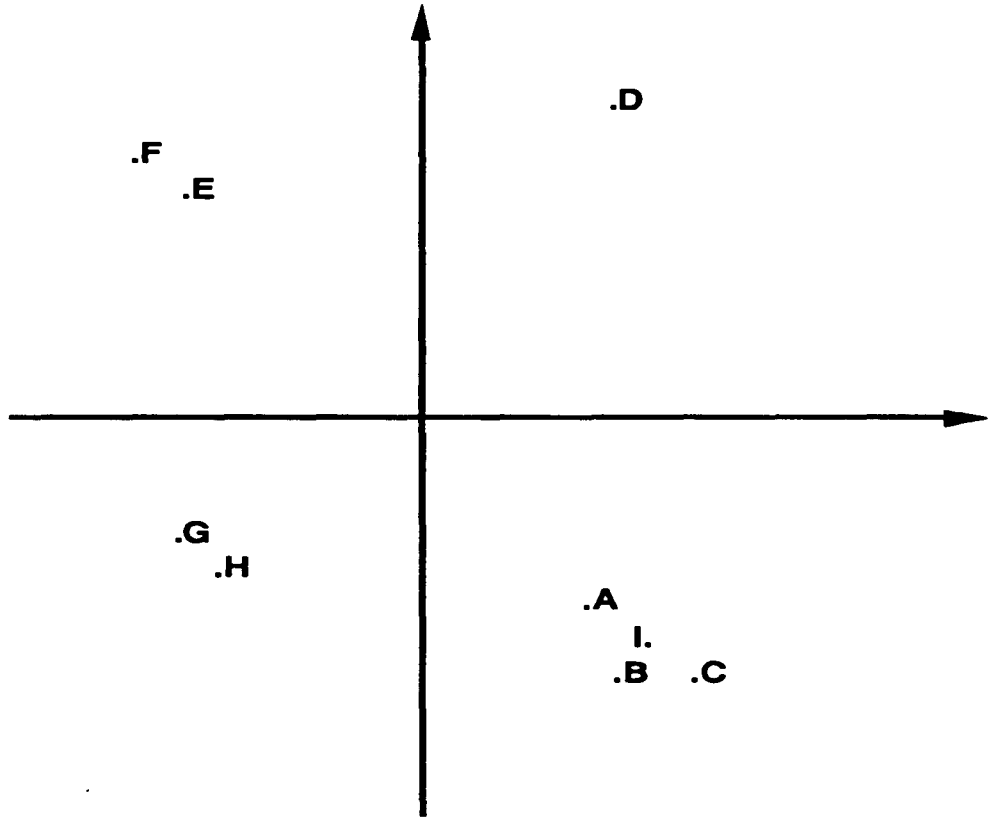
- si elles font l'objet d'une préférence suffisante : elles appartiennent à son ensemble de considération.
- si elles apportent toutes les deux une combinaison de bénéfices acceptable pour la situation d'usage particulière.

Comme le proposent Chandon et Strazzieri (1986), l'appartenance simultanée de deux marques à l'ensemble évoqué d'un répondant, ensemble évoqué qui a été prononcé pour une situation particulière, nous parait constituer une bonne mesure de substituabilité.

Deux marques faisant l'objet d'une forte préférence sont elles substituables ?

- si le contexte d'utilisation est précis, nous penchons pour l'affirmative: les marques sont préférées parce qu'elles satisfont à un certain nombre de bénéfices spécifiques à la situation.
- si le contexte d'utilisation est général, il est alors fort possible que les marques soient désirées pour des motifs différents, auquel cas elle ne sont plus forcément perçues comme étant substituables.

FIGURE 12 REPRESENTATION HYPOTHETIQUE D'UN MARCHE A L'AIDE D'UN MODELE INDIVIDUEL DE PREFERENCES



La carte hypothétique de marché précédente, issue d'une analyse des préférences externe (dans laquelle le point idéal, I, est fixé dans un espace issu d'une analyse des similarités perçues) illustre le fait que deux marques similaires ne sont pas forcément perçues comme substituables, pour une situation d'usage donnée.

Les groupes de marques (F,E), (G,H), (A,B,C) sont perçues comme similaires. Cependant, (F,E) ou (G,H) ne sont pas substituables parce qu'elles ne seront pas considérées par le répondant lors de son processus de choix. Par contre, les marques (A,B,C) le sont (selon notre définition). Elles sont en effet proches du point idéal I du consommateur (ou groupe de consommateurs).

SS3-3 SUBSTITUABILITE ET CONCURRENCE

La différence essentielle entre substituabilité et concurrence, est que la première se définit au niveau individuel, dans un contexte "*cognitif et affectif*" de traitement de l'information et la seconde se définit au plan agrégé, après avoir pris en compte les variables de la situation d'achat.

Les deux concepts sont pourtant utilisés alternativement dans la littérature. Un bon exemple est donné par Shocker (1986) qui, pour définir la substituabilité, utilise le terme de concurrence : "*La substituabilité établit que deux produits sont concurrents parce qu'un consommateur les considère comme interchangeables dans une situation d'usage particulière*". A notre avis, il est important de ne pas confondre ces deux concepts.

Comme nous l'avons noté à la section 2, le comportement de recherche de variété temporelle (ou de renforcement), va conduire le consommateur à évaluer différemment un produit et ses substituts d'une période à l'autre, par le simple fait d'avoir consommé ce produit à la période précédente.

Soit A et B deux produits substituables
A est consommé à la période t

Si un individu est "*chercheur de variété*" alors les bénéfices associés à A vont diminuer d'importance à la période t+1 et de la même manière, les bénéfices associés à B puisque A et B sont substituables. Quelle va en être la conséquence ? A la période suivante, en (t+1), il va acheter une marque apportant un ensemble de bénéfices différents, ce changement lui permettant de satisfaire son besoin de recherche de variété. Que signifie ce changement ? Le consommateur étant saturé par les bénéfices apportés par le produit A, l'attrait de ce produit baisse lors de son évaluation. Ce faisant, A va être "exclu" de l'ensemble évoqué du consommateur. Un autre produit C sera choisi par le consommateur. Les produits A et C peuvent-ils être considérés comme substituables ? Nous répondons par la négative puisque A et C ne sont pas apparus ensemble, au même instant, dans l'ensemble évoqué du consommateur.

Pour notre part, nous nous bornerons à constater que la substituabilité est le lien qui existe entre deux marques à un moment donné du temps. Ni en t, ni en t+1, les produits A et C ne sont considérés comme substituables.

Par contre, la concurrence est une mesure de la relation inter-temporelle. D'ailleurs, son acception classique le montre puisqu'il s'agit de la capacité d'une marque à influencer les résultats de ses concurrentes, il y a donc là une idée de durée. Les deux marques A et C de l'exemple précédent sont effectivement concurrentes. En t+1, la marque C a pris de la part de marché à la marque A. Nous pouvons ainsi considérer que le transfert d'une marque vers une autre est l'expression de la relation de concurrence existant entre ces deux marques, à partir du moment où il a été établi avec certitude que les deux marques appartiennent au même marché. Le transfert entre deux marques est une mesure de la concurrence entre deux marques, il n'est par contre pas forcément une mesure de la substituabilité (au niveau du consommateur) entre ces deux marques, notamment dans le cas où ce transfert est la manifestation de recherche de variété de la part du consommateur.

Ces réflexions montrent qu'il est indispensable de se poser la question du niveau d'analyse à adopter lors de l'étude la structure d'un marché. Si l'analyse est placée du point de vue de la "réalité consommateur", le marché sera structuré en sous-ensembles de produits considérés comme substituables par les consommateurs (ou groupes de consommateurs). Les produits A et C, dans l'exemple précédent, appartiendront alors à deux sous-marchés différents. Si l'analyse est placée du point de vue de la "réalité de marché", le marché sera structuré en sous-ensembles de produits effectivement concurrents sur le marché. Les produits A et C seront alors considérés comme faisant partie du même sous-marché.

Le problème majeur associé à cette double approche est que les structures caractérisées sur le "marché" risquent d'être bien différentes de ce que conçoit le consommateur en matière de substituabilité. Ceci s'explique pour plusieurs raisons. Sur le "marché" (à partir de données de panel), c'est le comportement des ménages qui est observé et non pas celui des individus. Les situations d'usage du produit ne sont pas prises en compte, l'effet des variables "magasin" est lui, par-contre, pris en compte, le niveau d'analyse est inter-temporel ...

Il n'y a pas de bonne ou de mauvaise approche, mais il est important de savoir quel concept on étudie et dans quelles conditions il vaut mieux utiliser un concept plutôt qu'un autre. L'approche centrée sur le consommateur (substituabilité) nous semble mieux adaptée dans un contexte "Etude Marketing" dont l'objectif est la compréhension du consommateur afin de lui fournir et de lui communiquer une "offre" adaptée à ses attentes. L'exemple le plus probant est certainement celui de la création d'un nouveau produit. L'approche centrée sur le marché (concurrence) correspond plus à une optique de contrôle sur le terrain de l'efficacité des actions menées par "nous et la concurrence", à l'étude du champ de bataille, à l'évaluation des forces relatives des concurrents.

SECTION 5 LA DEFINITION DES MARCHES

Comme nous le verrons dans le chapitre suivant consacré à l'exposé des modèles d'analyse de structure des marchés, la définition du marché est la première étape de toute analyse. Or, beaucoup d'auteurs passent ce problème sous silence et exposent la mise en oeuvre de leurs modèles en considérant cette étape comme résolue.

Définir un marché pour en étudier la structure, c'est définir l'ensemble des produits (ou des marques) qui lui appartiennent, c'est-à-dire les produits qui seront considérés pour mesurer des indices de concurrence, découper le marché en sous-marchés, tester des structures de concurrence particulières. Définir un marché équivaut finalement à en tracer les frontières et à fixer un niveau de généralité au type de relations entre marques qui sera étudié.

SS1 DEFINITIONS DU MARCHE: LES DIFFERENTS POINTS DE VUE

SS1-1 APPROCHE ECONOMIQUE

Dans les approches économiques classiques, le marché est la rencontre d'une offre avec une demande. Les produits du marché sont considérés comme homogènes, il existe un prix de marché résultant de la confrontation entre offre et demande. Les seules variables prises en compte sont les quantités et les prix des produits. Dans ce type d'approche, les relations entre les marques au sein d'un marché particulier n'ont pas besoin d'être étudiées puisque tous les produits au sein d'un marché sont parfaitement substituables. Seules sont étudiées les relations entre des produits appartenant à différentes catégories, puisque leur choix peut être le résultat de différentes allocations d'un même budget. Ces relations sont étudiées à partir des élasticités croisées (aux prix ou aux quantités).

Selon cette perspective (voir Micaleff 1982 pour une présentation du marché selon le point de vue des économistes), la définition du marché n'est pas présentée comme la recherche d'un ensemble de produits ou de marques constituant un ensemble commun mais plutôt comme les relations qu'entretiennent une Offre (catégorie de produits définie à priori) et une Demande.

La logique de cette approche est bien représentée par le modèle de Hicks et Samuelson dans lequel le consommateur est rationnel et effectue ses choix de manière à maximiser son utilité tout en tenant compte de ses contraintes économiques. Ce modèle, fixe et statique formalise une optimisation. Il a été considérablement assoupli par la suite, notamment avec le modèle de Lancaster (1966) qui établit une théorie de la réponse du consommateur aux variations des attributs physiques (dit aussi "écologiques"), non perçus du produit.

SS1-2 APPROCHE MANAGEMENT

Dans cette approche, le marché est souvent défini par rapport à l'entreprise et à ses objectifs. Par exemple, le manager désire t-il définir le marché actuel, ou bien le marché potentiel ?

Dans ce type d'approche centrée sur l'activité de l'entreprise il est possible de parler du "marché de la marque de l'entreprise". Deux optiques traditionnelles de définition du marché coexistent :

SS1-2-1 DEFINITION A PARTIR DU PRODUIT

C'est l'approche classique "Producteur" dans laquelle les produits sont groupés en fonction de la similarité de leurs attributs physiques, processus de production, circuits de distribution, type de technologie utilisée, matières premières consommées.

Cette approche est l'expression d'une logique de producteur. Elle est couramment utilisée en analyse stratégique lorsque le marché est découpé en "Unités de Base Stratégiques" (voir notamment Porter 1986) mais elle l'est aussi en marketing. A la différence des approches "consommateurs", les définitions du marché selon une logique "producteur" sont plus claires, plus parlantes pour le manager, mais surtout plus faciles à rendre opérationnelles et plus stables dans le temps. Les entreprises sont souvent organisées en fonction de telles définitions du marché (notamment les chefs de produit et les chefs de marché). Ces deux approches peuvent conduire à obtenir différentes définitions du marché.

SS1-2-2 DEFINITION ORIENTEE VERS LE CONSOMMATEUR

Un grand nombre de définitions du marché existent dans la littérature Marketing, nous n'en citerons que quelques unes. Nous pouvons observer qu'elles adoptent différents points de vue. Le marché est alternativement défini, selon les auteurs ou les cadres d'analyse :

- à partir des individus qui le composent ,
- à partir des produits qui le composent ,
- à partir d'une relation entre un ensemble de produits et un ensemble d'individus.

Des deux premiers points de vue est né le débat maintenant classique entre segmentation des produits et segmentation des marchés (voir Dickson et Ginter 1988 pour un éclairage nouveau de ce débat). Plutôt qu'adopter une position ferme, il nous parait plus judicieux de considérer que la définition d'un marché peut différer en fonction du "point de vue" et de "l'objectif" de cette définition".

Si il s'agit, par exemple, de mettre en oeuvre une stratégie de Segmentation (au sens Marketing et non Stratégique du terme), il parait naturel de définir le marché comme un ensemble de segments d'individus. Le marché est alors défini par rapport aux individus qui le composent. L'objectif est ensuite de communiquer une offre spécifique de la firme à une cible spécifique : le segment de marché choisi.

Par contre, lors de l'analyse la structure d'un marché, le responsable marketing s'intéresse aux relations qui existent entre les produits (marques) appartenant à ce marché, il est donc naturel de considérer le produit (la marque) comme l'unité d'analyse. Le marché est alors défini par rapport aux produits qui le constituent.

Le troisième point de vue, plus riche mais aussi plus difficile à opérationnaliser, est de considérer que l'ensemble des produits constituant un marché et ses sous-marchés, les

relations entre les produits au sein de ces sous-marchés, ne sont pas uniques mais varient en fonction des groupes de consommateurs considérés (et éventuellement les situations d'usage considérées).

Un marché est alors défini par *"un ensemble de consommateurs, un ensemble de produits et la mesure de réponse de ces consommateurs à ces produits"*. Le marché est le résultat de l'interaction entre offre et demande.

SS1-2-2-1 DEFINITION DU MARCHE PAR RAPPORT AUX UNITES QUI ACHETENT (individu, ménage, ..)

Une définition déjà ancienne et qui a le mérite d'être claire est celle donnée par Sissors (1966), dans un article intitulé "*What is a market* ?" Pour l'auteur, la définition est simple:"*les individus constituent le marché*", "*un marché est constitué par l'ensemble des consommateurs potentiels d'un produit*". Kotler et Dubois (1986) définissent le marché comme *"l'ensemble des individus et organisations qui achètent ou pourraient acheter un produit"*. Dussart (1985) définit le marché comme étant composé de *"toute unité d'achat qui, dans un avenir immédiat ou futur, peut être amenée à acheter un produit ou un service donné"*. Dans ce type de définitions, l'attention est focalisée sur la demande au marché, c'est-à-dire, la quantification du marché réel ou potentiel.

SS1-2-2-2 DEFINITION DU MARCHE PAR RAPPORT AUX PRODUITS (Unités achetées)

Selon cette perspective, le marché est considéré comme un ensemble de produits (marques) dont l'image, l'usage ou le choix présentent des éléments similaires. Nous distinguons différents niveaux dans le processus de choix du consommateur qui permettent d'opérer une définition: les perceptions, les attitudes, les préférences et finalement les choix.

DEFINITIONS BASEES SUR LE JUGEMENT

Le jugement des acheteurs réels ou potentiels d'un produit permet de définir le marché par rapport à la logique "*réalité du consommateur*". Le marché est alors défini non pas tel qu'il est mais tel qu'il pourrait être. Shocker et Srinivasan (1974) ont proposé une définition du marché à partir de la notion de similarité perçue. Les auteurs insistent sur la nécessité de définir le marché sur lequel l'entreprise a l'intention de lancer un nouveau produit. Pour cela, ils proposent d'effectuer une analyse des similarités perçues avec un grand nombre de produits et de voir comment se dessinent les frontières des marchés.

La faiblesse de ce type de définition est que de nombreux auteurs ont montré que la perception n'est pas homogène entre les individus (voir Rao 1972, Ritchie 1974, Glazer 1984, Gensh 1985). D'autre part, nous avons vu à la section 1 que la situation d'usage envisagée peut faire varier substantiellement la perception qu'a un individu d'un marché. Il semble donc que la définition d'un marché à partir des similarités perçues soit insuffisante si le contexte de l'analyse requiert une vision large de ce marché. Par contre, elle conviendrait bien pour définir un marché, au sens étroit, ou un sous marché particulier. En effet, dans ce cas là, peu de situations d'usage sont envisagées par le répondant.

Shocker et al (1988), ont formulé une définition du marché qui repose sur le processus de choix du consommateur et peut être mise en relation avec le concept "*d'ensemble évoqué*". Ils définissent le marché comme l'ensemble des produits considérés par le consommateur dans la deuxième phase de son processus de choix, la première phase étant un processus d'élimination hiérarchique permettant d'aboutir à la sélection de cet ensemble de marques

entre lesquelles il hésite (voir section 3). Le marché ainsi défini correspondrait à l'ensemble évoqué du consommateur, concept sur lequel repose l'approche proposée par Chandon et Strazziéri (1986).

Une définition qui repose sur le concept de "*substituabilité perçue*" est celle qui est donnée par Hruschka (1986) : "*un marché est un ensemble de marques (produits) perçus comme substituts*".

DEFINITIONS BASEES SUR LE COMPORTEMENT

Le marché est ici défini à partir des relations de concurrence telles qu'elles s'expriment réellement.

Fraser et Bradford (1983) définissent le marché à partir du concept de "substituabilité révélée". Cette substituabilité révélée est opérationnalisée par les auteurs à partir des périodes inter-achats des marques (voir Chapitre II section 3).

D'autres auteurs semblent préférer définir un marché à partir de la définition de ses sous-marchés, ceci constituant peut être un aveu d'impuissance à en cerner "*à priori*" les frontières. Les auteurs finissent ainsi par cerner les composantes d'un marché (les sous-marchés) sans avoir pour autant justifié au préalable la nature du tout !

Urban, Johnson et Hauser (1984), dans la présentation du modèle Prodegy, définissent un sous-marché par rapport à un produit particulier de la firme. Un sous-marché est alors défini par la réponse à la question : "*à quel sous marché appartient mon produit?*". Ils donnent la réponse: "*un sous marché est constitué par l'ensemble des produits qui sont le plus affectés par la stratégie Marketing de la firme*" ou encore "*un produit appartient à un sous marché s'il prend ses ventes, lors de son lancement, aux produits appartenant à ce sous marché*". Cette définition est intéressante parce qu'elle repose sur les fondements de la notion de concurrence. Deux marques sont en concurrence quand les actions de l'une influencent les résultats de l'autre.

Dans le modèle de Hendry, un sous-marché (partition) est constitué par l'ensemble des produits vérifiant la loi de proportionnalité aux parts de marchés: "*deux marques appartiennent au même sous-marché si les transferts entre ces deux marques sont proportionnels aux parts de marché de ces marques, au sein du sous-marché*".

SS1-2-2-3 DEFINITION DU MARCHE PAR RAPPORT AUX PRODUITS ET AUX CONSOMMATEURS

Day, Shocker et Srivastava (1979) ont donné une définition à la fois séduisante et opérationnelle d'un marché de produits . Ils le définissent comme "*un ensemble de produits jugés substituts dans des situations d'usages pour lesquelles des combinaisons identiques de bénéfices sont attendues, et les consommateurs pour qui de tels usages sont valables*". Il s'agit là d'une définition intégrant à la fois consommateurs et produits mais aussi situation d'usage dont nous avons souligné l'importance à la section 1 précédente.

Shocker a reformulé cette définition d'une manière plus empirique mais aussi plus opérationnelle : "*un marché de produits est constitué par l'ensemble des alternatives qui sont considérées de manière positive par un pourcentage minimum des consommateurs pour lesquels l'achat ou l'usage du produit présente un intérêt*".

Green et DeSarbo (1979 p 90) définissent le marché comme "*un ensemble d'attributs du produit (caractéristiques produit), d'attributs du consommateur (caractéristiques consommateur) et une mesure de réponse comme la préférence ou le choix de ces consommateurs concernant ces produits*".

Face à l'enchevêtrement des problèmes de segmentation et de définition d'un marché (lorsque par exemple différents groupes d'individus ont différentes perceptions des marchés et sous marchés), certains auteurs se sont tourné vers des approches simultanées de la segmentation et de la définition d'un marché (voir Grover et Srinivasan 1987, Grover et Dillon 1985, voir chapitre suivant, section2). Les segments de marchés sont alors définis à partir du choix des marques.

SS2 L'IMPORTANCE DE L'UNIVERS CONCURRENTIEL

L'importance de la définition du marché peut être mieux saisie à partir des questions suivantes relatives à deux types d'influences .

1 Influence du nombre de marques :

- est-ce que la concurrence entre deux marques et sa mesure sont affectées par l'introduction ou la suppression d'une autre alternative ?
- est-ce que les critères permettant d'expliquer le découpage du marché en sous-marchés ne sont pas modifiés par l'introduction ou la suppression d'alternatives ?
- est-ce que le regroupement des marques en sous-marchés est stable par rapport à ces introductions ou suppressions ?

2 Influence de l'hétérogénéité des marques appartenant au marché:

Cette hétérogénéité peut se concevoir en différents termes :

- au niveau de produits perçus comme différents
- au niveau de produits plus ou moins bien connus

Il faut alors se poser les questions suivantes :

- est ce que la perception de produits existant est modifiée par l'introduction, sur le marché de nouvelles alternatives, différentes des précédentes ?
- est ce que les processus de choix varient avec l'introduction d'une marque (stimulus) hétérogène parmi un ensemble de choix (on s'intéresse alors au processus de choix) ?

Les réponses aux questions précédentes peuvent être envisagées selon trois voies différentes :

1 - Quel est l'effet de la modification de l'univers concurrentiel sur les relations de concurrence. Par exemple, quel est l'effet sur les frontières du marché qui est dû à l'introduction d'une nouvelle marque ? Le raisonne se place alors au niveau de la "réalité du marché".

2 - Quel est l'effet de la modification du nombre et de l'hétérogénéité des alternatives présentées, sur le comportement de réponse de l'individu (lors de la collecte de données de jugement) ? Par comportement de réponse, on entend perceptions, préférences, processus de choix. Le raisonnement se place alors au

niveau de la réalité du consommateur, il s'agit d'un problème de nature théorique.

3 - Quel est l'effet de la modification du nombre ou de l'hétérogénéité des alternatives étudiées sur l'analyse statistique elle-même, l'univers concurrentiel réel (le marché) restant identique ? Il s'agit alors d'un problème d'analyse des données qui est de nature méthodologique.

Les trois effets précédents peuvent s'ajouter et même inter-agir. La définition du marché est un problème important à résoudre.

Le point 1 correspond à ce qui se passe réellement sur un marché lorsqu'un produit y est introduit ou retiré. Dans ce cas, l'univers concurrentiel s'impose de lui-même. Par contre, en Marketing, lors de l'analyse d'un marché particulier, nous serons:

- confrontés au point 3 avec l'utilisation de données de comportement (historiques d'achats qui par définition ne sont pas concernées par le point 2) et confrontés au point 1 si sur la période d'observation des marques ont été supprimées ou introduites.
- confrontés aux points 2 et 3 avec l'utilisation de données de jugement (l'analyse porte alors sur la réalité consommateur et non la réalité de marché). Placé en situation de réponse concernant un univers concurrentiel qu'on lui propose, le consommateur va exprimer son jugement concernant cet environnement précis.

Nous pouvons considérer que le point 2 est à la "réalité consommateur" ce que le point 1 est à "la réalité de marché". Il y a une certaine forme d'équivalence entre les types d'effets observés à ces deux niveaux. Nous étudierons donc simultanément les points 1 et 2 avant d'aborder le troisième point.

SS2-1 ETUDE DE LA MODIFICATION DES RELATIONS DE CONCURRENCE DUE A L'INTRODUCTION (OU AU RETRAIT) D'UN PRODUIT (MARQUE) DANS LE CHAMP CONCURRENTIEL

Un grand nombre de recherches ont été effectuées dans ce domaine, certains modèles ont même été construits selon cette perspective. Un exemple bien connu est le modèle Defender qui a pour but d'établir une stratégie de défense lors de l'arrivée d'un nouveau produit sur un marché (voir Hauser et Gaskin 1984).
Nous regrouperons ces modèles sous l'appellation générale de "modèles de nouveaux produits" (voir Urban et Hauser 1980 pour une présentation complète de ces modèles). Ces recherches peuvent être scindées en deux grands courants:

A-Le premier courant fait l'hypothèse qu'un nouveau produit introduit sur un marché prend sa part de marché en "proportion" des parts de marchés antérieures des produits existants. Le modèle de base de ce courant a été développé par Luce (1959) (voir aussi Gensh et Recker (1979), Pessemier et al (1971), McFadden (1980)).
Dans cette approche, l'effet de similarité du produit introduit parmi les produits existants n'est pas pris en compte lors de la modélisation des préférences ou des choix. Un produit est

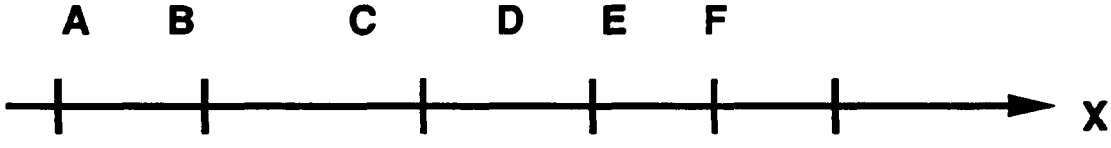
Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

supposé prendre autant aux marques qui lui ressemblent le plus qu'aux autres. Il s'agit de l'hypothèse IIA que nous avons vue à la section 3.
Ce type de modèle fournit en général de bonnes estimations des choix réels et a en plus l'avantage d'être compatible avec des modélisations de la préférence qui peuvent être soit probabilistes soit déterministes.

B-Le deuxième courant fait l'hypothèse qu'un nouveau produit va prendre sa part de marché sur les produits qui lui sont le plus similaires. Cette approche rejette donc l'hypothèse précédente des IIA. Il s'agit alors d'un effet négatif de la similarité: plus une marque est similaire à la nouvelle marque, plus celle-ci va lui prendre de sa part de marché. Cette approche est aussi très attrayante. Elle est compatible avec le Modèle de choix de Tversky d'élimination par les aspects.

La définition correcte du marché apparait comme fondamentale dans le premier cas (celui où on accepte les IIA). En effet, les gains en part de marché du nouveau produit se font sur l'ensemble des marques appartenant à l'univers de concurrence tel qu'il a été défini, il est donc important d'avoir correctement défini cet univers. Dans le deuxième cas où l'hypothèse des IIA, est rejetée, la similarité entre la marque introduite et les autres intervient dans le calcul de la part de marché prévisionnelle. C'est avec les marques les plus proches que s'effectuent les transferts de parts de marché. Les marques les plus proches constituent l'univers concurrentiel au sens strict, univers qu'il est alors plus aisé de définir. L'hypothèse IIA apparait, comme l'ont noté Jeuland (1979) et Bass et Pilon (1980) beaucoup plus tenable si l'analyse est menée résolument au sein d'un sous marché particulier sur lequel est lancé le nouveau produit. Dans un tel univers de concurrence défini au sens strict, les marques présentent des niveaux de substituabilité comparables et l'hypothèse des IIA reste tenable. Ceci n'est plus le cas si l'analyse se situe au niveau d'un marché défini de manière plus large.

C-Une troisième approche (Huber et Puto 1983, Huber, Payne et Puto 1983) introduit l'hypothèse d'un effet de similarité positif du produit introduit sur les items existants qui lui sont similaires. L'idée développée par les auteurs est que l'introduction d'un produit occupant une position extrême sur une dimension de jugement attirerait les préférences des consommateurs (et donc augmenterait les parts de marchés) pour les produits voisins.
Le schéma suivant représente le score perceptuel de plusieurs marques selon une dimension d'évaluation X particulière.



Soient A, B, C, D, E cinq marques existantes, l'introduction de la marque F, à l'extrême de la dimension X, va attirer les préférences vers la zone occupée par C, D et E. Ce faisant, les parts de marché des marques C, D et E vont profiter de ce transfert des préférences puisque ce sont les marques les plus proches de F. Ce phénomène est appelé *"effet d'attraction"*. Ajouter un nouvel item qui *"élargit"* les frontières du marché a deux conséquences selon les auteurs :

- un effet d'*attraction global* car les préférences sont "*attirées*" vers le nouvel item. Cet effet d'attraction bénéficie donc aux marques proches de l'item introduit. Les auteurs ont d'ailleurs admis l'hypothèse que cet effet fonctionne aussi lorsque le produit est introduit dans une zone qui n'est pas extrême sur la dimension d'évaluation.
- un effet de *substitution local* : c'est l'effet classique de non-proportionnalité, selon lequel le nouvel item prend ses parts de marchés essentiellement auprès des items qui lui sont similaires.

Ces deux effets sont justifiés de la manière suivante par les auteurs. Si on appelle "*marque leurre*" la marque F qui est introduite et "*marque cible*" la marque E déjà présente et concurrente des marques B, C et D, il est possible d'identifier (selon Huber et Puto 1983) les trois types d'effets suivants, liés à l'introduction de la marque leurre F :

- *effet de rang* : le faible score de la marque leurre F sur la dimension X où la marque cible E est mal placée va diminuer la faiblesse relative de cette marque E par rapport à ses concurrentes (B, C et D).
- *effet d'accroissement du poids de la dimension* sur laquelle la marque cible E est supérieure à ses concurrentes, ceci parce que l'introduction de la marque F introduit plus d'hétérogénéité des marques sur la dimension X (l'attribut devient alors plus discriminant).
- *effet d'ancrage* : la marque leurre sert de point d'ancrage pour l'évaluation des marques concurrentes, ce qui peut favoriser la marque cible.

Hauser et Gaskin (1986) ont observé cet effet d'ancrage avec l'utilisation du modèle Defender. Plus récemment encore, Ratneshwar, Shocker et Stewart (1987) ont essayé d'expliquer l'effet d'attraction en terme de difficulté de la tâche demandée au répondant. Ces auteurs ont montré que l'effet d'attraction est d'autant plus sensible que les répondants ont une mauvaise connaissance des produits étudiés. Quand le consommateur connaît mal un produit qu'on lui demande d'évaluer, il a tendance à le faire par référence au contexte de ce produit. Lorsque le répondant connait bien le produit qu'il doit évaluer, il porte son jugement indépendamment des autres produits, en quelque sorte de façon "absolue".

Est-ce à dire, comme le suggèrent ces auteurs, que le phénomène d'attraction se limite uniquement à des problèmes de tâche du répondant ? Nous répondons par la négative, parce que le même phénomène doit se retrouver dans la réalité, sur le marché, quand le consommateur compare et évalue des marques pour lesquelles il n'a qu'une connaissance imparfaite .

Huber et Puto (1983) notent encore que les effets d'attraction et de substitution peuvent se compenser et conduire à l'absence d'effet de la similarité de la marque introduite avec ses concurrentes sur la constitution de sa part de marché. Ce phénomène justifierait alors le principe de proportionnalité de Luce et expliquerait la bonne qualité des prévisions faites à partir des modèles d'ordre zéro (modèles qui reposent sur l'hypothèse des IIA).

Cependant, il est une question que semblent se poser peu d'auteurs : "*quelle est la définition qui est donnée au marché pour pouvoir prévoir (avec ou sans l'hypothèse IIA) sur quelles marques le nouveau produit prendra sa part de marché ?*" C'est dire que peu d'auteurs justifient leurs définitions de l'ensemble des marques dont les parts de marché sont influencées par l'introduction (suppression) d'une autre marque. Cette définition est pourtant

de première importance. Elle peut notamment expliquer les résultats contradictoires évoqués précédemment. L'hypothèse d'un processus d'ordre zéro peut être spécifique à une définition de l'univers de concurrence. Lorsque l'ensemble des produits étudié est très similaire, (l'univers de concurrence est défini au sens strict), l'hypothèse IIA parait tenable puisque tous les produits ont un degré de similarité voisin. Dans ces conditions, les processus de choix seraient d'ordre zéro au sein d'un sous marché alors qu'ils peuvent être d'un ordre supérieur quand l'univers de concurrence est élargi, c'est à dire quand on fixe à un niveau plus élevé le degré d'hétérogénéité des produits.

SS2-2 EFFET DE LA MODIFICATION DU NOMBRE D'ALTERNATIVES ETUDIEES SUR LES RESULTATS DE L'ANALYSE STATISTIQUE

Ce problème a été essentiellement étudié à partir des méthodes d'analyse des similarités (MDS). La question essentielle a été de savoir si la représentation d'un groupe particulier de produits est stable quand l'ensemble des produits avec lesquels on associe ce groupe varie (ce qui correspond à différentes définitions possibles du marché). Il s'agit là du problème de la fiabilité structurelle d'une méthode. Green, Maheshwari et Rao (1969) ont montré que la structure des distances entre 5 items reste stable quand on l'analyse successivement avec deux ensembles différents de stimuli. Il semblerait, d'après les auteurs, que ce résultat dépende du niveau de la similarité des items de base avec le référant que l'on fait varier. Si celui-ci est sensiblement différent, la structure des items de base pourrait varier.

Moore et Lehmann (1982) ont montré que l'introduction de nouvelles marques peu différentes des marques anciennes ne modifie pas la perception des marques existantes. Par contre, la stabilité perceptuelle des nouvelles marques introduites est largement fonction de la familiarité qu'ont les répondants avec les caractéristiques du produit. Ceci signifie que pour étudier la perception d'un nouveau produit, il est nécessaire que le répondant se soit préalablement familiarisé avec le produit.

Malhotra (1987), Jain, Malhotra et Pinson (1988), dans une recherche sur la validité et la fiabilité structurelle de l'analyse des similarités (MDS) ont étudié avec précision ce problème. Leurs résultats peuvent être résumés ainsi :

- la représentation par le MDS d'un groupe de marques (ou plus généralement d'objets perçus), quand le nombre d'items parmi lesquels ce groupe est inclus varie, est particulièrement robuste (stable) à une faible ou une forte variation du nombre de ces items.
- la représentation par le MDS d'un groupe de marques est par contre très sensible à la variation de la nature des autres stimuli parmi lesquels est inclus ce groupe de marques.

Que signifient ces résultats dans le cadre de l'analyse de la structure d'un marché ? C'est la définition du marché, c'est à dire la fixation d'un degré d'appartenance (degré d'hétérogénéité) des produits à un même ensemble, qui a le plus de conséquences sur la représentation des produits. Par contre, lorsque ce degré d'appartenance est fixé, le fait d'inclure dans l'analyse un nombre variable de produits a peu d'influence sur les relations existant entre les produits sur lesquels porte plus particulièrement. l'analyse.

Ceci est en concordance avec la réflexion de Shocker, Zahorik et Stewart (1984) suivant laquelle *"un marché est correctement défini (c'est à dire la spécification correcte de l'ensemble des produits à analyser) quand la structure du marché (l'ensemble des relations entre sous-ensembles de produits) existant à une époque donnée n'est pas altérée par l'introduction d'autres produits"*.

Il est donc très important de définir avec précision quel est le marché étudié, moins important ensuite d'inclure ou non la totalité des produits (marques) y appartenant, lorsque l'analyse est menée avec des techniques de MDS. Ceci est logique puisqu'une fois fixées les frontières d'un marché, le degré maximum d'hétérogénéité entre les marques est lui-même fixé et par là-même, la structure qui explique cette hétérogénéité .

Lors de la définition d'un marché, il est plus grave d'introduire une marque différente de toutes les autres que de ne pas prendre en compte une marque appartenant effectivement à ce marché. Dans le premier cas le risque est de modifier la structure reliant toutes les autres marques .

Concernant les autres méthodes et notamment des méthodes classiques comme l'analyse factorielle, l'analyse typologique, l'analyse discriminante qui sont les méthodes les plus utilisées dans les "analyses produit", peu d'études de ce type ont été menées. Nous citerons quand même les articles de Hauser et Koppelman (1979) qui ont effectué une étude comparative de ce type de techniques. D'autre part, Shocker et Srinivasan (1979) ont noté qu'avec l'analyse factorielle ou l'analyse discriminante, la structure obtenue n'est pas fonction du nombre d'alternatives proposées car la méthode repose sur des corrélations entre attributs (notons cependant que l'introduction d'une marque différente peut modifier ces corrélations).

SS3 METHODES DE DEFINITION DES MARCHES

Les méthodes de définition du marché peuvent être groupées en deux catégories en fonction du type de données qu'elles utilisent, données de jugement ou données de comportement.

SS3-1 METHODES PROCEDANT A PARTIR DE DONNEES DE COMPORTEMENT

Comme nous l'avons déjà souligné, le marché est défini ici par rapport à une "réalité de marché", c'est à dire tel qu'il est réellement. Les principales approches sont les suivantes :

* ANALYSES DES ELASTICITES CROISEES

C'est l'approche classique des économistes ; deux produits appartiennent au même marché si les élasticités croisées de leurs demandes par rapport au prix sont positives .

* ANALYSE DES SIMILARITES D'UTILISATION (Cocks et Virt 1975)

Les produits appartenant à un même marché doivent avoir des patterns d'utilisation identiques. Le problème majeur est que les données de panels ne permettent pas de connaître les occasions d'usage des produits. Il est donc nécessaire de mener des études ad hoc, ce qui est coûteux.

* ANALYSE DES TRANSFERTS ENTRE MARQUES

Deux marques appartiennent au même marché si elles entretiennent des relations qui se matérialisent sous la forme de transferts (brand switching). La faiblesse essentielle de cette approche est que le concept de transfert entre deux marques n'a de sens qu'après qu'une définition pertinente ait été donnée au marché, à savoir l'ensemble des marques entre lesquelles peuvent s'effectuer ces transferts. Il est donc, au contraire, nécessaire d'avoir correctement défini le marché avant de pouvoir calculer des transferts. Il ne s'agit donc pas d'une méthode de définition des marchés contrairement à ce qu'ont noté plusieurs auteurs (par exemple Hauser et Urban, 1980).

SS3-2 METHODES PROCEDANT A PARTIR DE DONNEES DE JUGEMENT

Il est important de retenir qu'avec ce type de données le marché est défini selon une logique "réalité consommateur": ce n'est pas le marché réel, mais le marché tel qu'il pourrait être.

* ANALYSE DES PROCESSUS DE DECISION METHODE DES PROTOCOLES

On retient ici les produits (marques) qui ont été envisagés par le consommateur lors de son choix, choix qui est étudié *"in situ"* à l'aide de la méthode des protocoles. Cette approche a été proposée par Bettman (1971, 1976, 1979) dans le but d'étudier les processus de traitement de l'information utilisés par le consommateur. La méthode consiste à accompagner le consommateur sur son lieu d'achat et à lui faire exprimer oralement toutes les étapes de son processus de choix. En particulier il est invité :

- à préciser les produits qu'il considère lors de son choix,
- à spécifier les attributs des produits considérés lors de son choix,
- à indiquer la séquence et la méthode de combinaison des attributs conduisant à son choix

Cette information permet par la suite de reconstituer les séquences d'achats sous la forme d'arbres de décision qui sont ensuite analysés. C'est en essayant de reconstituer le processus de choix que l'ensemble des marques appartenant au marché considéré sera déterminé. Plusieurs réserves sont à prononcer à l'encontre de cette méthode.

Tout d'abord, il existe un biais évident de réponse, l'interviewé n'étant pas en situation naturelle d'achat. D'autre part, les variables de situation d'achat ont une influence considérable sur la nature de l'information fournie par le répondant. La procédure est lourde et complexe. Les processus obtenus au plan individuel sont complexes et difficile à agréger.

*ANALYSE DES SIMILARITES PERCUES ENTRE LES MARQUES

Deux grandes approches sont ici possibles. Une approche dite *"agrégative"* qui consiste à collecter des données de similarités perçues entre les marques. A partir de ces similarités, les techniques d'analyse des similarités (MDS) et de classification permettent de dessiner les contours du marché.

Bourgeois, Haines et Sommers (1979) ont proposé une méthode de groupement direct des produits dans laquelle le consommateur, face à un large ensemble de produits, doit constituer

des groupes, expliquer la signification de ses regroupements et évaluer la similarité des produits dans chaque groupe.
Une deuxième approche dite *"désagrégative"* procède à partir de données perceptuelles de type marque x attribut. Les répondants expriment leurs perceptions des marques selon un ensemble d'attributs. Les techniques d'analyse factorielle ou d'analyse discriminante permettent alors de représenter la similarité perçue entre les marques de manière à déterminer, comme précédemment, les contours du marché.
Nous noterons que ces méthodes (surtout l'analyse des similarités directes) sont mal adaptées pour analyser des ensembles de produits dont les niveaux de similarité sont très variables (c'est à dire analyser simultanément des items très similaires et très dissimilaires).

* ANALYSE DE LA SUBSTITUABILITE PERCUE ENTRE LES MARQUES

Plusieurs approches ont été proposées:

1 ANALYSE DES JUGEMENTS DE SUBSTITUABILITE DES CONSOMMATEURS

Le consommateur est ici invité à exprimer le niveau de susbtituabilité qu'il perçoit entre des produits ou des marques .
Plusieurs approches existent quant à la manière de collecter l'information auprès des répondants.

a) Les techniques de réponses libres (Green, Wind et Jain 1973).

Le répondant doit associer librement les marques qu'il pense substituables entre elles. La méthode fournit deux types de données :

- l'ordre de substituabilité (au niveau individuel)
- la fréquence d'association (au niveau agrégé)

b) La technique des choix forcés **(Urban, Johnson et Hauser 1984)**

L'individu doit indiquer son choix quand sa marque préférée (ou achetée précédemment) est supprimée. Au plan agrégé, cela permet de calculer un indice de substitution entre les marques à partir des probabilités conditionnelles de choix. Ceci peut être aussi effectué à partir des rangs de préférences (voir Urban et Hauser 1980).

c) L'approche "dollarmetric" **(Pessemier et al 1971)**

Pour chaque couple de produits envisagé, il est demandé au consommateur quelle somme il lui faudrait pour renoncer à son choix.

2 ANALYSE DES MATRICES PRODUITS USAGES

Nous développerons ici cette méthode car elle nous semble être, parmi les techniques présentées ici, la méthode la plus prometteuse pour définir les frontières d'un marché à partir du jugement des consommateurs . Elle permet notamment de prendre en compte la variable situation d'usage, ce qui a pour avantage de placer le répondant dans un contexte de réponse précis. De plus, le concept de substituabilité doit être, comme nous l'avons dit à la section 4, défini relativement à une situation d'usage.

Cette méthode, proposée à l'origine par Stefflre (Stefflre 1971, Stefflre dans Myers et Tauber 1977, Stefflre 1979) a été reprise par Belk (1979), Srivastava, Shocker et Day (1978), Day, Shocker et Srivastava (1979). A l'origine cette méthode est destinée à définir, sélectionner et affiner des concepts de produit en vue du lancement d'un nouveau produit. Elle comprend deux grandes phases:

Phase 1 Sélection du concept de produit

Cette phase comprend elle même trois étapes :

a) La sélection d'une cible de marché. Il s'agit ici de définir le marché sur lequel le nouveau produit sera amené à évoluer. Ce marché se définit en terme de type de besoin particulier à satisfaire .

b) Auprès d'un échantillon de vingt à trente répondants on va (en utilisant des techniques d'interview en profondeur) essayer d'établir :

- la liste des produits appartenant au marché défini en terme de besoins
- la liste des différents usages relatifs à ces produits
- les autres produits, non encore cités, qui pourraient être utilisés dans les mêmes situations d'usage.

c) Auprès d'un nouvel échantillon (trente personnes) et sur la matrice obtenue en croisant les produits et les usages définis précédemment, on collecte *"l'adéquation perçue de chaque produit à chacun des usages"*. Cette procédure permet d'effectuer un diagnostic de l'adéquation des produits aux usages, des insuffisances existantes ce qui rend possible la détermination des concepts les plus intéressants.

Dans les phases suivantes, on va :

d) Collecter des jugements de similarité concernant les concepts de produit.
e) Effectuer des mesures de préférences à petite puis à large échelle.
f) Sélectionner le concept de produit le plus prometteur.

Phase 2 Affiner le concept de produit final

Il s'agit ici d'un ensemble de tests qualitatifs et quantitatifs permettant de définir le concept de manière opérationnelle : forme du produit, emballage, contenu de la communication .

Notons que l'objectif de la méthode de Stefflre n'était ni de définir ni d'analyser la structure d'un marché. Ce sont les premières étapes de la phase 1 qui ont été reprises à cet effet. L'avantage de la méthode est, en partant du coeur du marché (à la limite la procédure peut démarrer à partir d'un seul produit), de lister progressivement tous les produits d'un marché qui est défini successivement en termes de situations d'usage différentes.

Nous décrirons maintenant la méthode telle qu'elle a été modifiée par Day et al (1979). La procédure proposée comprend trois phases :

1 - Phase exploratoire

L'objectif est ici de définir, à l'aide de méthodes stimulatives, entretiens de groupes, les situations d'usage associées à un besoin spécifique, de manière à construire la matrice des produits x situations d'usage.

2 - Typologie des situations d'usage

Elle est effectuée à partir d'une Analyse en Composantes Principales de la matrice produit x usage, les situations d'usage tenant le rôle de variables. On obtient alors une typologie d'usages.

3 - Mesure de la substituabilité

Un deuxième échantillon indépendant permet d'évaluer le degré d'adéquation de chaque produit (ou marque ou concept) aux situations d'usage obtenues à partir de la typologie de la phase 2 et ainsi de cerner les contours du marché étudié.

* DEFINITION DU MARCHE A PARTIR DES ENSEMBLES EVOQUES DU CONSOMMATEUR PLACE DANS UN CONTEXTE DE CHOIX PARTICULIER (Chandon et Strazziéri 1986)

Le principe retenu par ces auteurs procède à partir de l'ensemble évoqué du consommateur pour définir les marques appartenant à un marché. Placé dans une situation d'usage particulière (préalablement définie), le consommateur est amené à définir les marques qui sont susceptibles de satisfaire son choix (nous développerons cette approche au Chapitre II).

Au vu de toutes ces approches et définitions d'un marché, nous constatons l'absence de procédure formalisée permettant de définir de manière certaine et unique les frontières d'un marché particulier. Comme le notent Day et Al (1979), *"les frontières d'un marché existent mais elles sont arbitraires et en perpétuel changement"*. De même Dussart (1985) insiste sur le caractère "*arbitraire*" de la frontière d'un marché, ce qui ne signifie pas pour autant qu'il ne faut pas en fixer une. La frontière d'un marché "*...existe pour répondre à des besoins qui sont de mieux comprendre la structure type d'un marché et d'éclaircir, en l'ordonnant, la complexité naturelle de tout environnement concurrentiel*".

"*Arbitraire*" ne doit donc pas être pris dans un sens péjoratif, mais simplement afin de reconnaître que le manager (ou le chercheur) doit lui-même fixer, à un moment donné, avec quel degré de généralité il désire définir son marché. En effet, à un certain niveau de généralité, tout est substituable avec n'importe quoi (un bateau avec un voyage organisé...). Le manager devra notamment fixer à quel niveau de la hiérarchie des produits il compte définir son marché (voir Lunn 1972, Srivastava, Leone et Shocker 1981) :

- différents types de produits, qui peuvent satisfaire des types de besoins assez différents, le lien de substitution est alors considéré à long terme.
- différentes variantes de produits (pour un type donné), le lien de substitution est à considérer à court terme.
- différentes marques (au sein d'une même variante) qui sont alors directement concurrentes.

Ainsi, il est possible de parler de marché soit en termes de types de produits, soit en termes de variantes de produits, soit encore en termes de marques. Nous retiendrons que ces trois niveaux dans la hiérarchie du produit ne servent qu'à fixer un schéma théorique (voir

figure 6 à la section 1). En réalité, ce niveau de généralité peut varier sur un continuum, à chaque niveau du continuum pouvant correspondre une définition du marché particulière. Toutefois, il existe vraisemblablement des discontinuités (les frontières du marché varient alors brutalement) lorsque le niveau de généralité de la relation entre marques varie. C'est peut être à ces niveaux là qu'existent des "entités de marché". Nous remarquerons enfin qu'avec la variation du degré de généralité de la relation entre marques, l'un des deux éléments suivants varie aussi:

- l'unité d'analyse du marché : marque, variante, catégorie de produit ..
- pour une unité d'analyse fixée, le nombre de ces unités dans le marché.

Il convient de souligner qu'entre un manager qui désire savoir contre qui il est concurrent à un moment donné et un manager qui désirerait lancer un produit pour satisfaire un besoin d'une nouvelle manière, ou savoir contre qui il risque d'être concurrent à moyen terme, les niveaux d'analyse ne peuvent pas être identiques. Comme le notent Shocker et al (1979) *"une définition unique du marché est un compromis entre des vues à long et à court terme, ce qui ne peut être cohérent avec les vues des différentes alternatives possibles considérées par les consommateurs, pour une situation d'usage particulière"*.

Une fois le marché défini, c'est à dire l'univers concurrentiel envisagé, les méthodes marketing semblent plus à l'aise pour étudier les relations de concurrence entre les produits de ce marché et le regroupement des produits (marques) en sous-marchés.

Pour conclure cette section consacrée à la définition des marchés nous insisterons sur les points suivants :

1 Dans le cadre de l'Analyse de la structure des marchés, la définition du marché doit se faire en termes de "produits" (marques) par opposition à consommateurs, produits (marques) qui constituent l'unité d'analyse privilégiée.

2 Il ressort de nos analyses que les définitions proposées du concept de *"marché"* ne permettent pas, en général, d'opérationnaliser la définition du marché, c'est à dire de fixer les frontières du marché. Une définition du marché est pourtant un préalable incontournable à l'étude de sa structure. Cette définition est pourtant absente, comme nous le verrons au Chapitre II, dans la majorité des modèles d'analyse de la structure du marché qui sont proposés dans la littérature. La définition du marché semble bien souvent *"s'imposer d'elle même"*.

3 Un marché peut se définir, selon nous, en fonction des deux types de réalités que nous avons étudiées précédemment : la *"réalité consommateur"* et la *"réalité de marché"*.

La définition du marché selon la réalité consommateur procède à partir de données de jugement. Le marché se définit alors en termes de besoins à satisfaire, de situations d'usage du produit. C'est une vision *"large"*, *"potentielle"* du concept de marché et qui convient particulièrement bien lorsque l'entreprise mène des études prospectives afin d'étudier non pas ce qui est mais ce qui pourrait être.

La définition du marché à partir de la réalité du marché adopte un point de vue résolument différent du précédent. Elle utilise des données sur les comportements de choix passés et définit le marché tel qu'il est. Cette démarche doit à notre avis se limiter à définir des univers concurrentiels très directs, c'est à dire à la définition du marché de manière très étroite. Nous pensons que cette démarche est moins riche que la précédente. Les données de comportement d'achat prennent, en effet, en compte les variables de situation d'achat et ne permettent pas de

différencier les différents individus constituant le ménage. Cependant, leur disponibilité en quantité et en précision ne fait que croître et il est nécessaire d'intensifier les recherches afin de mettre au point des méthodes de définition des marchés fiables, à partir de ce type de données.

4 Pour finir, nous présenterons une définition qui n'a pas pour but de donner la vérité en matière de définition d'un marché mais plutôt de préciser les éléments participant à cette définition. Nous proposons la définition suivante : "Un marché est constitué par un ensemble de produits ou de marques entretenant une forme particulière de relation".

Pour définir un marché, il est donc nécessaire :

- de définir la nature de la relation : pour l'étude de la "réalité du consommateur", cette relation s'analyse en terme de substituabilité. Pour l'étude de la "réalité de marché", cette relation s'analyse en terme de concurrence et d'influence respective des actions marketing des marques les unes sur les autres.
- de définir l'objet de la relation, ce qui revient à préciser le degré de généralité du marché défini. L'objet de la relation peut être la satisfaction d'un besoin du consommateur, deux marques sont substituables parce qu'elles se proposent de satisfaire le même besoin du consommateur. L'objet de la relation peut être défini en terme de puissance et de "relations conflictuelles entre marques" (correspondant à une volonté de prise de part de marché), il s'agit alors de concurrence (voir tableau suivant).

TABLEAU 7 NIVEAUX D'ANALYSE ET TYPE DE RELATIONS ENTRE MARQUES

		NATURE DE LA RELATION ENTRE DEUX MARQUES	
		SATISFAIRE LE MEME BESOIN DU CONSOMMATEUR	PRENDRE DES PARTS DE MARCHE SUR LES CONCURRENTS
REALITE ETUDIEE	CONSOMMATEUR	SUBSTITUABILITE	X
	MARCHE	SUBSTITUBILITE REVELEE	CONCURRENCE

CHAPITRE DEUXIEME MESURE DE LA RELATION ENTRE LES MARQUES

Dans ce chapitre, nous présenterons de manière exhaustive les modèles d'analyse de la concurrence qui ont été proposés dans la littérature. Lors de cette présentation nous nous attacherons à isoler l'indice de concurrence utilisé. En effet, bien souvent, les modèles d'analyse de la concurrence sont présentés comme un tout, sans distinction entre le choix de l'indice de mesure de la relation entre les marques et la méthode utilisée pour analyser ces indices. Il est pourtant, selon nous, indispensable de bien identifier la nature de cette mesure, puisque c'est elle qui revêt, du point de vue marketing, le plus d'importance. Nous étudierons les différentes mesures de la relation entre les marques qui ont été proposées dans la littérature au travers de cinq sections qui correspondent aux concepts servant de base à l'étude de ces relations. Ces concepts sont les suivants:

- élasticités croisées (section 1)
- transferts entre marques (section 2)
- part de choix (section 3)
- intervalles inter-achats (section 3)
- jugement des consommateurs: perception, ensembles évoqués, choix (section 4)
- modèles de préférences (section 5)

SECTION 1 MESURE DE LA CONCURRENCE PAR LES ELASTICITES CROISEES

Issue de la théorie économique classique (voir Hicks 1946) et des travaux économétriques, cette mesure était conçue à l'origine pour étudier les liens de substituabilité et de complémentarité entre des produits appartenant à différentes classes Le consommateur qui dispose de ressources limitées affecte son budget entre différentes catégories de biens. Les élasticités représentent alors un indice de mesure de substituabilité qui peut être qualifiée comme étant "horizontale" (entre différentes catégories de produits). En effet, comme les produits sont très peu différenciés, qu'il existe un prix d'équilibre unique et optimal, il ne présentait alors aucun intérêt d'étudier la substituabilité de produits appartenant au même marché .

Une telle approche de la substituabilité des biens peut être mise en parallèle avec les théories de Maslow (1952) sur la hiérarchie des besoins. Le consommateur peut effectuer des substitutions de type vertical en affectant son budget à des catégories de biens qui satisfont différents niveaux de besoins dans sa hiérarchie. Coursey (1988) a développé un modèle de choix compatible avec une telle approche.

La mesure de l'élasticité croisée des quantités achetées aux prix des produits est définie de la manière suivante.
Soient:

D_i : la fonction demande à la marque i
D_j : la fonction demande à la marque j
P_i et P_j : les prix respectifs de i et j
E_i/P_j l'élasticité de la marque i au prix de j :

$$E_i/P_j = \frac{\partial D_i / \partial P_j}{P_j / D_i} \qquad \text{de même: } E_j/P_i = \frac{\partial D_j / \partial P_i}{\partial P_i / D_j}$$

Deux produits sont considérés comme :

- substituables lorsque leurs élasticités croisées sont positives
- complémentaires lorsque leurs élasticités croisées sont négatives.

En marketing, l'objectif étant d'apprécier la sensibilité des ventes aux actions marketing ainsi que la mesure de la concurrence entre les marques, les élasticités sont calculées entre des marques qui sont directement substituables. Nous qualifierons ce type d'élasticité comme étant "verticale" (au sein de la même catégorie de produits). Il est ici nécessaire d'avoir opéré au préalable une définition du marché étudié.

Une fois replacé dans un contexte Marketing le concept de concurrence recouvre une réalité plus riche qu'en Economie. En effet, en Marketing, vont intervenir beaucoup d'autres facteurs que le prix sur les relations de concurrence entre les marques. Il est alors possible de concevoir et de calculer plusieurs types d'élasticités (simples ou croisées) : élasticité à la pression publicitaire, élasticité au linéaire en magasin, élasticité à la promotion en magasin, élasticité à la force de vente, etc...D'une manière plus générale il

est possible de concevoir un coefficient d'élasticité à l'ensemble des actions marketing, ce coefficient étant une mesure du lien concurrentiel existant entre deux marques.

L'élasticité est donc un concept à la fois très naturel et très séduisant au plan théorique. Elle caractérise, comme le note Cooper (1988), *"la manière dont les conditions du marché se transforment en changements des parts de marché ou des ventes"*. La concurrence mesurée par l'élasticité établit un lien entre les actions de l'entreprise et de ses concurrentes et les réactions du marché à ces actions. Malgré son intérêt théorique, l'élasticité est en pratique difficile à mettre en oeuvre. Les problèmes techniques liés à l'estimation des élasticités constituent certainement l'obstacle majeur à leur utilisation en tant que mesure de la concurrence entre les marques. Les approches de l'analyse de la concurrence entre les marques utilisant le concept d'élasticité sont en effet étonnamment rares compte tenu de la cohérence de ce concept avec ceux de substituabilité ou de concurrence. Nous citerons successivement les obstacles suivants à l'utilisation de l'élasticité :

1 Problème d'opérationnalisation des variables étudiées :

En effet, comment opérationnaliser de manière fiable l'élasticité des ventes à la distribution, à la communication? Le problème apparait déjà plus compliqué que l'estimation des élasticités aux prix.

2 Le choix de l'unité d'observation :

Doit-on estimer les élasticités au niveau d'une marque dans un point de vente particulier ou doit-on effectuer cette estimation au niveau des ventes totales de la marque ou encore, à l'extrême, agréger différentes marques de manière à faire l'étude sur une "Unité Stratégique de Base" (USB) ? Dans ces deux derniers cas, les analyses se heurtent à des problèmes d'agrégation. Par exemple, pour estimer une élasticité au prix sur plusieurs points de vente ne suivant pas la même politique de prix, comment opérationnaliser la variable prix ?

3 Les techniques d'estimation posent de sérieux problèmes concernant :

- le nombre de paramètres à estimer
- l'autocorrélation des erreurs
- la prise en compte de retards échelonnés (par exemple retards échelonnés de l'effet des variations des prix sur les comportements d'achats)

4 Les mesures des ventes peuvent être affectées par d'autres facteurs de l'environnement qui ne sont pas maîtrisés (disponibilité des produits, variation de prix d'autres marques, politique de promotion...).

5 On pose en général l'hypothèse que la marque concurrente j, ne réagit pas en terme de prix aux variations de prix de la marque i, pour le calcul de Ei/Pj.

6 Cette mesure est inopérante lorsque les prix sont stables ou que tous les prix varient en même temps .

7 L'élasticité n'est en général, pour une marque i, pas constante (concavité des courbes de demande), ce qui peut compliquer l'utilisation de l'indice. Pour deux marques, la concurrence peut varier ainsi en intensité par le simple fait d'une variation globale du niveau de la demande (par exemple l'effet d'une promotion), les structures de préférence des consommateurs restant stables par ailleurs).

8 Les études économétriques nécessaires sont en général longitudinales ce qui pose le problème de l'hypothèse de la stabilité des structures de préférences (une modification des goûts des consommateurs étant à craindre pendant la période d'estimation des élasticités).

Cet indice met, par contre, clairement en avant que la relation de concurrence n'est pas symétrique. Une mesure d'élasticité permettant de prendre en compte les critiques apportées plus haut serait ainsi une mesure fort intéressante de la concurrence. Cooper (1988), avec le modèle qui est présenté plus loin, a apporté une réponse à certains de ces problèmes :

- l'effet de retard échelonné
- la prise en compte de l'évolution structurelle des préférences due à l'écoulement du temps

D'une manière générale, l'information disponible dans les panels de consommateurs "modernes" facilitera de plus en plus l'estimation de ce type de paramètres.

SS1 ESTIMATION DES ELASTICITES EN COUPES INSTANTANEES

Cette méthode d'estimation nécessite l'agrégation des produits en "Unités Stratégiques de Base". L'agrégation en unités stratégiques de base (USB) se fait sur la base de leur similarité en terme de type d'industrie, position dans le cycle de vie du produit, fréquence d'achat du produit. De par son caractère agrégé, cette analyse relève essentiellement du domaine du Marketing Stratégique. Nous présenterons ici le modèle de Hagerty, Carman et Russell (1988) qui illustre cette approche et qui est estimé à partir de données PIMS. Ce modèle exprime les ventes d'une unité stratégique de base particulière en fonction des ventes totales de l'industrie et des variables du mix relatives à cette unité.

$$S_t = \beta 0 \cdot I_t^{\mu} \cdot X_{1t}^{\beta_1} \ldots\ldots X_{nt}^{\beta_n} \cdot \exp(u_t)$$

Modèle multiplicatif qui peut s'écrire :

$$\log S_t = \log \beta 0 + \mu . \log I_t + \sum_{i=1}^{n} \beta_i . \log X_{it} + u_t$$

S_t Ventes en l'instant t de l' USB
I_t Ventes totales de l'industrie de l'USB en l'instant t
X_{it} ième variable du mix
u_t erreur aléatoire distribuée selon la loi normale
μ.est le paramètre caractérisant l'influence de l'industrie sur l'UBS
Les β_i (β_1 à β_n) représentent les élasticités nettes, c'est à dire l'effet des dépenses marketing de la USB sur les quantités vendues, nettes des effets dûs à la réponse des concurrents.

SS2 ESTIMATIONS DES ELASTICITES A PARTIR DE DONNEES LONGITUDINALES

Le niveau d'analyse dans cette approche est la marque. Elle permet notamment d'estimer des élasticités croisées entre marques, élasticités qui sont la mesure de la relation concurrentielle entre les marques.

Vanhonnacker (1984) est un des premiers auteurs à avoir proposé l'utilisation des coefficients d'élasticité croisée aux variables du mix comme indice de mesure de la concurrence entre les marques. Selon l'auteur, les estimations traditionnelles de l'élasticité souffrent des limites suivantes:

- elles nécessitent beaucoup d'observations
- il existe des liens de multi-colinéarité entre les variables (par exemple entre les prix des marques i et j au cours du temps).
- les variables explicatives possèdent peu de variance

Cependant, selon l'auteur, la disponibilité croissante des données de panels fournit une alternative intéressante à toutes ces limites :

- on dispose de beaucoup d'informations sur une période courte ce qui limite les problèmes de multi-colinéarité et de modification des élasticités au cours du temps.
- comme on observe les données à un niveau faiblement agrégé, les variables explicatives contiennent plus de variance.

L'auteur a insisté sur les qualités de cette mesure de la concurrence qui tient au fait qu'elle prend directement en compte les effets des variables du mix : prix, promotion etc

Il propose le modèle d'équations simultanées suivant. Pour chaque marque il est possible d'écrire :

$$V_{it} = \alpha . V_{it-1}^{\gamma_{0i}} \; P_{1t}^{\gamma_{1i}} \ldots P_{jt}^{\gamma_{ji}} \ldots P_{nt}^{\gamma_{ni}} \; A_{1t}^{\gamma_{1i}} \ldots A_{jt}^{\gamma_{ji}} \ldots A_{nt}^{\gamma_{ni}}$$

où :

V_{it} sont les ventes de la marque i à la semaine t (i = 1 ...n)
P_{it} est le prix de la marque i à la semaine t
A_{it} est la quantité de publicité dépensée par la marque i à la semaine t (mesurée par la pige publicitaire)

L'exposant de P_{jt}: γ_{ji} est l'élasticité croisée du prix de la marque i relativement au prix de la marque j, j variant de 1 à n.

L'exposant de A_{jt}: γ_{ji} est l'élasticité croisée de la communication de la marque i relativement à la publicité de la marque j.

Les coefficients significativement non-nuls sont retenus par la suite comme mesure de la concurrence. L'auteur considère et c'est important pour la signification de la mesure

qu'il propose, que des coefficients d'élasticité négatifs et positifs ne constituent pas les opposés d'un même continuum mais deux facettes du concept de concurrence. Nous regrettons pour notre part que l'auteur n'ait pas mieux développé ce point. Ceci pourrait signifier par exemple que substituabilité (élasticité croisée négative) et complémentarité (élasticité croisée positive) ne sont pas les opposés d'un même continuum mais deux concepts différents.
Ainsi, deux types d'analyses sont menées, en constituant des matrices de concurrence à partir d'indices positifs d'une part et négatifs d'autre part. Ces matrices sont par la suite analysées séparément en utilisant l'analyse des similarités.
L'intérêt de cette approche est de permettre de pratiquer des mesures de concurrence selon différentes variables du mix (prix, publicité), ou bien globalement si l'ensemble des élasticités aux différentes variables du mix (d'un signe donné) sont analysées simultanément. L'auteur a insisté d'autre part sur le caractère non-symétrique de la mesure de la concurrence qui constitue selon lui une donnée fondamentale de la concurrence.

D'autres approches ont été proposées pour estimer les élasticités à partir de données inter-temporelles.
Cooper (1988) a utilisé le modèle d'attraction suivant pour estimer des structures d'élasticités croisées :

$$A_{it} = \exp(\alpha_i).\prod_{k=1}^{K}[f_t(X_{kit})^{\beta_{ki}}].\prod_{(k^*j^*)\varepsilon\, Ct}[f_t(X_{k^*j^*t})^{\beta_{k^*ij^*}}].\exp(u_{it})$$

Ait est l'attraction de la marque i en t

$$Mit = \frac{A_{it}}{\sum_{j=1}^{N} A_{jt}}$$

i, j =1 à N sont les marques
t = 1 à T représente le temps
k = 1 à K représente les actions marketing possibles
αi est l'intercept de la marque i
X_{kit} est la valeur de l'action marketing k pour la marque i en l'instant t
ft est une fonction transformant X_{kit} en échelle de ratio
βki est un paramètre de sensibilité de la part de marché de la marque i en t (M_{it}) aux variations de ft(X_{kit})
$\beta_{k^*ij^*}$ est le paramètre de l'effet croisé sur i de la variable marketing k* utilisée par la marque j
Ci ensemble des effets croisés relatifs à la marque i pour l'action marketing k*, relativement à la marque j*
u_{it} est un terme d'erreur aléatoire

Estimation des élasticités à partir de ce modèle, l'exemple des prix

La matrice E_t des élasticités au prix a pour terme général: e_{ij}, l'influence du prix de la marque j sur la part de marché de i . Il est possible d'écrire:

$$E_t = (I - J.M'_t)\ B.S_t$$

I Matrice identité (N,N)
J Vecteur unitaire (N,1)
M_t Vecteur (N,1) des parts de marché en t
B Matrice (N,N) des coefficients de prix, de terme général $ß_{kij}$*
S_t Matrice (N,N) dont le terme général S_{ijt} représente le pourcentage de changement de $f_t(X_{kit})$ pour une variation de X_{kjt} de un pourcent (k étant le prix).

$$S_{ijt}^{(k)} = \frac{\partial f_t(X_{kit})}{f_t(X_{kit})} \cdot \frac{X_{kjt}}{\partial X_{kjt}}$$

$e_{ijt}^{(k)}$ peut finalement s'écrire :

$$e_{ijt}^{(k)} = ß_{ki}.S_{ijt}^{(k)} + \sum_{kj^* \varepsilon\ Ci} ß_{kij^*} \cdot S_{j^*jt}^{(k)} - \sum_{i'=1}^{N} M_{i't}.ß_{ki'}S_{i'jt}^{(k)} + \sum_{k,j' \varepsilon\ Ci'} ß_{ki'j'}S_{j'jt}^{(k)}$$

La matrice S_t est donnée par la forme :

$$S_t = \frac{1}{\sigma_{Xt}} \cdot \left[(I - \frac{1}{N} \cdot J.J') - \frac{1}{N} \cdot Z_t.Z'_t\right] \cdot D_{Xt}$$

σ_{Xt} est l'écart-type des prix sur la période
Z_t est le vecteur des prix standardisés des marques, de terme général :

$$z_{it} = (X_{it} - X_t)/\sigma_{Xt}$$

X_{it}: prix de la marque i en t
X_t: moyenne des prix des marques étudiées en t
D_{Xt}: matrice diagonale (N,N) contenant les prix des marques .

Pour chaque période d'observation, on obtient une matrice Et (N,N) des élasticités croisées entre les marques. Cooper a proposé une décomposition très intéressante du coefficient d'élasticité en "*puissance de la marque*" et "*réceptivité de la marque*". Ces deux notions caractérisent la non-symétrie de la relation de concurrence. La puissance caractérise la capacité générale de la marque à influencer les autres marques. La réceptivité caractérise la propension générale de la marque à être influencée par les autres marques.

Ainsi, Eij est écrite comme étant le produit scalaire entre le vecteur de puissance de i sur les autres marques et le vecteur de réceptivité de la marque j par rapport à l'ensemble des autres marques. Cette modélisation nous parait très pertinente.

SS3 ESTIMATION DES ELASTICITES A PARTIR DES FONCTIONS D'UTILITE DES CONSOMMATEURS

Kamakura et Russell (1989) ont repris l'approche de Grover et Srinivasan (1987) qui consiste à segmenter les consommateurs et représenter la structure du marché à partir d'une analyse en classes latentes (voir section 2). L'intérêt de leur approche est d'intégrer l'élasticité par rapport au prix qui permet notamment de lever l'hypothèse de stationnarité des préférences, puisqu'il est ensuite possible d'estimer l'utilité d'une marque pour différents niveaux de prix.
L'utilité pour une marque est exprimée, comme dans le modèle d'utilité classique, en y ajoutant cependant un paramètre lié à l'élasticité aux prix.

$$V_{jkt} = u_{jk} + \beta_k . X_{jkt} + \varepsilon_{jkt}$$

V_{jkt} est l'utilité de la marque j pour le consommateur k, en l'instant t .
u_{jkt} est l'utilité intrinsèque de la marque. Ce concept n'est malheureusement pas développé par les auteurs, il correspond à notre avis à un fourre-tout

ε_{jkt} est la composante aléatoire de l'utilité. Les ε_{jkt} sont indépendants et de même loi .
X_{jkt} est le prix de la marque j pour le consommateur k en l'instant t
β_k représente la sensibilité du consommateur k au prix

Bucklin et Srinivasan (1989) proposent l'estimation des élasticités à partir de données de jugement et d'un modèle d'utilité probabiliste qui a la forme classique du modèle Multinomial Logit (MNL).

$$P_{cih} = \frac{\exp (\beta_c{}^h . V_{ci}{}^h)}{\sum_h \exp (\beta_c{}^h . V_{ck}{}^h)}$$

$P_{ci}{}^h$ est la probabilité de choix de la marque i par le ménage h appartenant au groupe d'usage c . Les groupes d'usage sont établis après analyse des déclarations d'usage des marques dans le ménage, par les différentes personnes composant ce ménage .
βch est un paramètre d'échelle estimé pour chaque groupe c .

$V_{ci}{}^h = U_{ci}{}^h - Pi$ où:

P_i est le prix de la marque i

V_{ci}^h est la valeur perçue de la marque i pour le ménage h appartenant au groupe c
U_{ci}^h est l'utilité de la marque i pour le même ménage

Une variation du prix (les auteurs proposent une variation de 20%) permet d'estimer P*cih qui est la nouvelle valeur de Pcih après variation du prix, les paramètres du modèle restant constants. Il est alors possible d'estimer l'élasticité de la marque au prix .
Cette analyse est selon nous très novatrice. Elle propose (entre autres) une estimation des coefficients d'élasticité à partir de données de jugement et d'une modélisation de l'utilité qui nous apparait tout à fait réaliste.

Allenby (1988) a proposé une démarche dans laquelle les élasticités croisées sont, comme précédemment, établies à partir du modèle d'utilité probabiliste. La probabilité de choix d'une marque i, au sein d'un ensemble homogène de marques S, par un consommateur h, est fonction de son utilité. Cette utilité est la différence entre l'évaluation de la marque selon ses caractéristiques et le logarithme de son prix.

$$Pr(i) = \frac{\exp [v_i/l_S]}{\sum_j [v_j/l_S]}$$

l_S est une mesure du niveau moyen d'indépendance entre les marques, au sein du sous-marché S
Pr(i) est la probabilité de choix de la marque i
v_i est l'utilité associée à la marque i . V_i peut s'exprimer de la manière suivante :

$$v_i = X'_i . \beta_k - \beta_{Ph} . \ln (P_i)$$

où:
X_i est le vecteur des attributs de la marque i
β_k est le vecteur des coefficients d'importances des attributs pour le consommateur k
β_{Ph} représente l'aversion du consommateur pour le prix.

Le coefficient d'élasticité croisée est obtenu par différenciation de Pr(i) par rapport au prix. Le modèle est estimé à partir des parts de marché des marques. L'avantage essentiel de cette approche (ainsi que la précédente proposée par Bucklin et Srinivasan) est qu'elle permet d'estimer les élasticités dans un cadre conceptuel compatible avec la théorie économique et marketing du comportement du consommateur et non par une simple régression des quantités selon les prix. De ce fait découle un avantage pratique important, il n'est plus nécessaire d'observer des variations de prix et des quantités pour estimer les élasticités. Celles-ci sont le résultat de simulations, une fois que les paramètres du modèle d'utilité individuel.ont été estimés. Une partie des problèmes techniques que nous avons exposés précédemment ne se posent plus. Par contre, il est nécessaire de s'assurer que les paramètres estimés au plan individuel sont fiables et surtout ne varient pas avec des variations des variables du mix.

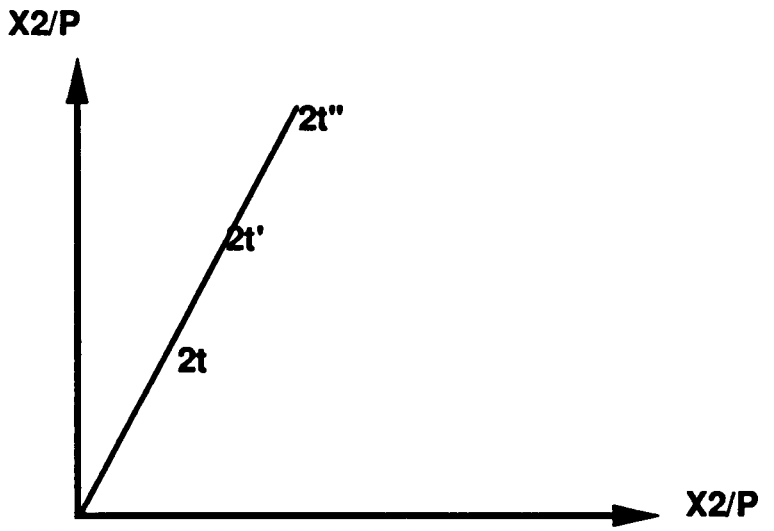
Dans ce modèle, les coefficients d'élasticité sont par la suite rendus symétriques en appliquant une transformation mathématique de manière à pouvoir les analyser à l'aide des techniques classiques de scaling (ALSCAL, voir SAS 1981). Ceci nous parait être une réduction de la richesse de l'information contenue dans ces coefficients.

Le domaine de la recherche portant sur la mesure de la concurrence à partir de coefficients d'élasticité s'enrichit, comme nous avons pu le constater en étudiant les publications récentes, ce qui contraste avec le peu d'intérêt qu'avait suscité ce domaine de recherche tout au long de la dernière décennie. C'est, selon nous, une voie très prometteuse pour la mesure de la concurrence entre les marques et la représentation de la structure des marchés.
Pour terminer cette section, nous présenterons une approche particulière de la mesure de la concurrence entre les marques qui repose en grande partie sur le concept d'élasticité. Il s'agit d'une version du modèle Defender proposé par Hauser et Shugan (1983).

SS4 MODELE D'ANALYSE DE LA STRUCTURE DU MARCHE DERIVE DU CONCEPT D'ELASTICITE

Shugan (1987) a proposé une extension du modèle Defender (voir Hauser et Shugan 1983, Hauser et Gaskin 1986, pour une présentation du modèle) qui permet d'obtenir des cartes des positions concurrentielles des marques. Defender est une modélisation du comportement du consommateur d'essence essentiellement *"économique"*, il dérive du modèle de Lancaster (voir Lancaster 1966).
Nous allons nous attacher ici à présenter la mesure de concurrence entre les marques qui est utilisée dans ce modèle. Son intérêt réside dans le fait qu'il ne nécessite pas de données d'historiques individuels d'achats mais simplement des données de parts de marché des marques en différents points du temps.
Cette approche utilise les cartes par dollar du modèle Defender. La carte par dollar est obtenue en standardisant les positions des marques selon chacun des attributs par leur prix (imaginons une carte d'analyse perceptuelle sur laquelle la position de chaque marque sur chaque attribut serait divisée par le prix de la marque). Le graphique suivant représente une marque dans un espace par dollar à deux dimensions, deux attributs X1 et X2 . Chaque marque est représentée par une droite passant par l'origine O (puisque les attributs sont standardisés par le prix, à un prix nul correspond une valeur par dollar infinie pour n'importe quel attribut et n'importe quelle marque).

GRAPHIQUE 1 ESPACE ATTRIBUTS PAR DOLLAR



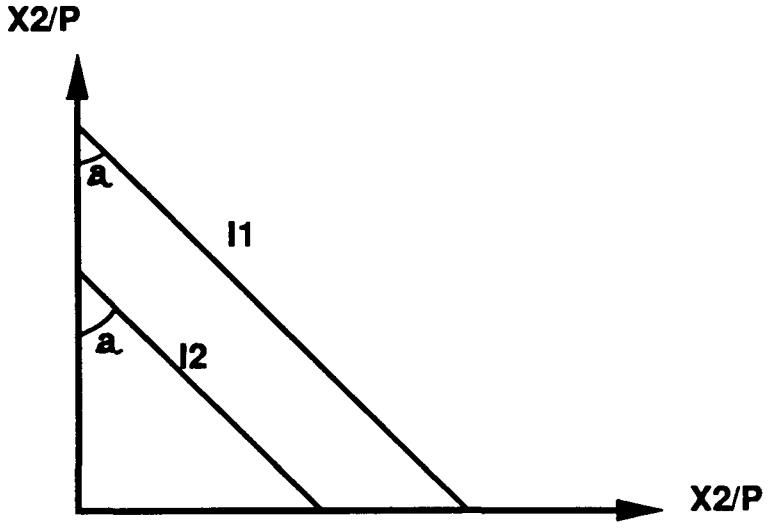
Lorsque le prix d'une marque varie, la trajectoire de la marque suit une droite passant par l'origine. Dans cet espace par dollar, les consommateurs choisissent la marque qui maximise leur fonction d'utilité par dollar. Pour deux attributs X1 et X2, j est choisie si :

$$\max (u) = \sigma_1 . X_{1j}/P_j + \sigma 2 . X_{2j}/P_j$$

σ_1 et σ_2 représentent l'importance des attributs X_1 et X_2 pour un consommateur aléatoire .
Pj est le prix de la marque j

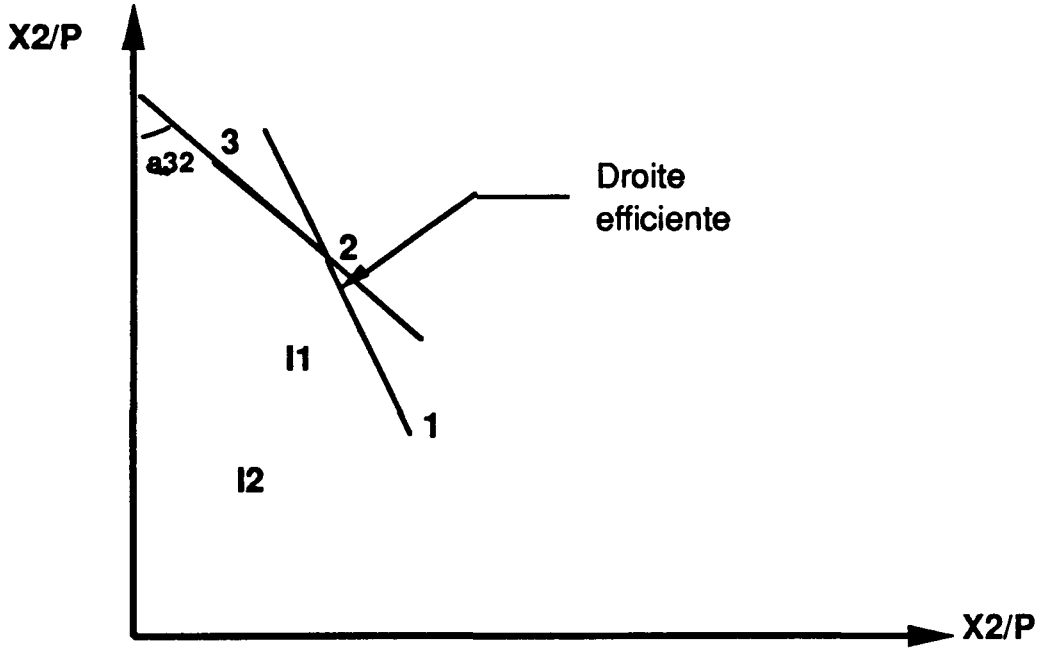
Chaque individu est alors caractérisé par sa droite d'indifférence (droite d'iso-utilité). Cette droite est fonction des poids qu'il accorde aux différents attributs, soit σ_1 et σ_2. L'angle que fait cette droite avec l'axe des ordonnées est donné par: $\alpha = \text{Arctg} (\sigma_1 / \sigma_2)$
Chaque individu est ainsi défini par l'angle α que fait sa droite d'indifférence avec l'axe 2. Si α est faible, le consommateur préfère une unité de X_1 à une unité de X_2 .

GRAPHIQUE 2 DROITE D'ISO-UTILITE PAR DOLLAR



Toute marque située sur la courbe d'iso-utilité la plus éloignée de l'origine, ici I1, est préférée à toute marque située sur une droite moins éloignée. Le consommateur est indifférent à des marques situées sur la même droite.

GRAPHIQUE 3 DROITE EFFICIENTE ET POSITION DES MARQUES



La droite efficiente joint les marques 1, 2 et 3 qui ne sont dominées sur la totalité des dimensions par aucune autre marque .

Pour deux marques i et j, la condition pour qu'un consommateur choisisse la marque j par rapport à la marque i sera :

$$(\sigma_1.X_{1j} + \sigma_2.X_{2j}) / P_j > (\sigma_1.X_{1i} + \sigma_2.X_{2i})/P_i$$

Si l'ensemble des marques est ordonné de la gauche (faibles valeurs sur X1) vers la droite (fortes valeurs sur X_1) on dira que j est préféré à i si :

$\alpha < \alpha_{ijt}$ pour $j < i$ (j est à gauche de i)
ou si $\alpha > \alpha_{ijt}$ pour $j > i$ (j est à droite de i)

Il est alors possible d'écrire la probabilité qu'un consommateur aléatoire choisisse la marque j. Cette probabilité est la part de marché de j sur la période t : mjt.

$$mjt = F(\alpha_{j,j+1}, t) - F(\alpha_{j-1,j}, t)$$

F est la fonction de répartition cumulée des préférences des consommateurs (les angles α que font les droites de ces consommateurs avec l'axe des ordonnées).
j+1 est la marque située immédiatement à droite de j selon l'axe X_1
j-1 est la marque située immédiatement à gauche de j selon l'axe X_1

Graphiquement, ce résultat signifie que la part de marché de la marque j est égale à la proportion d'individus ayant leur angle α d'iso-préférence compris entre $\alpha_{j,j+1}$ et $\alpha_{j-1,j}$.

De la même manière il est possible d'écrire que la relation entre deux marques pour la période t est fonction de l'angle que font ces deux marques j et j+1 entre elles :

$$\alpha_{j,j+1,t} = F^{-1} \left(\sum_{i=1}^{j} mit \right)$$

Cependant, connaître l'angle que font deux marques entre elles ne suffit pas pour les replacer sur une carte. Le modèle les placera finalement en étudiant l'évolution de leur trajectoire sur plusieurs périodes lorsque les prix, donc les parts de marché, varient.
Il est important de noter que l'objectif de la méthode étant de placer les marques sur une carte, ne connaissant pas cet espace, on ne sait pas ordonner les marques de la gauche vers la droite. Par conséquent, Shugan propose de tester l'ajustement du modèle pour tous les ordres différents possibles (il y a N! ordres différents entre N marques !) et de conserver l'ordre pour lequel l'ajustement est le meilleur.
Selon cette approche, la relation entre deux marques adjacentes est évaluée par l'angle que font ces deux marques $\alpha_{j,j+1}$ en l'instant t. Nous remarquerons que la localisation d'une marque n'est pas un indicateur de sa part de marché. Par contre, la proximité entre deux marques est un indicateur d'une forte élasticité croisée au prix entre les deux marques. La proximité sur l'espace par dollar s'interprète donc en terme d'importance de l'élasticité croisée. Ainsi, deux marques dont on sait qu'elles ont une forte élasticité croisée au prix devraient se retrouver proches dans l'espace par dollar.
Le concept théorique sous-jacent, dans ce modèle, est l'élasticité croisée au prix même si celle-ci n'est pas directement spécifiée. En effet, l'analyse repose, dans un premier temps,

sur la construction de courbes d'iso-utilité et de la frontière efficiente dans un espace attribut par dollar. La position des marques (donc la frontière efficiente) est estimée à partir d'une étude dynamique puisque finalement une seule position par marque est compatible avec les variations de parts de marché liées aux variations de prix.

LES AVANTAGES ET LIMITES DE CETTE APPROCHE

Cette approche est très séduisante au plan théorique et pratique. En particulier, les auteurs ont repris et enrichi tout un corps d'hypothèses issu de la micro-économie et en ont fait un instrument destiné à l'étude de problème de Marketing. Ce pont entre les deux disciplines mérite d'être souligné car il est rare.

Au plan pratique, les seuls intrants nécessaires sont les parts de marché et les prix correspondants au cours du temps. L'approche de Shugan souffre pourtant de plusieurs limites :

- les courbes d'iso-utilité sont linéaires, ce qui signifie qu'il existe une relation linéaire constante entre utilité et attribut par dollar. Les économistes considèrent pourtant, en général, que cette relation est non linéaire.
- puisque l'ordre des marques selon l'attribut X_1 n'est pas connu, il est nécessaire de renouveler l'analyse autant de fois qu'il y a d'ordonnancements possibles des marques selon cet attribut. Ceci limite considérablement le nombre de marques qu'il est possible d'analyser. D'autre part, différents ordres peuvent aboutir à des qualités d'estimation comparables, dans ce cas la solution n'est plus unique.
- l'étude de la relation entre marques n'est possible que pour deux marques adjacentes selon un ordre hypothétique de marques. Une mesure partielle pour deux marques est donc fortement sujette à l'aléa résident sur le choix de l'ordre des marques.
- le modèle fait l'hypothèse que toutes les marques se trouvent sur la frontière efficiente (voir graphique 3). Ceci signifie que toutes les marques ne sont jamais vendues à un prix tel qu'elles soient dominées sur l'ensemble des attributs. L'hypothèse du modèle est donc que toute marque non efficiente a une part de marché nulle et ne peut donc entrer dans le modèle.
- comme la majeure partie des analyses reposant sur les élasticités, il faut avoir observé des variations significatives de prix sur la période d'étude. Or, dans une analyse agrégée au niveau des parts de marché, il est peu probable d'observer une variance pertinente des prix.
- l'analyse n'a été développée que dans un espace à deux dimensions ce qui est une limitation considérable du phénomène concurrentiel. Sa généralisation à trois, voire plus, dimensions est, de l'avis de l'auteur, fort difficile.

Pour conclure cette section, nous soulignerons que les nouvelles approches que nous avons présentées dans les paragraphes SS2 et SS3 nous paraissent très attractives. Elles méritent et connaissent actuellement de nouveaux développements en matière de recherche. Dans la section suivante, nous allons présenter des indices de concurrence construits à partir de données de transferts entre marques.

SECTION 2 INDICES CONSTRUITS A PARTIR DES TRANSFERTS ENTRE MARQUES

Cette approche utilise, pour mesurer la concurrence entre les marques, les historiques d'achats des consommateurs qui sont disponibles dans les panels de consommateurs. Depuis la proposition de Lehmann (1972) d'utiliser des indices de transferts entre les marques pour mesurer la concurrence, ce type d'indice à connu un grand succès auprès des chercheurs en Marketing mais aussi des praticiens. A la base de l'utilisation de ce type d'indices, il y a la proposition suivante : " *La concurrence entre deux marques peut être appréciée par une étude des transferts d'achats de l'une de ces deux marques vers l'autre*". Lorsqu'un individu change d'une marque pour une autre, ce comportement de transfert est l'expression de la relation de concurrence qu'entretiennent ces deux marques. La caractéristique de cette mesure est d'être dynamique.

Un grand nombre de modélisations de la concurrence ont été proposées à partir de l'observation des transferts entre marques. On peut les étudier en distinguant leur caractère :

- déterministe ou stochastique
- individuel ou agrégé
- reposant ou non sur une modélisation du processus de choix du consommateur

Ces indices sont généralement calculés (ou estimés) à partir de données d'historiques d'achats des consommateurs qui figurent dans les panels du même nom. Avant d'étudier ces indices, nous commencerons par présenter les problèmes méthodologiques liés à l'utilisation de données d'historiques d'achats.

SS1 PROBLEMES METHODOLOGIQUES LIES A L' ANALYSE DES HISTORIQUES D'ACHATS DES CONSOMMATEURS

La disponibilité de ce type de données ainsi que leur qualité (notamment pour les données du type "Store level panel data") font que leur utilisation génère un flot de recherche très important. Il est cependant nécessaire, pour les utiliser correctement, de prendre un certain nombre de précautions.

1 Le choix de l'unité d'observation : Par exemple, Lesieur Arachide 100 cl et Lesieur Arachide 300 cl sont-elles deux unités d'observation différentes ou identiques ? Dans l'affirmative, on sera amené à traiter un très grand nombre de références ce qui peut poser des problèmes de validité des matrices de transitions entre marques (voir Partie II Chapitre 1 section2 à ce sujet). En effet, le nombre de relations de transferts entre marques augmente avec le carré du nombre de marques étudiées.
Le nombre d'actes d'achats permettant d'estimer chaque relation de transferts diminue donc très vite avec l'augmentation du nombre de marques étudiées, remettant alors en cause la fiabilité des estimations.

2 Il est nécessaire de savoir à quel type de comportement de choix on s'intéresse :

- choix d'une marque parmi n
- choix de k (constant) marques parmi n
- choix d'un nombre quelconque de marques parmi n. C'est le cas notamment sur le marché des Soft Drink où le consommateur va acheter tout un ensemble de marques pour satisfaire différents besoins liés à des situations d'usage différentes, les goûts de différentes personnes dans le ménage, la recherche de variété (notamment au niveau des parfums).

3 Il est nécessaire de définir la période d'observation ainsi que *"l'occasion d'achat"*. La période d'achat doit être suffisamment longue pour observer correctement les comportements mais pas trop afin de ne pas prendre en compte la modification éventuelle des structures de préférences. Celle-ci peut être due à un changement du consommateur lui même ou à une modification de l'offre sur le marché (lancement ou suppression de marques). Cette période devrait être normalement choisie en fonction :

- du rythme de consommation moyen dans la catégorie de produits étudiée
- de la fréquence moyenne de réapprovisionnement qui prend en compte le nombre moyen d'achats simultanés dans la même catégorie de produits .
- du type d'indice utilisé, pour un indice qui est estimé au plan individuel il est nécessaire d'avoir une période d'observation plus longue.

4 Il est nécessaire d'avoir défini *"l'occasion d'achat"* : pénétrer dans le magasin, acheter effectivement dans la catégorie de produits... De plus, on sera amené à se poser la question de l'équivalence sur le plan théorique entre trois unités d'une marque achetées en une seule fois et les trois mêmes unités achetées en trois occasions différentes.

5 Il est nécessaire d'avoir choisi un mode d'agrégation des observations et des individus. On peut en effet :

- ne sélectionner que les ménages ayant effectué un nombre d'achats (ou ayant eu un nombre d'occasions d'achat) suffisant.
- pondérer chaque ménage par son nombre d'achats ou d'occasions d'achat.
- sélectionner pour chaque ménage une occasion d'achat au hasard.
- pondérer chaque ménage par son nombre moyen d'achats par occasion.

Nous allons maintenant présenter les principaux travaux qui ont été proposés en matière de mesure de la concurrence utilisant des données de transferts. Notre plan de présentation suivra la distinction (qui nous parait importante) que nous avons faite entre indices déterministes et indices stochastiques ainsi que la distinction entre indices individuels et agrégés.

SS2 INDICES DETERMINISTES

Sous cette appellation nous regroupons un ensemble d'indices qui sont calculés directement à partir des flux de transferts observés entre les marques. A la différence des indices stochastiques, il n'y a pas constitution d'un corps d'hypothèses statistiques concernant le comportement d'achat du consommateur.
Lehmann (1972) est un des premiers auteurs à avoir proposé une mesure de la similarité entre marques basée sur l'observation des données de panels. A l'origine cet indice était présenté comme une alternative aux analyses classiques de MDS sur des données de similarité perçue. Il permettait d'établir un "pont" entre les modèles stochastiques de part de marché et les modèles de choix déterministes. L'auteur proposait les deux indices suivants :

$$Sij = \frac{Pij + Pji}{2} \quad \text{ou encore} \quad Sij = \frac{Nij + Nji}{Ni + Nj}$$

avec

Sij : Similarité entre les marques i et j
Pij : probabilité de changer de la marque i pour la marque j
Pji : probabilité de changer de la marque j pour la marque i
Nij : Nombre de consommateurs ayant effectué au moins un transfert de i vers j sur la période d'observation
Nji : Nombre de consommateurs ayant effectué au moins un transfert de j vers i sur la période d'observation
Ni : Nombre de consommateurs ayant choisi la marque i sur la période
Nj : Nombre de consommateurs ayant choisi la marque j sur la période

Les matrices de *"transition"* entre deux périodes sont traitées ensuite par la technique classique du MDS. En comparant les représentations obtenues avec celles d'une analyse de similarités perçues, l'auteur concluait sur une différence entre les structures de marché obtenues. Ce résultat laissait présager que l'utilisation de données perceptuelles (ici les similarités) ne fournit pas forcément une représentation de la concurrence telle qu'elle se manifeste réellement sur le marché ou du moins, qu'on ne mesure pas la même chose dans les deux cas. Les problèmes associés avec cette modélisation de la concurrence sont liés aux constatations suivantes :

- comme pour tout indice construit à partir de transferts, il est nécessaire d'avoir justifié au préalable une définition pertinente du marché, ce que ne fait pas l'auteur dans son article.
- le modèle pose l'hypothèse que les préférences (manifestées en terme de changements entre marques) sont stables sur la période d'étude. Il faut en effet observer une période suffisamment longue pour calculer de manière fiable les indices de transferts entre marques. Dans le cas où cette stabilité ne serait pas vérifiée, ce que le chercheur ou responsable marketing prend pour une relation de concurrence entre marques pourrait n'être dû qu'à une modification des structures

de préférences. Notons toutefois que pour les produits à fréquence d'achat élevée, ce problème ne doit pas se poser.

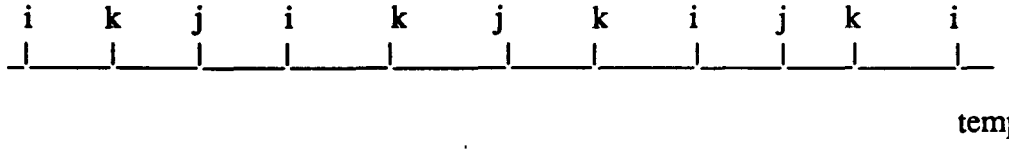
- les individus du panel étudié doivent être homogènes. Dans le cas contraire, on se heurte à un problème d'agrégation de structures de marché différentes. C'est le cas lorsqu'à différents groupes d'individus correspondent différentes structures de concurrence.

-les structures de concurrence caractérisées sont elles stables en fonction des diverses situations d'usage du produit ? Ce point est malheureusement bien difficile à prendre en compte avec ce type de données et il constitue une faiblesse pour la majeure partie des indices construits à partir de ce type de données.

Il est à noter que d'autres indices de concurrence pourraient être construits à partir des données précédentes. En effet, l'indice de Lehmann est un indice d'association particulier parmi la batterie d'indices d'association qui ont été proposés dans la littérature statistique. Il correspond en fait à l'indice de similarité de Russet et Rao (1940, voir Chandon et Pinson 1981). Cet indice s'écrit en effet comme le rapport P/T du nombre de co-présences d'attributs sur le nombre total d'attributs. Dans notre cas T figure le nombre total d'acheteurs des deux marques et P le nombre total d'acheteurs pour qui il y a présence d'un transfert d'une marque vers l'autre. De nombreux autres indices existent et peuvent présenter des propriétés intéressantes pour mesurer la concurrence entre les marques. Les tableaux figurant à l'annexe 2, extraits de Chandon et Pinson (1981 p 67 et 74), présentent un panorama (non exhaustif) de ces indices.

La prise en compte de l'effet de taille (part de marché de la marque), conduisant à des taux de fréquence d'association dus au simple hasard, pourrait par exemple être pris en compte avec un indice du même type que celui développé par Lerman (voir Chandon et Pinson 1981, chapitre 3). D'autre part, il serait intéressant d'utiliser pour ce type d'analyse des indices reflétant la non-symétrie des relations de concurrence entre les marques.

Totten et Elrod (1988) ont proposé l'utilisation d'un indice de concurrence défini à partir de données de transferts, indice qui est peu sensible à une définition inexacte du marché. Les auteurs partent de la constatation (qui nous parait très importante) que peu de méthodologies d'analyse de la structure du marché permettent de définir le marché qui est étudié. Or, cette phase est pourtant très importante. En effet, l'introduction d'une marque n'appartenant pas à l'univers de concurrence peut avoir des effets très importants sur le calcul de l'indice comme le montre l'exemple suivant.



Dans l'historique d'achat précédent, les deux marques i et j sont réellement concurrentes alors que la troisième, k, a été introduite par erreur. L'introduction erronée de k a pour conséquence d'établir un lien de concurrence entre les marques i et k et j et k (à partir des transferts) mais aussi de sous-estimer la relation de concurrence entre les

marques i et j qui sont, elles, réellement concurrentes. On peut ainsi aboutir à la constatation qu'il y a plus de transferts entre deux marques appartenant à deux marchés différents qu'entre deux marques appartenant au même marché. Les mesures de concurrence établies à partir de transferts entre marques sont donc très sensibles à une définition inexacte de l'univers de concurrence.

Elrod et Totten proposent une mesure de concurrence moins sensible à une mauvaise définition de l'univers de concurrence. Cet indice repose sur l'hypothèse suivante : *"Si deux marques sont similaires (concurrentes) elles le sont aussi dans leurs relations (leurs transferts) avec les autres"*. Sur un marché comportant cinq marques (voir tableau suivant), la concurrence entre les marques 1 et 2 peut être établie à partir de la similarité entre les deux profils de transferts de ces deux marques, c'est à dire (n13, n14, n15) et (n23, n24, n25). n13 représente le nombre de ménages chez qui on a observé au moins un achat de la marque 1 suivi d'un achat de la marque 3.

L'indice proposé est une mesure *indirecte* de la similarité entre deux objets. On dira que les marques 1 et 2 sont similaires parce qu'elles ont toutes les deux les mêmes patterns de ressemblance aux autres marques. La mesure de la concurrence entre deux marques ne dépend plus du lien direct, en terme de transferts, entre ces deux marques.

TABLEAU 8 MATRICE DE TRANSFERTS ENTRE CINQ MARQUES

	1	2	3	4	5	
1	n11	n12	n13	n14	n15	SR1
2	n21	n22	n23	n24	n25	SR2
3	n31	n32	n33	n34	n35	SR3
4	n41	n42	n43	n44	n45	SR4
5	n51	n52	n53	n54	n55	SR5

La concurrence entre deux marques i et j est finalement mesurée par la distance du chi2 calculée sur le sous-tableau formé par les lignes i et j du tableau d'échanges, après suppression des colonnes i et j. On obtient alors la sous-matrice des profils de transferts de i et j suivante :

	3	4	5	
1	n13	n14	n15	L1
2	n23	n24	n25	L2
	C1	C2	C3	T

L_i est la somme de la ligne i de cette sous-matrice
C_i est la somme de la colonne j de cette sous-matrice
T est le total des sommes en lignes ou en colonnes

Le terme général du chi2 s'écrit : $(n_{ij} - \sigma_{ij})^2/\sigma_{ij}$ où σ_{ij} est le transfert théorique attendu sous hypothèse d'indépendance entre les deux marques i et j.

$$\sigma_{ij} = L_i \, . \, L_j / T$$

Pour éliminer l'effet de part de marché des marques, les sous-matrices sont standardisées par les parts de marché des marques.

L'avantage de cet indice est donc :

- de ne pas être sensible aux parts de marché des marques étudiées
- d'être plus robuste que d'autres à l'introduction de marques n'appartenant pas à l'univers de concurrence car il repose sur le lien entre deux marques par l'intermédiaire de toutes les autres.

SS3 INDICES STOCHASTIQUES

Par indices de type stochastique, nous faisons référence à des indices qui sont développés à partir d'une modélisation de type stochastique du comportement du consommateur. L'hypothèse de base sur laquelle reposent ces modèles est que la préférence des consommateurs est de nature stochastique et non pas déterministe (voir Massy et al 1970, Bass 1974).

Nous présenterons les indices stochastiques de mesure de la concurrence présents dans la littérature en fonction du type de modèle de choix qu'ils utilisent. Nous avons vu au premier chapitre (section 3) qu'il existe plusieurs types de modèles stochastiques de choix. Principalement, on rencontre les modèles d'ordre zéro (modèles de Bernoulli) ou d'ordre un (modèles de Markov).

SS3-1 INDICES CONSTRUITS A PARTIR DE MODELES DE CHOIX D'ORDRE ZERO

SS3-1-1 LE MODELE DE HENDRY

Le modèle de Hendry (voir Kalwani et Morrison 1977, Rubinson, Vanhonacker et Bass 1980 pour une présentation détaillée) est une méthode d'analyse de la structure du marché de type *"confirmatoire"*. Comme tel, il développe deux types d'indices :

- des indices de mesure de la concurrence calculés à un niveau théorique, c'est à dire sous l'hypothèse d'un ensemble d'hypothèses régissant le comportement de choix du consommateur, pour une structure particulière du marché.
- des indices de mesure de la concurrence calculés à partir de données empiriques.

La confrontation entre ces deux types d'indices permet d'évaluer la qualité des hypothèses faites quant à la structure du marché.

D'autre part, le modèle de Hendry pose l'hypothèse que le marché se structure sous la forme de partitions (exclusives), ce qui permet de calculer deux types de mesure de la concurrence :

- à l'intérieur d'une partition (c'est à dire un sous-marché) : indice intra-partition
- entre différentes partitions : indices inter-partitions

METHODOLOGIE DE CALCUL

Le modèle repose sur deux hypothèses :

h1 : Les achats des consommateurs suivent un processus d'ordre 0, ce qui signifie qu'il n'y a pas d'effet d'apprentissage au cours du temps. A l'équilibre, les consommateurs ont des vecteurs de préférences stables et donc des probabilités d'achat des différentes marques qui sont constantes.

h2 : Les probabilités de transferts entre marques sont proportionnelles aux parts de marché des marques correspondantes (hypothèse de proportionnalité).

SS3-1-1-1 Constante de Changement

Soit :

N individus

g marques appartenant à un marché (la première étape consiste donc à fixer l'univers concurrentiel) i=1 à g

Un consommateur choisi au hasard choisit entre g marques avec un vecteur de préférences (probabilités) : $[\pi_1, ..., \pi_g]$.

π_g est la probabilité pour que cet individu choisisse la marque g .

Sur la population totale, $E(\pi_i)=p_i$, p_i est la part de marché de la marque i.
D'autre part, à l'équilibre (Stationnarité des préférences dans un modèle d'ordre 0) :

$$E(\pi_i\pi_j) = E(\pi_j\pi_i)$$

Les transferts entre deux marques sont symétriques .

L'essence du modèle de Hendry est de dire que la probabilité de transfert entre deux marques au sein d'une même partition est proportionnelle à leurs parts de marché respectives :

$$E(\pi_i\pi_j) = K.p_i.p_j \quad i \text{ différent de } j \quad (1)$$

K est appelée constante de changement sur le marché, elle est indépendante des marques i et j; c'est une caractéristique de ce marché .

(1) est vraie si la distribution multivariée de $[\pi 1, - - \pi g]$ est Dirichlet (voir Bass, Jeuland et Wright 1976) .

De plus $K \in [0,1]$

$$Var(\pi_i) = (1-K).p_i.(1-p_i)$$

Bass (1974) a montré que (1-K) est la corrélation entre les achats successifs, c'est à dire "le facteur de fidélité à la classe de produit étudiée". Cet indice n'est pas spécifique à une marque particulière.

Si la relation (1) est vraie, on peut écrire que la fraction de la population changeant de la marque i pour une autre marque du marché est :

$$TS_i = Kw.p_i. \sum_{j \neq i} p_j$$

$$\text{comme } \sum_{j=1}^{g} p_j = 1$$

$$TS_i = Kw.p_i.(1-p_i)$$

On écrira de la même manière, en sommant sur i, le changement total empirique entre toutes les marques du marché, TS_{emp} :

$$TS_{emp} = Kw \sum_i p_i.(1-p_i)$$

Kw peut alors s'écrire :

$$Kw_{emp} = TS / \sum_i p_i.(1-p_i) \qquad (2)$$

(2) peut être estimée à partir du changement total empirique observé (TSemp) et des parts de marché des marques.

Kalwani et Morrison (1977) ont montré que, sous hypothèse d'un processus stochastique d'ordre 0, le nombre de consommateurs changeant d'une marque i pour une autre marque est égal au nombre d'individus changeant d'une marque quelconque pour la marque i.

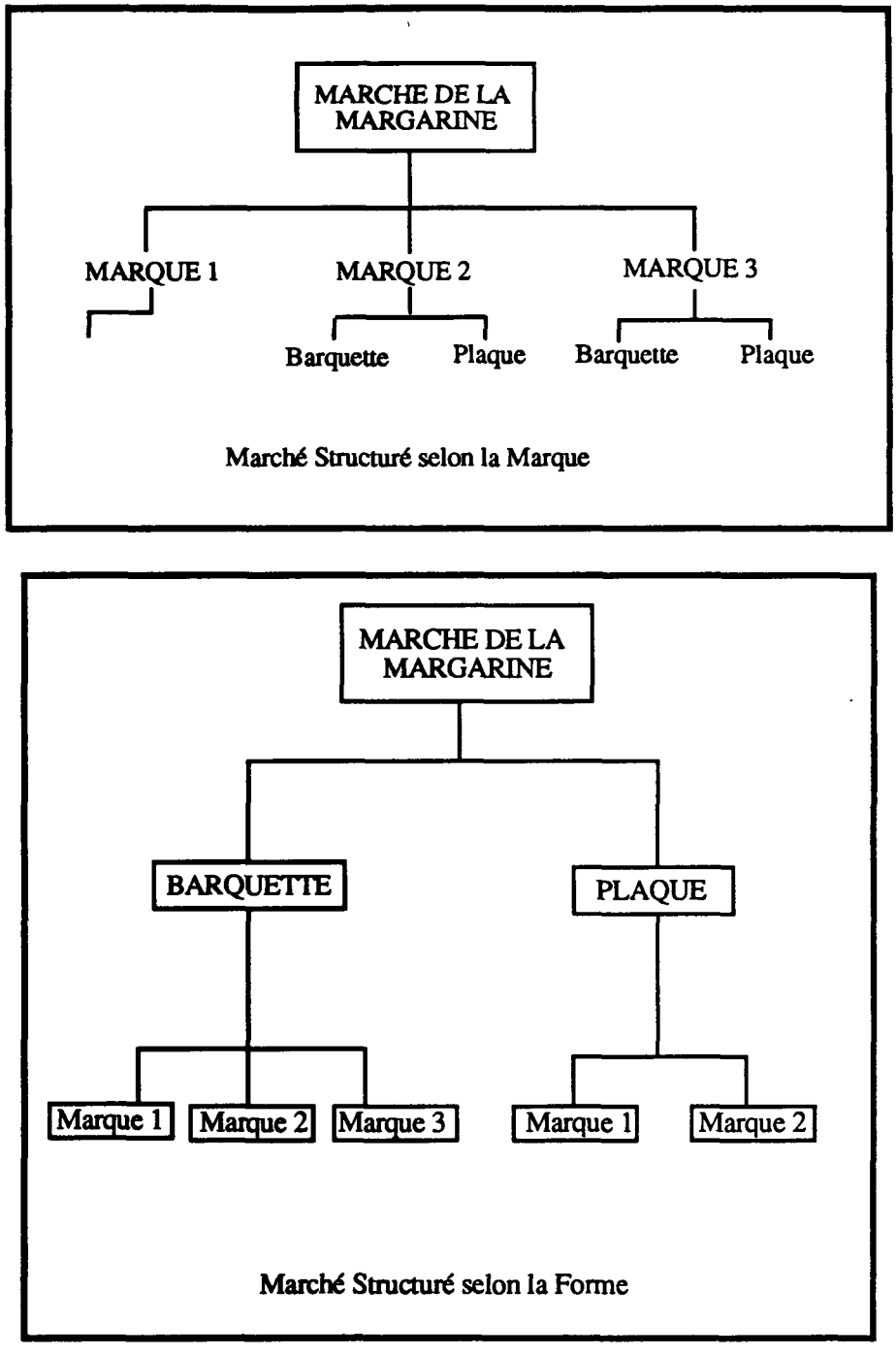
SS3-1-1-2 Structure Théorique

Un ensemble de structures hypothétiques sont identifiées à partir des attributs de la classe de produits étudiée. Trois types de partitions sont à considérer (voir Graphique 4)

A Partitions basées sur la marque :

Dans ce cas, le changement entre les marques se calcule sans tenir compte de la "forme" des marques (c'est à dire leurs attributs physiques). Il y a une seule constante de changement pour toutes les marques, le changement entre deux marques est proportionnel aux parts de marché par rapport à l'ensemble du marché.

FIGURE 14 TYPE DE STRUCTURES POSSIBLES SUR UN MARCHE



"Adapté de Kalwani et Morrison 1977"

B Partitions basées selon la forme :

Les transferts entre marques appartenant à la même *"partition en terme de forme"* (attribut physique du produit, par exemple boisson gazeuse) sont plus forts, relativement aux parts de marché, que les transferts entre marques appartenant à des partitions différentes. Kw peut alors varier selon que le transfert s'effectue à l'intérieur ou entre deux partitions.

C Partitions mixtes :

Elles reprennent les deux critères précédents. Dans une partition de type "*marque-forme*", le consommateur choisit d'abord sur la base de la marque et ensuite sur la base d'une forme particulière de la marque. On peut, à l'inverse, rencontrer des marchés où le consommateur choisit d'abord en terme de forme (boisson gazeuse) puis ensuite en terme de marque (Orangina).

SS3-1-1-3 Indices Théoriques dans le Modèle de Hendry

Le modèle de Hendry fournit la valeur théorique de TS, le changement total au sein d'un marché (voir le détail des calculs en annexe 3).

$$TS_{theo} = \sum_{i=1}^{g} \frac{-p_i.\ln(p_i)}{1 - p_i.\ln(p_i)} \quad (3)$$

Cette valeur est obtenue (voir Bass et al 1980) en exprimant le comportement de N consommateurs aléatoires sans préférences particulières concernant g marques.

A - Si le marché est structuré "*marque*", (3) donne la valeur unique du changement total.

B - Si le marché est structuré "*forme*" :

* Le changement total entre formes est donné par (3') :

$$TS = \sum_{k} \frac{-p_k.\ln(p_k)}{1 - p_k.\ln(p_k)} \quad (3')$$

où pk est la part de marché de la partition k dans le marché total.

*Le changement total entre marques appartenant à la même "*forme*" k est donné par (3") :

$$TS_k = \sum_{i \in k} \frac{-p_i.\ln(p_i)}{1 - p_i.\ln(p_i)} \quad (3")$$

où p_i est la part de marché de la marque i au sein de sa partition k

C - ,Si la structure est de type "*mixte*" :

* Le changement entre "*formes*" est donné par la valeur de TS_{theo} dans le cas d'un marché structuré "*marque*", c'est à dire par (3)

*Le changement au sein d'une "*forme*" est donné par la valeur théorique de TS pour le cas d'un marché "*forme*", c'est à dire (3").

SS3-1-1-4 Indices de Concurrence Développés dans le Modèle Hendry

Les indices de concurrence développés dans le modèle Hendry se situent à un double niveau :

1 Au niveau théorique, avec des indices de mesure de la concurrence tels qu'ils devraient être observé si les hypothèses formulées sur la partition étaient justes. Ces indices ont une nature théorique.

2 Au niveau empirique, avec des mesures de la concurrence réellement observée.

La particularité de ce modèle est de distinguer deux types de relations de concurrence:

- la concurrence à l'intérieur d'une partition (c'est à dire un sous-marché). Cette concurrence, pour une partition particulière, est définie par un seul indice appelé *"indice d'entropie"* (ou constante de changement intra-partition). Cette valeur est unique pour tous les produits de la partition, ce qui signifie que le niveau de la concurrence entre les marques est unique au sein d'une même partition.
- la concurrence entre partitions (inter-partitions). Le niveau de concurrence entre toutes les marques d'une partition et toutes les marques d'une autre partition est unique et est donné par la constante de changement entre les deux partitions considérées.

Cette approche de la mesure de la concurrence sur un marché, mérite les remarques suivantes :

1 La méthode procède au plan agrégé, ce qui permet d'avoir des calculs assez simples. De plus les seuls intrants nécessaires pour faire fonctionner le modèle sont les parts de marché et les transferts entre marques. A input identique, cette approche est beaucoup plus simple à mettre en oeuvre que le modèle Defender étudié précédemment.

2 L'hypothèse que le comportement de choix des consommateurs suit un processus d'ordre zéro est bien sûr critiquable. Cependant, la robustesse des modèles d'ordre zéro a été démontrée dans beaucoup de recherches (voir section 3 du premier chapitre).

3 Les hypothèses à formuler concernant la nature des partitions peuvent être complexes et nombreuses quand il y a beaucoup d'attributs.

4 Pour former des partitions il est nécessaire d'avoir des attributs de type discret ce qui, sur certains marchés, est peu réaliste.

5 Les concepts d'entropie et de constante de changement sont intéressants car ils permettent de caractériser un marché ainsi que la "*force*" de sa structure (en comparant les niveaux de KS inter et intra-partitions). KS est un concept intéressant car c'est un indice comparable d'un marché à l'autre. Ses valeurs sont en effet standardisées et varient entre 0 et 1.

6 Une incertitude notable concernant ce modèle provient, comme l'ont montré Urban et Hauser (1980 chapitre V), du fait que l'hypothèse de proportionnalité des changements au

sein d'une partition peut dans certains cas être encore vraie pour des marques appartenant à différentes partitions. Ceci peut conduire alors à l'obtention de différentes structures .

SS3-1-5 Développements récents du modèle de Hendry

Kumar et Sashi (1989) ont présenté un modèle d'analyse de la structure des marchés qui reprend en grande partie les principes du modèle de Hendry. La concurrence entre deux marques y est mesurée à partir de la probabilité conditionnelle (Pi/j) de changement entre ces deux marques. L'expression de cette probabilité conditionnelle est différente selon que les deux marques sont supposées appartenir ou non à la même partition. Pour deux marques au sein de la même partition on a :

$$\pi_{j/i} = Ki \,.\pi_j$$

K_i est la constante de changement au sein de la partition
π_j est la probabilité marginale de choix de la marque j (sa part de marché).

Pour deux marques n'appartenant pas à la même partition, la probabilité conditionnelle de choix s'écrit :

$$\pi_{j/i} = \mu_{kl}.\partial_i.\pi_j$$

μ_{kl} est une constante positive plus petite que 1 qualifiant l'échange entre les partitions l et k
∂_i est une constante, dont la valeur est supérieure à K_i

Après avoir posé des hypothèses concernant la structure du marché, en termes d'attributs physiques du produit, les paramètres sont estimés dans un modèle log-linéaire. Le critère d'information d'Akaike permet de mesurer l'adéquation entre les transferts observés et prédits. La valeur de ce critère permet de tester quelle est la structure hypothétique qui est la mieux confirmée par les données, ce que ne fait pas le modèle de Hendry.

Il est a noter que les structures étudiées sont comparées à "l'*absence de structure*", à l'aide du critère d'Akaike. On cherche à voir si une structure particulière permet de mieux expliquer les transferts observés que l'absence de structure.

Comme pour le modèle Hendry, puisque les structures testées sont construites à partir d'attributs "***objectifs et discrets***" (par exemple café moulu ou en grain), il n'est pas nécessaire de poser d'hypothèses concernant l'homogénéité des individus en terme de transferts.

Notons pour finir, que d'autres auteurs ont proposé l'utilisation des probabilités conditionnelles de transferts entre deux marques i et j, ($P_{i/j}$), comme une mesure de la concurrence entre deux marque. Grover et Dillon (1988) ont ainsi proposé d'introduire ces probabilités conditionnelles dans une méthode de décomposition spectrale (méthode proche de l'analyse factorielle des correspondance) afin d'analyser la vitesse de retour à l'équilibre du marché (par exemple après avoir lancé une nouvelle marque).
Le concept de probabilité conditionnelle de choix est en effet tout a fait proche du concept "concurrence".

SS3-1-2 LE MODELE DE RAO ET SABAVALA (1981)

Rao et Sabavala (1981) ont présenté une modélisation du comportement de choix du consommateur, qualifiée par les auteurs de *"hiérarchique et stochastique"*. Cette modélisation permet d'estimer des indices de concurrence entre marques.
Les hypothèses suivantes permettent de modéliser le processus de choix du consommateur :

H1 A chaque occasion d'achat, le consommateur partitionne l'ensemble des marques présentes sur la base de leurs attributs (objectifs ou subjectifs), jusqu'à ce qu'un item soit sélectionné dans la partition restante.

H2 Le modèle de partition est le même pour tous les individus appartenant au même segment étudié (la même séquence d'attributs est considérée), seuls les niveaux des marques selon les attributs peuvent varier. Il n'y a donc qu'un seul arbre de décision pour chaque segment.

H3 Le choix comporte une composante stochastique: l'individu peut suivre différentes hiérarchies en différentes occasions d'achat, même si ses préférences sont stables dans le temps. Le degré de changement entre deux marques sur des occasions consécutives est proportionnel aux parts de marché de ces marques. La constante de proportionnalité du changement (voir modèle de Hendry plus haut) est un paramètre de la structure, il varie d'un niveau de la structure à l'autre.

H4 Le niveau dans la hiérarchie représente le degré de concurrence entre partitions, les changements entre partitions sont plus élevés en haut de la hiérarchie qu'en bas .

On peut constater que les partitions des marques par le consommateur sont atteintes en suivant une logique identique à celle du modèle d'élimination par les aspects (EBA) de Tversky (1972). La structure obtenue est une *"partition"* (au sens mathématique du terme), il n'y a pas de recouvrement entre les partitions.
Les auteurs proposent d'utiliser des données de comportement de choix pour alimenter ce modèle. L'observation d'historiques d'achat doit permettre *"d'inférer des interactions dans le comportement"*.
Pour chaque segment homogène et pour T occasions d'achats on va construire T-1 matrices de transitions N_t, t=1...T-1. Ces matrices Nt_t ont pour terme général n_{ij}, où n_{ij} est le nombre de consommateurs qui ont acheté i en t1 et j en t2. Afin d'avoir une estimation fiable, les auteurs proposent de calculer la moyenne de ces T-1 matrices de transitions à deux périodes . On obtient alors la matrice agrégée N des transferts de marques entre deux périodes.

L'index de mesure de la concurrence proposé par Rao et Sabavala s'écrit alors :

$$Fij = \frac{n_{ij} \, . \, n..}{n_{i.} \, . \, n.j}$$

n.. est le nombre de consommateurs appartenant au segment homogène, c'est à dire les consommateurs ayant des arbres hiérarchiques de choix identiques. Ces consommateurs ont le même ordre de considération des attributs

physiques ou perçus qui leur permet de fractionner successivement le marché de manière à aboutir à leur choix, même si ce choix final diffère d'un consommateur à l'autre.

$n_{i.}$ est le nombre de consommateurs, dans le segment homogène, qui ont acheté la marque i à la première occasion d'achat.

$n_{.j}$ est le nombre de consommateurs, dans le segment homogène qui ont acheté la marque j à la deuxième occasion d'achat

n_{ij} est le nombre de consommateurs, dans le segment homogène, qui ont acheté la marque i à la première occasion d'achat puis la marque j à la seconde occasion d'achat.

L'index de concurrence proposé par les auteurs est finalement le ratio du nombre empirique de consommateurs ayant changé de i pour j par le nombre attendu de consommateurs changeant de i pour j sous hypothèse d'indépendance entre les deux occasions de choix (les probabilités marginales de choix des marques étant données par les parts de marché relatives de ces marques).

Si $F_{ij} > 1$ cela signifie qu'il y a plus de transferts entre i et j que ne le laisse espérer la loi de proportionnalité, les deux marques sont en situation de sur-concurrence.
Si $F_{ij} < 1$, on observe moins de transferts entre i et j que ne le laisse espérer la loi de proportionnalité, les deux marques sont en situation de sous-concurrence.
F_{ij} est appelé facteur de proportionnalité dans la relation de changement. F_{ij}, tel qu'il est proposé par les auteurs n'est pas symétrique, même lorsque les parts de marché sont à l'équilibre (c'est à dire quand $E(n_{ij})=E(n_{ij})$ pour tout i et j). Selon les auteurs, cette non symétrie serait due à la nature "stochastique" du processus. Cette explication ne nous parait pas suffisante. Elle réduit la richesse du concept de concurrence qui selon nous n'est pas symétrique.

Une des hypothèses du modèle, bien que valable au plan théorique, nous parait difficilement tenable dans la réalité, c'est l'hypothèse de calcul des indices sur des échantillons homogènes de consommateurs, homogènes en terme de processus de choix. Comment savoir à l'avance que les individus d'un segment ont la même hiérarchie d'attributs ?

Dans leur méthodologie, les auteurs constituent ces segments sur la base des taux d'achat et de fidélité aux marques. Comment être sûr dans ces conditions que les individus ainsi regroupés ont la même hiérarchie de considération des attributs ? Autrement dit, s'il est tenable qu'au plan individuel les matrices de changements sont bien le reflet d'un processus de choix particulier, que peut-on dire d'un processus de choix caractérisé à partir d'une matrice de transferts agrégée sur des individus dont on ne sait pas à l'avance contrôler l'hétérogénéité? Enfin, il existe une incohérence entre la notion d'arbre hiérarchique de choix et la hiérarchie mesurée à partir de cette approche :

- d'une part, dans un arbre hiérarchique de choix, c'est la réalité "consommateur" qui est étudiée, alors qu'à partir des données de transferts agrégées, on observe une réalité de "marché" qui est largement influencée par les variables de situation d'achat (par exemple la marque i est achetée parce qu'elle est en promotion).

- d'autre part et cela est encore plus grave, la notion de hiérarchie des attributs considérés lors d'une situation de choix est un processus de type "*instantané*". La hiérarchisation du choix s'effectue en respectant un arbre de décision. Dans le modèle présenté ici, le choix hiérarchique s'effectue à différentes périodes, en respectant un arbre de décision défini à-priori. Quelle est la signification conceptuelle d'un tel modèle ? Comment justifier que c'est le même processus qui est caractérisé sur un historique d'achat observé durant toute une année que celui qui prévaut au niveau d'un choix ponctuel, à un instant donné ?

SS3-1-3 SEGMENTATION DES CONSOMMATEURS ET ANALYSE DE LA STRUCTURE DES MARCHES

LE MODELE DE DROVER ET DILLON

Grover et Dillon (1985) ont proposé une approche de l'analyse de la structure de la concurrence entre les marques qui procède de manière assez analogue à celle proposée par Rao et Sabavala .

Cette approche possède deux particularités notables. Tout d'abord il s'agit d'une démarche confirmatoire, au même titre que le modèle de Hendry. Ensuite, elle permet à la fois de structurer le marché et de segmenter les consommateurs. Nous avons noté au chapitre I, que l'analyse de la structure des marchés se heurte à un problème méthodologique important lorsque le marché étudié présente une structure "*composite*". C'est le cas où coexistent plusieurs segments de consommateurs possédant chacun leur propre structure de marché. Dans ces conditions, une analyse agrégée est forcément biaisée et d'autre part, on ne peut pas segmenter les consommateurs sur la base de leur structure de marché puisque c'est exactement ce qui est recherché. La seule solution envisageable est d'effectuer simultanément les deux opérations .

C'est ce que proposent Grover et Dillon. Le modèle repose sur l'hypothèse suivant laquelle les changements observés entre marques sont le reflet de l'appartenance des individus à des "*classes latentes*". Chacune de ces classes est caractérisée par une hiérarchie particulière de critères de choix et donc par un enchaînement particulier de transferts entre les marques. Le comportement de changements entre marques observé au cours de plusieurs périodes serait donc dicté par l'appartenance de l'individu à une classe latente particulière compatible avec un processus hypothétique de choix particulier.

DESCRIPTION DU MODELE

Soit

N_1 le nombre de marques entre lesquelles l'individu peut choisir lors de sa première occasion d'achat.

N_2 le nombre de marques entre lesquelles l'individu peut choisir lors de sa deuxième occasion d'achat.

"V_1" le vecteur à N_1 composantes de ses achats et non-achats à la première occasion d'achat (1 si achat de la marque 0 si non achat).

"V_2 " le vecteur à N_2 composantes de ses achats et non-achats à la deuxième occasion d'achat (1 si achat de la marque 0 si non achat).

M classes latentes caractérisant un processus hiérarchique de choix particulier, m = 1...M.

Le vecteur des achats du consommateur sur deux occasions peut s'écrire sous la forme du vecteur: $O_{1,2}$, vecteur à (N1+N2) composantes. Il y a donc $2^{(N1+N2)}$ vecteurs $O_{1,2}$ différents

$Em_{(1,2)}$ le vecteur de réponse à (N1+N2) composantes, qui reflète le comportement d'achat attendu (espéré) pour deux occasions d'achats, pour un consommateur appartenant à la classe latente m. Cette classe latente m est caractérisée par un enchaînement de changements entre marques bien précis.

Les auteurs expriment alors $P(O_{1,2})$, la probabilité d'observer le vecteur d'achats $O_{1,2}$ pour un consommateur (ou ménage) de la manière suivante:.

$$P(O_{1,2}) = \sum_{m=1}^{M} P(O_{1,2} / E^{m}_{1,2}) \cdot P(m) \quad (1)$$

P(m) est la probabilité qu'un ménage utilise la stratégie de choix associée à la classe latente m. P(m) est en fait la proportion de ménages appartenant à la classe latente m.

$$\sum_{m=1}^{M} P(m) = 1 \text{ pour chaque ménage}$$

$P(O_{1,2} / E^{m}_{1,2})$ est la probabilité conditionnelle d'observer le vecteur de comportement $O_{1,2}$ sachant que le ménage appartient à la classe latente m.

$$P(O_{1,2} / Em_{1,2}) = \prod_{t=1}^{2} \prod_{i=1}^{N_t} P(it/m)^{\alpha_{it}} \cdot (1-P(it/m))^{1-\alpha_{it}} \quad (2)$$

P(it/m) est la probabilité conditionnelle pour que la ième marque soit achetée au t ième achat (t=1 ou 2) sachant que le ménage appartient à la classe latente m.

α_{it}= 1 si la marque i est achetée en t

α_{it}= 0 dans les autres cas

Il est à noter que les P(it/m) sont des probabilités conditionnées à l'appartenance du consommateur à une classe latente particulière et non des probabilités de transfert d'une marque vers une autre marque . Ce sont ces probabilités conditionnelles qui donnent au modèle son caractère confirmatoire.

Les probabilités conditionnelles de transferts entre deux marques peuvent être établies à partir de la relation (2). En effet :

$$P(m/it) = [P(it/m)/P(i_t)].P(m)$$

$P(i_t)$ est la probabilité marginale d'acheter la marque i à la t ième occasion d'achat
$P(m/i_t)$ est la probabilité conditionnelle d'appartenir à la classe latente m quand on a acheté la marque i en t
t_{i1i2}, la probabilité de transition de la marque i en l'instant 1 vers la marque j en l'instant 2 s'écrit :

$$t_{i1i2} = \sum_{m=1}^{M} P(m/i_1) \, . \, P(i_2/m)$$

t_{i1i2} est donc la mesure, toutes classes latentes confondues, de la concurrence entre les deux marques i et j.

Si on note n_{12} le nombre de transferts observés entre les marques i et j pour les deux périodes et n le nombre total d'achats on peut noter que :

$P(O_{1,2})$ est un estimateur (estimé par le maximum de vraisemblance) non contraint de n_{12}/n.
$PT(O_{1,2})$ est un estimateur de $P(O_{1,2})$ sous hypothèse d'un modèle particulier de choix. Cette valeur est estimée en utilisant l'analyse en classe latente (méthode MLLSA, Maximum Likelihood Latent Structure Analysis, voir Clogg 1979).

Chaque classe latente reflète une stratégie d'évaluation utilisée et caractérise les enchaînements de transferts observés. Si un enchaînement de choix est parfaitement suivi par l'individu, alors il y a égalité entre $O_{1,2}$ et $Em(1,2)$. Cependant, en général, $O_{1,2}$ sera différent de $Em(1,2)$ par le simple effet des variables de situation d'achat à cause desquelles l'individu n'adopte pas la hiérarchie de choix qui est la sienne.

CALCUL DES INDICES

Il s'effectue en quatre phases.

1 Sélectionner dans le panel les individus ayant effectué au moins k achats sur la période et la catégorie de produits étudiées (les auteurs préconisent k=5).

2 Construire les tables de transitions pour les k-1 transferts entre marques, ces tables ont pour terme général n_{12}, le nombre de ménages effectuant un transfert de 1 vers 2. Ces tables sont réunies pour former une matrice I.I.K (I étant le nombre de marques et K nombre d'occasions d'achats).

TABLEAU 9A EXEMPLE DE MATRICE I.I.K

Cross-Classification Tables for the Hypothesized Tree Structure of Figure 4, Panel A

Table	k	Past Purchase/Switches	Next Purchase T1	T2	T3	T4	T5	
(i)	1	T1	0	141	392	36	55	626
		T2	112	0	15	31	4	162
		T3	276	22	0	14	9	321
		T4	18	28	10	0	1	57
		T5	23	4	3	0	0	30
			429	195	420	83	69	1196
(ii)	2	(T3,T1)(T5,T1)	0	13	254	8	26	299
		(T4,T2)	9	0	1	18	0	28
		(T5,T3)(T1,T3)	371	12	0	8	4	395
		(T2,T4)	5	21	4	0	1	31
		(T3,T5)(T1,T5)	51	4	7	2	0	64
			436	50	266	34	31	870
(iii)	3	(T1,T3,T1)(T1,T5,T1)	0	29	333	9	43	414
		(T2,T4,T2)	5	0	1	15	0	21
		(T3,T5,T3)(T3,T1,T3)	236	6	0	7	5	254
		(T4,T2,T4)	4	14	0	0	0	18
		(T5,T3,T5)(T5,T1,T5)	19	0	1	1	0	21
			264	49	335	32	48	728
(iv)	4	(T3,T1,T3,T1)(T5,T1,T5,T1)	0	4	229	1	21	255
		(T4,T2,T4,T2)	3	0	2	9	0	14
		(T5,T3,T5,T3)(T1,T3,T1,T3)	316	2	0	2	3	323
		(T2,T4,T2,T4)	0	15	0	0	0	15
		(T3,T5,T3,T5)(T1,T5,T1,T5)	27	0	1	0	0	28
			346	21	232	12	24	635
(v)	5	(T1,T3,T1,T3,T1)(T1,T5,T1,T5,T1)	0	14	291	0	31	342
		(T2,T4,T2,T4,T2)	7	0	0	8	0	15
		(T3,T5,T3,T5,T3)(T3,T1,T3,T1,T3)	244	1	0	3	0	228
		(T4,T2,T4,T2,T4)	0	9	0	0	0	9
		(T5,T3,T5,T3,T5)(T5,T1,T5,T1,T5)	19	0	0	0	0	19
			250	24	291	17	31	613

Conditional Probability Specifications for the Hypothesized Tree Structure of Figure 4, Panel A

Latent Class	Switch Occasion	Conditional Probabilities: First Purchase T1	T2	T3	T4	T5	Second Purchase T1	T2	T3	T4	T5
1	1	0	0	F	0	F	1	0	0	0	0
2		0	0	0	1	0	0	1	0	0	0
3		F	0	0	0	F	0	0	1	0	0
4		0	1	0	0	0	0	0	0	1	0
5		F	0	F	0	0	0	0	0	0	1
6		F	0	F	0	F	0	F	0	F	0
7		0	F	0	F	0	F	0	F	0	F
8		F	F	F	F	F	F	F	F	F	F
1	2	0	0	F	0	F	1	0	0	0	0
2		0	0	0	1	0	0	1	0	0	0
3		F	0	0	0	F	0	0	1	0	0
4		0	1	0	0	0	0	0	0	1	0
5		F	0	F	0	0	0	0	0	0	1
6		F	0	F	0	F	0	F	0	F	0
7		0	F	0	F	0	F	0	F	0	F
8		F	F	F	F	F	F	F	F	F	F
1	3	0	0	F	0	F	1	0	0	0	0
2		0	0	0	1	0	0	1	0	0	0
3		F	0	0	0	F	0	0	1	0	0
4		0	1	0	0	0	0	0	0	1	0
5		F	0	F	0	0	0	0	0	0	1
6		F	0	F	0	F	0	F	0	F	0
7		0	F	0	F	0	F	0	F	0	F
8		F	F	F	F	F	F	F	F	F	F
1	4	0	0	F	0	F	1	0	0	0	0
2		0	0	0	1	0	0	1	0	0	0
3		F	0	0	0	F	0	0	1	0	0
4		0	1	0	0	0	0	0	0	1	0
5		F	0	F	0	0	0	0	0	0	1
6		F	0	F	0	F	0	F	0	F	0
7		0	F	0	F	0	F	0	F	0	F
8		F	F	F	F	F	F	F	F	F	F
1	5	0	0	F	0	F	1	0	0	0	0
2		0	0	0	1	0	0	1	0	0	0
3		F	0	0	0	F	0	0	1	0	0
4		0	1	0	0	0	0	0	0	1	0
5		F	0	F	0	0	0	0	0	0	1
6		F	0	F	0	F	0	F	0	F	0
7		0	F	0	F	0	F	0	F	0	F
8		F	F	F	F	F	F	F	F	F	F

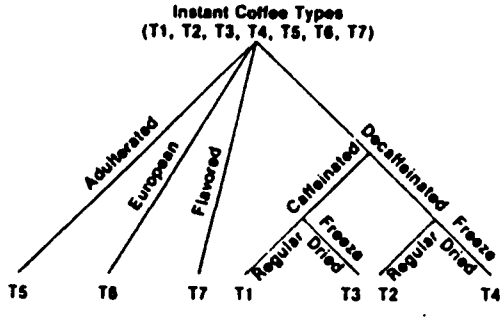
Note: A "0" or "1" designates that the probability estimate is restricted to be zero or one, respectively; "F" designates a free parameter to be estimated.

"Extrait de Grover et Dillon 1985 p328"

3 Convertir l'arbre hypothétique de choix en une forme cohérente avec le modèle probabiliste des équations 1 et 2 puis caractériser les classes latentes compatibles avec tous les enchaînements possibles sur cet arbre pour les K périodes.

4 Estimer les probabilités de transferts par la méthode du maximum de vraisemblance.

TABLEAU 9B EXEMPLE D'ARBRE HIERARCHIQUE ET CHOIX HYPOTHETIQUES CORRESPONDANT



Panel A. Hypothesized Structure

Character of the Latent Aspect Classes Associated with the Tree Structure Shown in Figure 5, Panel A.

Latent Aspect Classes	Permissible Switching Patterns From	To	Level in Hierarchy
1,11,21,31,41	T3,T5,T6,T7	T1	Low
2,12,22,32,42	T4,T5,T6,T7	T2	Low
3,13,23,33,43	T1,T5,T6,T7	T3	Low
4,14,24,34,44	T2,T5,T6,T7	T4	Low
5,15,25,35,45	T1,T2,T3,T4,T6,T7	T5	Low
6,16,26,36,46	T1,T2,T3,T4,T5,T7	T6	Low
7,17,27,37,47	T1,T2,T3,T4,T5,T6	T7	Low
8,18,28,38,48	T1,T3	T2,T4	Higher
9,19,29,39,49	T2,T4	T1,T3	Higher
10,20,30,40,50	Any Type	Any Type	Highest

"Extrait de Grover et Dillon 1985 p329 et 330"

Cette approche mérite les commentaires suivants :

*Au plan méthodologique :

Est-il valide de ne considérer que les individus ayant effectué au moins k achats ? Sur de nombreux marchés, un pourcentage important des quantités achetées (70% par exemple) est le fait d'un faible pourcentage des consommateurs (30%). A-t-on alors une structure représentative à partir de si peu d'acheteurs ?
Les auteurs, à ce niveau, proposent de faire des tests de Chi2 entre les matrices de transferts d'individus ayant effectué plus de k ou moins de k achats.

*Au plan conceptuel :

Nous pouvons renouveler les remarques faites concernant l'approche de Rao et Sabavala. Les auteurs mettent sur un même plan un processus de traitement de l'information de type instantané et un historique d'achat inter-temporel.

LE MODELE DE GROVER ET SRINIVASAN

Dans le même objectif d'obtenir simultanément la segmentation des consommateurs et la segmentation des produits, Grover et Srinivasan (1987) ont proposé un modèle stochastique d'ordre zéro dans lequel les consommateurs sont regroupés sur la base de leurs préférences révélées (leurs fréquences d'achats des marques).
Au sein de chaque partition de marques, la probabilité de transfert d'une marque vers une autre est proportionnelle aux parts de marché de ces deux marques. Le modèle est estimé au plan agrégé en écrivant :

$$S_{ij} = \sum_{h=1}^{H} \beta_h . r_{ih} . c_{jh}$$

h est l'indice de partition hypothétique à laquelle appartiennent les marques i et j
β_h est le poids de la partition h
r_{ih} est la probabilité pour que la marque i dont on change appartienne à la partition h
c_{jh} est la probabilité pour que la marque j pour laquelle on change appartienne à la partition h
S_{ij} est la proportion de consommateurs ayant changé de i pour j sur deux occasions d'achats consécutives de la période d'observation.

Les paramètres sont estimés en utilisant l'analyse en classes latentes. Il est nécessaire de fixer a priori le nombre de segments. Pour tester le nombre de segments significatifs, on peut étudier la progression de la qualité de l'estimation, au fur et à mesure que le nombre de segments pris en compte augmente, à l'aide du test de chi2. On obtient finalement une probabilité d'appartenance de chaque individu à chaque segment de consommateurs et une probabilité d'appartenance de chaque marque à chaque partition. La structure du marché obtenue est donc floue. Nous retiendrons que c'est l'hétérogénéité des préférences des consommateurs qui est la cause de l'obtention de partitions floues.

Jain, Bass et Chen (1990) ont repris ce modèle en tenant compte explicitement de l'hétérogénéité des consommateurs. En effet, au sein de chaque segment, les consommateurs sont encore hétérogènes. Cette hétérogénéité est modélisée en utilisant la distribution multinomiale Dirichlet qui décrit la distribution des vecteurs de préférences P des consommateurs pour les marques, au sein d'un segment de consommateurs (il s'agit de préférences révélées, c'est à dire de probabilités d'achats des marques). Le vecteur de préférences $P = P(p_1,p_2,..p_n)$, observé pour un ménage particulier, est la réalisation de la loi Dirichlet de paramètres $(\alpha_1,..., \alpha_n)$. Ces préférences sont considérées comme stationnaires sur la période d'étude. Il est intéressant de noter que cette modélisation enrichit le modèle en permettant d'estimer la variabilité au sein de chaque segment et illustre nos opinions du chapitre premier concernant l'interpénétration croissante des modèles de choix. On a là un bon exemple dans lequel les lois utilisées dans les modèles stochastiques de part de marché (la loi Dirichlet, voir Jeuland et al 1980) sont utilisées dans un modèle dont l'objectif n'est pas de prévoir les parts de marché mais d'analyser la structure concurrentielle des marques.

SS3-1-4 INDICE DE CONCURRENCE EXPLIQUE PAR LES VARIABLES DE CONTROLE DU MIX

Carpenter et Lehmann (1985) ont proposé une modélisation de la concurrence très intéressante parce que novatrice. Selon les auteurs, le flux (transferts) entre deux marques au cours d'une période exprime la relation de concurrence entre ces deux marques. L'originalité des auteurs réside dans leur approche *"explicative"* de ce flux observé et non dans une reconstitution de la hiérarchie des attributs ayant conduit à ce flux.
Le transfert exprimant la concurrence entre deux marques est exprimé comme étant le résultat d'une relation de concurrence qui se manifeste à trois niveaux :

- une concurrence qui repose sur les attributs des deux marques. Il s'agit alors de la "forme" du produit, de sa marque etc ..
- une concurrence qui repose sur les variables de contrôle marketing comme les prix relatifs, la pression publicitaire relative etc ...
- enfin, l'interaction entre les attributs des marques et les éléments de leur mix.

Il est à noter que dans les approches que nous avons vues précédemment (comme celle de Rao et Sabavala, Grover et Dillon, le modèle de Hendry), seul le premier de ces trois points est considéré (les attributs du produit) pour étudier la concurrence entre les marques. Cette limite présente l'avantage d'avoir un modèle qui reste *"en apparence"* conforme à la théorie des choix et du traitement de l'information selon laquelle un individu effectue ses choix à partir des attributs du produit.
Cependant, comme ces modèles ne fonctionnent qu'à partir de données de comportement, les auteurs ne considèrent dans ces modèles que des attributs *"physiques"* et non *"perceptuels"*, étant bien entendu qu'ils ne peuvent pas considérer les attributs perçus du produit. Ceci constitue une vision appauvrie de la théorie du traitement de l'information et constitue un retour vers des modèles où le consommateur est conçu de manière sommaire. Les attributs perçus sont fonction des caractéristiques physiques du produit mais aussi de l'action marketing qui a été menée (et d'autres facteurs bien sûr ..). Si dans les modèles de traitement de l'information développés en analyse conjointe on étudie directement le

lien entre les caractéristiques physiques du produit et les préférences (ou les choix) exprimées par le répondant, il est à noter qu'une telle approche a été développée dans le cadre de l'étude de concepts de produits nouveaux et non pas de produits existants sur un marché. Les répondants ne connaissent le produit qu'à partir de la description qui leur en a été faite. Les variables d'action marketing perçues n'ont pas à jouer un rôle dans ces conditions. Carpenter et Lehmann, en intégrant les éléments d'action marketing dans leur modèle, enrichissent la vision du comportement de choix et se rapprochent ainsi d'une réalité "*perçue*". Ils utilisent des attributs physiques du produit (ou encore attributs écologiques, voir Holbrook,1981). Ces attributs n'étant pas suffisants pour expliquer les relations de concurrence, les auteurs prennent en compte les éléments du mix de ces produits.
Ce point est très important et mérite d'être souligné. Dans de nombreux modèles d'analyse de la concurrence (comme le modèle de Hendry) les auteurs contournent le problème de la "*pauvreté des données de comportement*" en n'utilisant que des attributs physiques. C'est ainsi que la structure des marchés est expliquée à partir d'attributs dichotomiques (café en grain / café moulu, boisson gazeuse / non-gazeuse) pour lesquels les problèmes de perception ne se posent pas: un café est soit "en grains" soit "moulu". Une telle approche, bien qu'elle soit orientée vers le consommateur, peut être qualifiée d'approche "*producteur*" puisque l'élément fondamental de la théorie du consommateur, sa perception et ses attitudes sont laissées pour compte. Que faire si le marché étudié se structure sur la base d'attributs de type continu (par exemple goût amer/goût sucré) ?

DESCRIPTION DU MODELE

Les auteurs ont construit un indice de mesure de la concurrence à partir du modèle classique d'utilité. Le transfert d'une marque vers une autre est fonction de l'utilité qu'apporte ce changement.
Soit i le segment de consommateurs ayant acheté la marque i en t.
En t+1, parmi l'ensemble des N marques disponibles, la marque j procure une utilité uij qui peut s'écrire :

$$u_{ij} = v_{ij} + e_{ij}$$

v_{ij} est la composante déterministe de l'utilité
e_{ij} est la composante aléatoire de l'utilité, elle suit une loi "*exponentielle double* ".

La probabilité conditionnelle du transfert de i vers j pour les individus appartenant au segment i s"exprime de la manière suivante:.

$$\pi_{ij} = \frac{\exp^{v_{ij}}}{\sum_{k \in M} \exp^{v_{ij}}}$$

π_{ij} a les propriétés suivantes :

- il représente l'utilité de changer de la marque i pour la marque j divisée par l'utilité de changer pour toutes les marques du marché.
- c'est un modèle dynamique, πij représente le processus dynamique qui aboutit à la formation des parts de marché.

La composante déterministe de l'utilité est exprimée en fonction des attributs et des actions marketing de *"l'ensemble"* des marques appartenant au marché. vij s'exprime comme une combinaison linéaire des trois éléments constitutifs de la concurrence que nous avons cités précédemment.

$$v_{ij} = \sum_{k\varepsilon C} \alpha_{ik}.X_{ijk} + \sum_{l\varepsilon F} \beta_{il}.W_{ijl} + \sum_{m\varepsilon C} \sum_{n\varepsilon F} \gamma_{imn}.X_{ijm}.W_{ijn}$$

(1) (2) (3)

C représente l'ensemble des variables contrôlées par le marketing
F représente l'ensemble des caractéristiques des produits
X_{ijk} est la valeur de la marque j, relativement à i, selon la variable de contrôle k
W_{ijl} est la valeur de la marque j, relativement à i, selon la caractéristique l
α_{ik} , β_{il} et γ_{imn} sont respectivement les poids des variables d'action marketing, des attributs du produit et des éléments d'interaction, dans la formation de l'utilité.

Le terme (1) est relatif aux variables d'action marketing
Le terme (2) est relatif aux attributs des produits
Le terme (3) est relatif aux interactions possibles entre les deux termes précédents .

En notation matricielle, cette relation peut s'écrire :

$$V_{ij} = Z_{ij}.D_i$$

Z_{ij} est la matrice des variables explicatives
D_i est le vecteur des poids correspondants. :

$$\pi ij = \frac{\exp Z_{ij}.D_i}{\sum_{k\varepsilon M} \exp Z_{ik}.D_i}$$

Après transformation par les logarithmes, la probabilité conditionnelle peut s'écrire :

$$\log\pi_{ij} = Z_{ij}.D_i - \log\left(\sum_{k\in M} \exp^{Z_{ij}.D_i}\right) \qquad (1)$$

Le deuxième terme ne dépendant pas de i, on pose alors :

$$Mui = -\log\left(\sum_{k\in M} \exp^{Z_{ij}.D_i}\right) \qquad (2)$$

En remplaçant (2) dans (1) il vient :

$$\log\pi_{ij} = Mui + Z_{ij}.D_i \quad (3)$$

(3) est estimé par régression multiple en utilisant les moindres carrés généralisés .

La probabilité π_{ij} est opérationnalisée par $T_{ij}/T_{i.}$ qui en est l'estimateur par le maximum de vraisemblance (voir chapitre I, section 3).

T_{ij} est le nombre d'individus ayant effectué un transfert de i vers j sur la période.
$T_{i.}$ est le nombre d'individus ayant effectué un transfert de i vers une autre marque sur la période .

Les variables "prix relatif" et "publicité relative" (respectivement Pij et Aij) sont opérationnalisées de la manière suivante :

$$P_{ij} = P_i/P_j$$
$$A_{ij} = A_i/A_j$$

P_i, P_j, A_i ,A_j sont les valeurs moyennes des prix et des publicités des marques i et j sur la période .

L'estimation de (3) fournit quatre types de résultats :

- Mu_i est l'indice de fidélité à la marque. Plus il est faible et plus la fidélité est forte. En effet Mui représente l'effet agrégé de toutes les autres marques sur πij. Une faible valeur de Mui signifie que la marque, en moyenne, échange peu de consommateurs avec les autres marques du marché.
- les α_{ik} représentent la propension de la marque i à retenir ses acheteurs par l'action des variables de son mix, si cette valeur est négative. Dans le cas inverse, elle représente la propension de la marque à perdre des acheteurs par l'utilisation de ses variables d'action marketing.
- de la même manière, les ßil représentent la propension à changer pour des marques possédant les mêmes attributs, si les ßil sont négatifs. Par exemple si ßil représente une caractéristique physique du produit (boisson gazeuse ou non

gazeuse) cela signifie que le changement de marque s'effectue pour rechercher une autre caractéristique du produit. Ceci peut caractériser des situations d'usage multiples du produit ou bien de la recherche de variété.

- les coefficients d'interaction γ_{lmn} caractérisent l'effet joint d'une variable du mix sur un attribut en terme de changement de marque. Si cette valeur est négative entre publicité et forme du produit, cela signifie que la publicité pour la marque i augmente la propension à changer pour un produit de même forme. Le signe et la taille de ces coefficients permettent d'effectuer un diagnostic intéressant de l'influence des variables d'action sur la *"solidité concurrentielle de la marque"*.

Nous soulignerons que le modèle de Carpenter et Lehmann a la vertu d'établir un pont entre les approches stochastiques dans lesquelles le comportement est observé mais est considéré en quelque sorte comme "inexplicable" (les probabilités d'achat sont alors calculées sans essayer d'expliquer ces probabilités comme étant le résultat d'un processus psychologique de traitement de l'information par le consommateur) et les approches déterministes qui cherchent avant tout à expliquer le processus conduisant à un comportement particulier.
Replacé dans le contexte actuel de développement des Modèles à Source Unique dans lesquels les variables d'action auxquelles sont soumis les consommateurs sont de mieux en mieux maîtrisées (voir deuxième partie chapitre I, section 1), ce modèle nous apparait extrêmement prometteur. Il est à noter qu'en terme de mesure de la concurrence, ce modèle n'apporte rien de nouveau puisque la concurrence y est mesurée par les probabilités conditionnelles de transferts entre marques (Pi/j). Par contre, le modèle est novateur en ce qu'il apporte une vision explicative de la concurrence. C'est la raison pour laquelle nous l'avons fait figurer ici.

SS3-2 INDICES CONSTRUITS A PARTIR DE MODELES DE CHOIX D'ORDRE 1 : PRISE EN COMPTE DE LA RECHERCHE DE VARIETE

Comme nous l'avons noté dans le premier chapitre (section 2), la recherche de variété peut avoir dans certaines conditions une influence importante sur le comportement de choix du consommateur.
Nous avons montré que les concepts de substituabilité et de complémentarité sont des concepts relatifs au niveau de la *"réalité du consommateur"* alors que le concept de concurrence est relatif à la *"réalité de marché"*, ce qui explique peut être pourquoi les modèles de recherche de variété sont estimés au plan individuel. Ainsi, le concept de recherche de variété n'a (toujours selon nous) pas d'intérêt pour mesurer le niveau de la concurrence entre les marques, il en a par contre lorsque le responsable marketing cherche à expliquer cette concurrence.
Pour analyser la structure d'un marché, il est nécessaire de définir des mesures de la concurrence entre les marques, il était donc logique que les chercheurs ayant étudié le concept de recherche de variété se soient tournés vers des applications de leurs modèles à l'analyse de la structure des marchés .
Nous présenterons ici deux approches de l'analyse de la structure des marchés qui utilisent le concept de recherche de variété . Il s'agit des approches proposées par Givon (1985) et Lattin et McAlister (1985). Nous verrons que ces deux approches diffèrent de

par la nature même de la définition que donnent les auteurs au concept de recherche de variété.
Ces approches sont présentées en tant qu'indices stochastiques puisque le comportement du consommateur y est modélisé de manière stochastique. Les modèles utilisés sont "d'ordre 1" puisqu'un achat influencé par la recherche de variété fait référence à un (ou plusieurs) achat(s) antérieur(s).

SS3-2-1 LE MODELE DE GIVON

Givon (1985) a développé un modèle d'analyse de la structure des marchés dans lequel la mesure de la concurrence entre les marques fait appel à un modèle de recherche de variété qu'il a lui-même développé (voir Givon 1984) et repris des travaux de thèse de Reibstein (1975, cité dans Givon 1985). A la différence de Lattin et McAlister (1985), pour Givon, la recherche de variété est un concept multi-dimensionnel. L'individu est supposé avoir des attentes en matière de recherche de variété qui varient d'un attribut du produit à un autre. La concurrence entre les marques est estimée à partir des probabilités conditionnelles de changement entre marques et ce, à deux niveaux :

- entre deux partitions du marché
- au sein d'une même partition (notons qu'il s'agit là de la logique du modèle de Hendry).

Le comportement du consommateur est modélisé à partir des hypothèses suivantes :

H1 Le processus de choix du consommateur est hiérarchique, il choisit en fonction des attributs (qui constituent des partitions de marché) jusqu'à la détermination du choix d'une marque particulière. Cette hiérarchie de choix peut être caractérisée de la manière suivante :

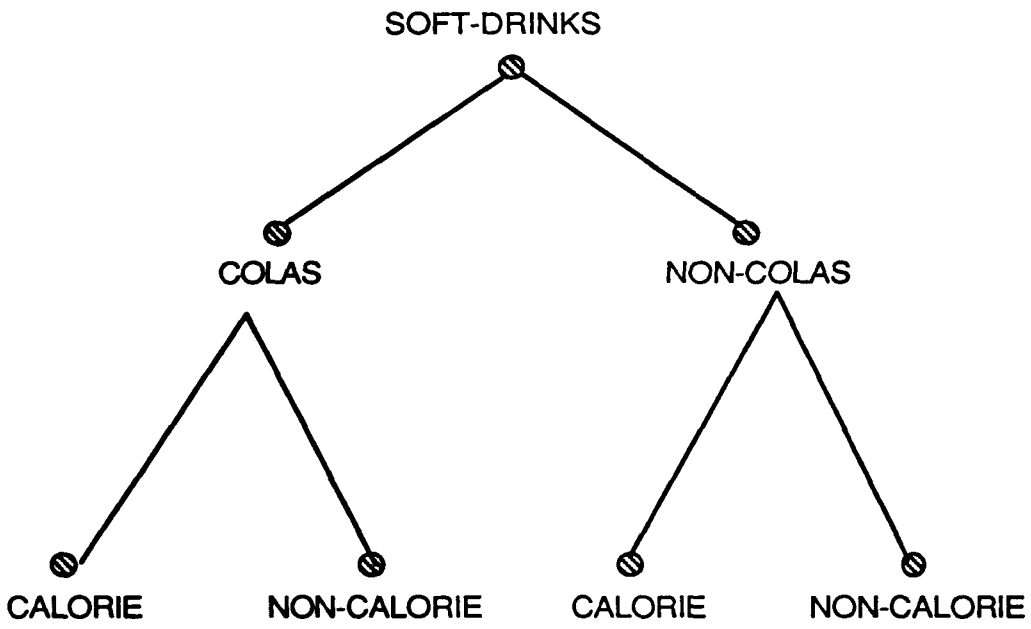
- au sommet de l'arbre de choix (c'est à dire au niveau des attributs qui sont pris en compte en premier lieu) il y a beaucoup de transferts entre des marques appartenant à des partitions différentes et peu entre des marques appartenant à la même partition. Par exemple, sur l'arbre de la figure 15, il y aura beaucoup de transferts entre des marques appartenant à la partition des boissons au cola et à la partition des boissons sans cola. Ces marques seront qualifiées de complémentaires.
- à la base de l'arbre de choix (c'est à dire au niveau des attributs qui sont considérés en dernier), il y a beaucoup de changements au sein des partitions et peu entre les partitions. Par exemple, au sein de la branche des boissons au cola, il y a beaucoup de transferts entre des marques appartenant à la partition des boissons sans calories, marques qui seront qualifiées de substituables.Par contre, on observe peu de transferts entre des marques appartenant aux partitions avec calories et sans calorie.

H2 Les individus partitionnent les marchés sur la base d'attributs et satisfont à leur besoin de recherche de variété en changeant d'une partition à une autre. Il y a donc deux types de changements liés à des motivations différentes :

- des changements au sein d'une même partition, qui caractérisent une relation de substituabilité. Sur notre exemple, cela correspondra à des changements entre des marques de cola ou entre des marques de non-cola. Ces marques seront dites complémentaires. Par contre, il y a peu de changements entre les marques appartenant à la même partition.

- des changements entre partitions différentes, qui caractérisent un besoin de recherche de variété (et donc des relations de complémentarité). En effet, au sein d'une même partition, toutes les marques sont équivalentes du point de vue de l'attribut (physique) qui définit cette partition. Pour satisfaire leur besoin de recherche de variété relatif à cet attribut, les consommateurs doivent changer de partition.

FIGURE 15 ARBRE INDIVIDUEL DE DECISION ET STRUCTURE D'UN MARCHE



La recherche de variété conduit ainsi à changer pour des marques appartenant à des partitions différentes alors que la concurrence conduit à changer pour des marques appartenant à la même partition. Cette définition permet d'opérationnaliser une mesure de la concurrence qui sera fonction de la recherche de variété entre les marques. La recherche de variété est caractérisée pour un attribut particulier et un individu particulier.
Un critère de choix (attribut) auquel correspond un faible niveau de recherche de variété, apparait plus haut dans la hiérarchie de choix qu'un critère caractérisé par un fort niveau de recherche de variété.

DESCRIPTION DU MODELE

Dans le modèle, le changement entre deux marques est le résultat de deux composantes de l'utilité:

- l'utilité dérivée de la marque pour laquelle le consommateur change

- l'utilité dérivée du changement lui même, cette utilité pouvant être positive ou négative

La préférence pour une marque i est obtenue à partir du modèle de Luce (1959) :

$$\pi_i = \frac{u_i}{\sum_{i=1}^{n} u_i}$$

π_i est la préférence pour la marque i ou encore la probabilité de choix de la marque i dans le cadre du modèle multinomial d'ordre 0 .
u_i est l'utilité procurée par la marque i
n est le nombre de marques dans l'ensemble évoqué du consommateur

La préférence π_i est supposée constante au cours du temps (cas d'un modèle d'ordre 0). Cependant, l'introduction d'une deuxième composante dans le choix, qui est l'utilité dérivée du changement "*pour lui même*", procure au modèle les caractéristiques d'un modèle de choix d'ordre 1.

Le choix en l'instant t d'une marque dépend de l'utilité de cette marque et du fait d'avoir acheté ou non cette marque en l'instant t-1. Il s'agit donc d'une forme particulière de modèle d'ordre 1 où le choix est une combinaison linéaire de :
- l'importance de la recherche de variété
- la préférence pour la marque

<u>* PROBABILITE CONDITIONNELLE DE PASSAGE ENTRE DEUX MARQUES APPARTENANT A DEUX PARTITIONS DIFFERENTES: Pij</u>
Pour normaliser les préférences, on pose :

$$\sum_{i=1}^{n} \pi_i = 1$$

$$\Pi_k = \sum_{i \varepsilon k} \pi_i$$

Π_k est la préférence globale pour la partition k

<u>1 Cas d'un "individu chercheur de variété"</u>
Soit V l'indice de recherche de variété du consommateur, V ε [0,1]. Si le niveau de recherche de variété est extrême, alors Pij ne dépend pas de la préférence pour la marque vers laquelle le consommateur change.

$P_{ij} = 1 / (n-n_k)$ pour tout j différent de k
P_{ij} est la probabilité de choix de la marque j quand le consommateur a choisi i à l'occasion précédente

nk est le nombre de marques de la partition k dans laquelle se trouve i, la marque dont le consommateur change.
$P_{jj} = 0$ pour tout j

Si le niveau de recherche de variété est intermédiaire, le changement pour une marque dépend en partie de la préférence pour cette marque et en partie du désir de changement :

$$P_{ij} = V \,.\, 1/(1-\Pi_k) + (1-V).\pi_j$$
$$P_{jj} = V.0 + (1-V)\pi_j = (1-V)\pi_j$$

2 Cas d'un individu refusant la variété

Soit I le coefficient de renforcement dans le comportement de choix, $I \in [0,1]$. Si I=1, cas extrême,

$$P_{ij} = 0$$
$$P_{jj} = 1$$

Le consommateur ne tolère pas de passer d'une marque à une autre.

Si I est à un niveau intermédiaire :

$$P_{ij} = I \,.\, 0 + (1-I)\pi_j = (1-I)\pi_j$$
$$P_{jj} = I + (1-I).\pi_j$$

Tout ce qui précède peut être réécrit de la manière suivante en posant (VP : indice de recherche de variété ou de renforcement) :

$VP = V$ si $V \in [0,1]$
$VP = -I$ si $I \in [0,1]$

La probabilité de changer de la marque i pour la marque j s'écrit alors à partir de la connaissance de la préférence pour la marque j et de la manière dont se caractérise le consommateur en terme de recherche de variété ou de renforcement.

$$P_{ij} = \frac{(|VP| + VP).\pi_j}{2.(1-\Pi_k)} + (1-|VP|).\pi_j \quad (1)$$

$VP \in [-1,1]$
$\pi_j \in [0,1]$
i appartient à la partition k
j n'appartient pas à la partition k

Nous pouvons observer que (1) ne dépend que de πj et Πk et ne dépend pas de πi.

Si l'individu est caractérisé par une tendance à la recherche de variété qui est positive, P_{ij} s'écrit :

$$P_{ij} = VP(\pi_j/(1-\Pi_k) + (1-VP).\pi_j \quad (1')$$

Si l'individu est caractérisé par une tendance à la recherche de variété qui est négative, Pij s'écrit :

$$P_{ij} = (1-VP).\pi_j \quad (1'')$$

* PROBABILITE CONDITIONNELLE DE PASSAGE ENTRE DEUX MARQUES APPARTENANT A LA MEME PARTITION

Dans ce cas, la marque j est évaluée par rapport à sa partition k, soit $\pi_j/\Pi k$

$$P_{ij} = \frac{(|VP| - VP).\pi_j}{\Pi_k} + (1 - |VP|).\pi_j \quad (2)$$

i et j appartiennent à la partition k

Si l'individu est caractérisé par une recherche de variété positive (2) s'écrit :

$$P_{ij} = (1-VP).\pi_j \quad (2')$$

Si l'individu est caractérisé par une recherche de variété négative (2) s'écrit :

$$P_{ij} = VP(\pi_j/\Pi k) + (1-VP).\pi_j$$

Au travers de ces relations, il ressort que la recherche de variété (quand elle est positive) a pour effet de diminuer la préférence pour les marques qui bénéficiaient au départ d'une forte préférence (π_j est fort) et d'augmenter la préférence pour les marques qui à l'inverse bénéficient d'une préférence faible (π_j est faible).

* ESTIMATION (pour chaque individu)

Les π_j et VP sont estimés par la méthode du maximum de vraisemblance à partir de données d'historiques d'achats (notamment à partir des n_{ij} nombre de transferts de i vers j sur la période). On peut ensuite estimer :

- la valeur de VP pour chaque partition définie a priori.
- la valeur de VP pour chaque critère (attribut du produit) susceptible d'appartenir à la hiérarchie.

Cette estimation est donc coûteuse en temps puisque tous les indices sont définis au niveau individuel.

L'avantage d'un tel indice, sur les indices de transfert classiques, comme celui proposé par Rao et Sabavala, est de prendre en compte (pour partie au moins) les motifs du changement, notamment la lassitude à l'égard des attributs du produit et le fait que deux marques soient désirées pour des usages différents.

Une des limites sérieuse du modèle est qu'il ne prend pas en compte le temps de manière explicite. L'indice de recherche de variété, pour un attribut et un individu, est supposé être constant au cours du temps, quelle que soit l'expérience accumulée de l'individu avec cet attribut pendant les périodes antérieures au choix.

Le problème lié au calcul de tels indices de concurrence est, nous l'avons déjà souligné, qu'ils reposent sur des partitions de marché définies à partir de niveaux d'attributs physiques de type discret. Une telle hypothèse est-elle cohérente avec une approche de la formation des préférences à partir d'attributs perceptuels qui ne possèdent pas, en général, ce caractère discret ?
Enfin, si cette approche est très riche par sa modélisation du consommateur (modélisation très élégante d'un point de vue statistique), il est à craindre que son utilité pour le manager soit limitée puisque l'approche a un caractère à la fois "*confirmatoire*" et "*individuel*", (il faut faire des hypothèses de partitions au niveau "*individuel*").

SS3-2-2 LE MODELE DE McALISTER ET LATTIN

Lattin et McAlister (1985) ont proposé une approche de la mesure de la concurrence entre les marques à partir des relations de substituabilité et de complémentarité existant entre ces marques. Les auteurs procèdent à partir des considérations suivantes :

- les mesures traditionnelles de la concurrence, comme l'élasticité des économistes, ne permettent pas de faire la différence entre biens substituables et biens complémentaires .- changer de produit pour cause de "*recherche de variété*", n'implique pas que les produits considérés soient substituts. Au contraire, le changement peut avoir pour objet de satisfaire différentes composantes du besoin (besoin de composition exprimé à travers la recherche de variété). Le changement de marque exprimerait alors le degré de complémentarité entre les marques plutôt que la substituabilité.

Ainsi, selon les auteurs :

- si différents produits sont choisis pour satisfaire la même composante du besoin, ils seront dits "*substituts*".
- si différents produits sont choisis pour satisfaire différentes composantes du besoin, ils seront dits "*complémentaires*".

McAlister et Lattin développent ainsi un modèle où les probabilités conditionnelles d'achat d'une marque quand le consommateur en a choisi une autre à la période précédente permettent de détecter si deux produits (marques) sont complémentaires ou substituts. En effet, la recherche de variété est une variable dont la présence inverse la relation entre les probabilités conditionnelles d'achat et la substituabilité.

Si un individu est "*chercheur de variété*""(recherche de variété positive), trois cas peuvent se présenter :

- si la marque A est très similaire à la marque B, alors l'achat de A diminue la probabilité d'acheter la marque B à la période suivante, A et B sont dits substituts.
- inversement, si A et C sont très différents, alors l'achat de A à une période augmente la probabilité d'acheter B à la période suivante, A et C sont dits complémentaires.

- si A et D sont indépendants, la consommation de A n'influe pas sur la consommation future de D, A et D sont dit indépendants.

Les relations ainsi présentées seront inversées si l'individu est caractérisé par une tendance au renforcement du comportement de choix (recherche de variété négative). Le modèle distingue en fait deux types de complémentarité :

- la complémentarité qui se manifeste de manière instantanée. Elle correspond à la complémentarité classique définie par les économistes (par exemple cigarettes et allumettes) qui n'a de sens que pour des produits appartenant à différentes catégories. En reprenant la terminologie de la section 1 précédente, il s'agit de *"substituabilité horizontale"*.
- la complémentarité dans le temps qui elle est liée à l'existence d'un besoin de composition. Elle conduit le consommateur, par l'intermédiaire de sa recherche de variété, à satisfaire les différentes composantes dans le temps de son besoin et donc à choisir des marques *"différentes"*, marques qui ici appartiennent à la même catégorie de produit (*complémentarité verticale).*

Nous rappelons que selon nous, Substituabilité et Complémentarité sont deux concepts qui n'ont de valeur qu'au niveau du consommateur. Par contre, au niveau du marché (pour une catégorie de produit particulière) nous ne considérons qu'un seul type de relation entre les marques : la concurrence.

DESCRIPTION DU MODELE

Le modèle repose sur quatre hypothèses :

H1 Un produit est un ensemble de caractéristiques satisfaisant des besoins. La préférence pour un produit est la somme des évaluations de ce produit selon ses principales caractéristiques (il est donc fait référence à un modèle multi-attributs de type compensatoire).

H2 Les préférences pour les produits sont influencées par le dernier achat . Les choix suivent donc un modèle de Markov d'ordre 1.

H3 La consommation d'une caractéristique particulière du produit au cours d'une période diminue l'attrait de celle ci au cours de la période suivante. L'importance de cette diminution est fonction de la recherche de variété caractérisant chaque individu.

H4 La probabilité conditionnelle d'achat d'un produit par un consommateur est proportionnelle à sa préférence relative pour le produit. Il s'agit donc du modèle de Luce, mais une fois que la recherche de variété à été prise en compte par le modèle d'ordre 1.

Soit :

E_C: l'ensemble des marques choisies par le consommateur c au cours de la période.

π_{jc}: la préférence (non conditionnelle) du consommateur c pour la marque j Ce n'est pas une part de marché car celle-ci doit prendre en compte l'effet de la recherche de variété.

On pose $\sum_{j \in Ec} \pi_{jc} = 1$ (1) afin de normaliser les préférences

S_{ijc}: la valeur de toutes les caractéristiques que le consommateur c value et qui sont partagées par les marques i et j.
V_c, l'indice de recherche de variété qui caractérise le consommateur c,
$V_c \in]0, 1]$
$\mu_{i/jc}$, la préférence conditionnelle de l'individu c pour la marque i sachant qu'il a acheté (choisi) j à la période précédente.

$$\mu_{i/jc} = \pi_{ic} - V_c \cdot S_{ijc} \quad (2)$$

Notons que: $\mu_{i/jc} = \pi_{jc} - V_c.S_{jjc} = \pi_{jc}.(1 - V_c) < \pi_{jc}$ pour tout j. Ceci signifie qu'à cause de la recherche de variété, la probabilité de ré-acheter une marque est inférieure à la probabilité de choix (marginale) de cette marque.

$P_{i/jc}$, la probabilité conditionnelle pour que le consommateur c achète la marque i sachant qu'il a acheté j à la période précédente s'écrit finalement:

$$P_{i/jc} = \frac{\mu_{i/jc}}{\sum_{k \in E_c} \mu_{i/jc}} = \frac{\pi_{ic} - V_c \cdot S_{ijc}}{1 - V_c \cdot \sum_{k \in E_c} S_{kjc}} \quad (3)$$

La mesure de la substituabilité (complémentarité) est alors opérationnalisée par la différence :

$$P_{i/jc} - \pi_{ic} \quad (4)$$

Trois cas sont possibles :

1 - Si l'expression (4) est de signe négatif, cela signifie que les produits sont substituables

en effet, $\pi_{ic} < \frac{S_{ijc}}{\sum_{k \in E_c} S_{kjc}}$

Ceci signifie que i partage avec j une quantité de caractéristiques désirées plus importante que ne le laisse prévoir la simple proportionnalité.
2 - Si l'expression (4) est de signe positif, cela signifie que les produits sont complémentaires (dans le sens où nous l'avons défini précédemment) i et j partagent alors en commun moins de caractéristiques désirées que n'en partage en moyenne j avec l'ensemble des autres marques de Ec.
3 - si (4) = 0, cas qui se présente
- si l'indice de recherche de variété V_c =0, l'indice de mesure de la substituabilité n'est alors plus utilisable.
- si les deux produits sont complètement indépendants ($S_{ijc} = 0$) alors ils appartiennent à des sous-marchés différents (S_{ijc} est peu différent de 0, les deux marques n'ont rien en commun).

$P_{i/jc} = \pi_{ic}$ et $P_{i/jc} = \pi_{jc}$

- si le produit n'a jamais été acheté, $\pi_{ic} = 0$, les deux produits sont alors indépendants.

ESTIMATION

πic, S_{ijc} sont estimés à partir de l'historique d'achat de chaque consommateur c.

Soit, $\tilde{N}_{i/jc}$, la probabilité de changement de i vers j estimée à partir des transferts entre i et j effectués par le consommateur c. Il s'agit alors d'estimer Vc, πic et Si/jc de façon à minimiser :

$$\sum_{i \in E_c} \sum_{j \in E_c} \tilde{N}_{i/jc} - \frac{\pi_{ic} - V_c \cdot S_{ijc}}{1 - V_c \cdot \sum_{k \in E_c} S_{kj}}$$

sous les contraintes

$$\sum_{i \in E_c} \pi_{ic} = 1 \text{ (contrainte d'échelle)}$$

$$\sum_{\substack{k \neq j \\ j \in E_c}} S_{kjc} < \pi_{jc}$$

$$\pi_{ic} > 0 \quad \text{pour tout } i \in E_c$$

$$0 < V_c < 1$$

$$S_{ijc} > 0$$

L'agrégation est effectuée par la suite en effectuant la moyenne des $(P_{i/j} - \pi_i)$ pondérée par le poids des individus.

$$(P_{i/j} - \pi_i) = \sum_c w_c.(P_{i/jc} - \pi_{ic})$$

$$w_c = n_c / \sum_c n_c ,$$

- w_c est un facteur de pondération des individus en fonction du nombre d'achats qu'ils ont effectués. Plus ces achats sont nombreux (gros consommateurs) et plus les valeurs estimées sont fiables, ce qui justifie le fait que leur poids soit plus fort dans le modèle agrégé.
- n_c est le nombre de produits choisis par le consommateur c sur la période étudiée

LIMITES ET AVANTAGES DE CETTE APPROCHE

Au plan statistique, la procédure d'estimation est lourde puisqu'il est nécessaire d'effectuer les estimations au niveau individuel avant de passer à l'agrégation. D'autre part, les propriétés des estimateurs obtenus par la programmation non-linéaire sont mal connues. Il n'est donc pas possible d'établir leur fiabilité.

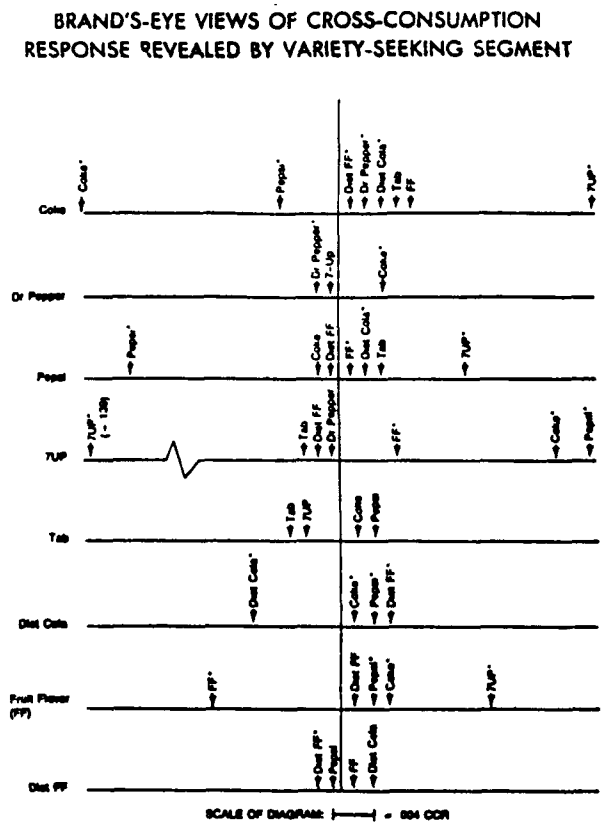
Au plan pratique, les données collectées sont de classiques données d'historiques d'achats. Le modèle permet de représenter la relation de chaque marque sur un continuum allant de "*très substituable*" à "*fortement complémentaire*" (voir graphique suivant).

Au plan conceptuel, nous retiendrons les point suivants :

- Le modèle apporte peu de conclusions quand $\pi ic = \pi i/jc$, ce qui peut être dû à une indépendance entre les marques i et j ou à une annulation des effets de substitution et de complémentarité (les marques sont similaires selon certaines caractéristiques, complémentaires selon d'autres).
- De même quand la recherche de variété est proche de zéro (à partir de quel moment est-il possible de considérer que le niveau de recherche de variété est nul ?) le modèle n'est plus utilisable. Ces restrictions limitent le champ d'utilisation du modèle.
- $P_{i/jc} - \pi_{ic} < 0$ si $\pi_{ic} < S_{ij} / \sum_{k} S_{kjc}$,

C'est le cas lorsque i partage avec j moins de bénéfices que la moyenne des produits n'en partagent avec j. Le problème d'une moyenne est que si tous les produits sont substituables, l'analyse conduit à donner pour complémentaires des produits qui à l'origine étaient substituables (ou moins substituables que la moyenne). La mesure de la substituabilité est donc relative et non pas absolue. `

FIGURE 16 REPRESENTATION DE LA SUBSTITUABILITE



"Extrait de Lattin et McAlister (1985) p 337 "

L' intérêt de cette approche est de tenir compte explicitement de la recherche de variété des consommateurs pour mesurer la concurrence entre les marques. Cependant, nous pensons que cette approche n'a vraiment d'intérêt qu'au seul plan individuel. Au plan global, l'agrégation par la moyenne des estimateurs individuels peut aboutir à une sous-estimation des liens de concurrence entre les marques (si deux marques sont substituables pour certains consommateurs et complémentaires pour d'autres).

A l'opposé, la non-prise en compte de la recherche de variété lors du calcul de la concurrence entre les marques (avec les données d'historiques d'achats) peut conduire à sous-estimer la relation de substituabilité entre deux produits à partir du moment où les consommateurs *"recherchent de la variété"*.

Finalement, cette approche nous semble très riche dans la mesure où elle porte un regard nouveau sur les concepts de substituabilité et de complémentarité. Elle amène à poser le problème de la correspondance entre substituabilité, complémentarité et concurrence comme nous en avons débattu au chapitre I section 4 :

- à un moment donné du temps
- sur une période de temps

Cette approche a l'avantage de permettre d'effectuer une segmentation des consommateurs sur la base de leurs structures individuelles de substituabilité telles qu'ils

les manifestent à travers leurs historiques d'achats. Le modèle permet de prendre en compte la non-symétrie des relations de concurrence, point que nous avons déjà noté comme important. Dans l'exemple donné par les auteurs, Pepsi-Cola est fortement substitut de Coca-Cola, l'inverse étant moins vrai et ce pour 80% des consommateurs.

SECTION 3 AUTRES APPROCHES DE LA MESURE DE LA CONCURRENCE UTILISANT DES DONNEES D'HISTORIQUES D'ACHATS

Bien que la majorité des approches procédant à partir de données de panels utilisent le concept de *"transfert entre marques"* comme base de la mesure de la concurrence, d'autres méthodes ont été proposées. Nous présenterons dans cette section deux approches qui permettent d'estimer des indices individuels de concurrence.
Par indices individuels, nous désignons tous les indices qui sont calculés au plan individuel (en fait au niveau du ménage) quitte à être ensuite agrégés, comme c'est souvent le cas . Cette agrégation peut se faire au niveau d'un segment particulier ou de la population totale.

Comment agréger les indices calculés au plan individuel en indices représentatifs du marché ou d'un segment du marché ? Ce problème est en général résolu en effectuant la moyenne, pondérée par les quantités achetées, des indices calculés au plan individuel. Une telle procédure ne garantit cependant pas la fiabilité statistique des estimateurs agrégés et n'a de valeur que si la population est réellement homogène.

SS1 UTILISATION DES PARTS D'ACHATS COMME MESURE DE LA CONCURRENCE

Merunka et Bourgeat (1988) ont présenté un indice calculé au plan individuel, à partir des parts de choix des consommateurs.

$$I_{AB}^{l} = \frac{\text{Min}^{l}(A,B)}{\text{Max}^{l}(A,B)}$$

I_{AB}^{l} est un indice d'intensité de la concurrence entre les marques A et B pour le ménage l.
$\text{Min}^{l}(A,B)$ est la part de choix la moins élevée parmi les marques A et B, pour le consommateur l.
$\text{Max}^{l}(A,B)$ est la part de choix la plus élevée parmi les marques A et B pour le consommateur l.

Lorsque deux marques se partagent à l'identique le panier de la ménagère (elles représentent chacune 50% des parts de choix du ménage) la concurrence est maximale, l'indice est égal à 1. Elle est égale à 0 dans le cas contraire .

Dans la méthodologie proposée par les auteurs (le modèle Compétitor), les indices individuels sont agrégés au niveau du marché en effectuant la moyenne, pondérée par les volumes de consommation de chaque ménage, des indices calculés au plan individuel.
Cet indice peut aussi être calculé à partir de données de préférences (données de jugement) après calibrage de ces préférences en utilisant un modèle Logit, selon la méthodologie proposée par Hauser, Urban et Johnson (1984).

Cet indice présente un certain nombre d'avantages :

- il repose sur très peu d'hypothèses
- il ne nécessite pas d'avoir effectué au préalable une définition du marché, puisque le calcul des I_{AB} pour chaque ménage ne dépend pas d'une telle définition. Nous présentons cet indice plus en détail dans la deuxième partie, chapitre 1 section 3.
- il est très simple à calculer, ce qui est un avantage pour son utilisation pratique .

Il comporte cependant un certain nombre de limites qu'il convient de souligner:

- tout d'abord, cet indice ne prend pas appui sur une théorie du choix clairement établie, que mesure-t-il précisément ?
- il ne prend pas en compte la dynamique dans le comportement de choix puisqu'il ne considère que les quantités achetées des marques sans tenir compte de la manière dont ces achats s'inscrivent dans l'historique d'achat du consommateur.
- il considère que la concurrence est symétrique, ce qui à notre avis est un appauvrissement sérieux du concept .
- il ne tient pas compte des parts de marché des marques achetées par le ménage . Une marque peut être moins présente dans le panier de la ménagère parce qu'elle est en moyenne moins consommée qu'une autre. Il pourrait être intéressant de tenir compte de ce facteur dans le calcul de l'intensité concurrentielle.
- enfin, les auteurs ne posent pas le problème de la pertinence (d'un point de vue statistique) de l'agrégation d'indices obtenus au plan individuel. L'hétérogénéité des individus permet elle de faire une agrégation sans plus de précautions qu'avec une simple moyenne pondérée ?

SS2 UTILISATION DES PERIODES INTER-ACHATS COMME MESURE DE LA CONCURRENCE (MODELE D'INCIDENCE)

Les modèles d'incidence constituent une alternative aux modèles stochastiques de choix. C'est la loi des périodes inter-achat qui y est étudiée et non celle des actes d'achat, une fois placé à une occasion particulière.
Fraser et Bradford (1983) définissent comme substituts des produits (marques) achetés pour la satisfaction d'un besoin particulier. Constatant qu'une élasticité croisée positive est une condition nécessaire mais non suffisante de substituabilité, ils définissent un indice de "substituabilité révélée", l'IRS (Index of Revealed Substituability). Cet indice est calculé à partir des périodes inter-achats des produits (marques) considérés.
Soit pour un consommateur:

X_{ii} : la période de temps séparant deux achats (ou période de consommation selon les cas) de la marque i
X_{ij} : la période de temps séparant deux achats des marques i et j

Le modèle prend pour hypothèse de base que l'intervalle de temps séparant deux achats (ou utilisation, ou consommation) d'un produit (marque) est influencé par les intervalles inter-achats de ses substituts.

* Si deux marques i et j sont des *"substituts parfaits"* :

$E(X_{ii}) = E(X_{jj}) = E(X_{ij}) = E(X_{ji})$ (1)

$E(X_{ii})$ est l'espérance de X_{ii} sur l'ensemble de la période de temps considérée.

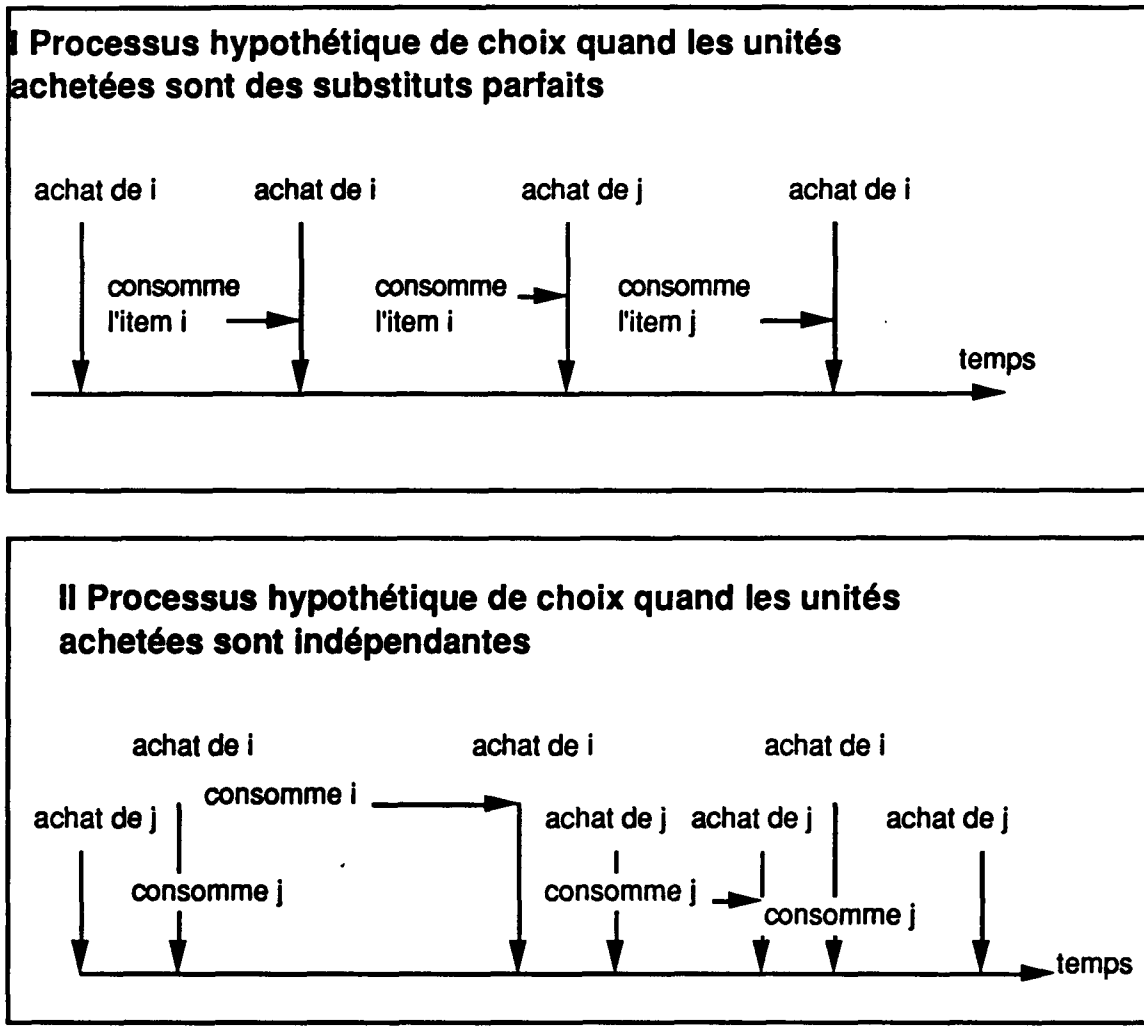
$E(X_{ii}) - E(X_{ij})$ (2) mesure la non substituabilité entre i et j.

* Si deux marques i et j sont des *"non substituts parfaits"*, leurs occasions d'achat sont indépendantes et l'on observe alors (voir Figure 17 :

$E(X_{ii}) > E(X_{ij})$

et $E(X_{jj}) > E(X_{ji})$ (3)

FIGURE 17 DEUX EXEMPLES DE PERIODES INTER-ACHATS



"Adapté de Fraser et Bradford 1983, p17"

Précisons que :

- sur les marchés où le choix entre substituts est fait au moment de l'achat, la période inter-achat est utilisée comme une mesure de la substituabilité .
- sur les marchés où le choix est fait au moment de l'occasion d'usage, c'est la période inter-consommation qui doit alors être prise en compte.

Pour avoir un index appartenant à (O,1), il faut tenir compte :

- du degré de substitution : $E(X_{ii}) - E(X_{ij})$
- du taux d'usage moyen des marques i et j
- du degré de régularité dans le processus d'achat de ces marques

En effet, ces taux varient d'un ménage à l'autre ou (et) d'un produit à l'autre, il est donc nécessaire de standardiser l'indice.

Les périodes inter-achats sont supposées suivre une loi gamma :

- d'espérance : $E(X_{ii}) = r_i/\alpha_i$
- de variance : $V(X_{ii}) = r_i/\alpha_i^{\,2}$

où :

α_i reflète le taux d'usage moyen du produit. C'est un paramètre d'échelle.

r_i reflète la régularité dans le processus d'achat (consommation). C'est un paramètre de forme, plus le taux d'usage est variable et plus r est petit (en général r est compris entre 1 et 2).

•Si i et j sont des *substituts parfaits :*

$$E(X_{ii}) - E(X_{ij}) = O \qquad (4)$$

•Si i et j sont *parfaitement indépendants :*

$$E(X_{iii} - E(X_{ij}) = \frac{(r_i - 1)}{2(\alpha_i + \alpha_j)} \qquad (4a)$$

•Si i et j sont des *substituts imparfaits*, on a un mélange de (4) et (4a), soit :

$$E(X_{ii}) - E(X_{ij}) = \frac{(1 - p)(r_i - 1)}{2(\alpha_i + \alpha_j)} \qquad (5)$$

Pour obtenir une valeur standardisée de $E(X_{ii}) - E(X_{ij})$ il est donc nécessaire :

- de contrôler l'influence de taux d'usage différents en divisant par $E(X_{ii})$.
- de contrôler l'influence de la régularité d'achat (usage) en divisant par la valeur de (4a), cas de l'indépendance parfaite.

Sij s'écrit alors :

$$S_{ij} = \frac{[E(X_{ii}) - E(X_{ij})] / E(X_{ij})}{(r_i - 1) / 2.r_i} \qquad (6)$$

- $S_{ij} = 0$, si i et j sont substituts parfaits
- $S_{ij} = 1$, si i et j sont parfaitement indépendants
- $S_{ij} = (1-p)$, si i et j sont substituts imparfaits

Afin d'obtenir un indice symétrique entre i et j, les auteurs proposent de calculer la moyenne des sij et sji pondérée par w, coefficient donnant plus de poids au processus le plus régulier.

$$w = \frac{r_i - 1}{r_i + r_j - 2} \qquad (7)$$

ou bien $w = 0,5$ de façon arbitraire

Finalement, le sens de la relation est changé pour obtenir un indice de similarité, aisément interprétable (égal à 1 dans le cas de substituts parfaits).

$$S_{ij} = S_{ji} = 1 - \frac{w.[E(X_{ij})-E(X_{ii})] + (1-w).[E(X_{jj})-E(X_{ji})]}{w\,[(r_i-1)/2r_i] + (1-w).[E(r_j-1)/2r_j]}$$

ESTIMATION DE L'INDICE IRS

Soit T le nombre d'individus ayant utilisé les deux marques i et j sur la période d'étude. Pour le ménage t, ayant effectué N_t achats, on calcule les moyennes empiriques inter-achats :

$$\bar{X}_{jjt} = \bar{X}_{ijt} = \bar{X}_{jit} = \bar{X}_{iit} = \frac{1}{N_t} \cdot \sum_{l=1}^{N_t} X_{kk'l} \qquad (8) \qquad (k,k'=i,j)$$

avec X_{iil}, la période inter-achat pour la marque i, à l'achat l.

On calcule ensuite les variances empiriques inter-achats, σ_{iit}^2, σ_{jjt}^2, σ_{ijt}^2 et σ_{jit}^2 :

$$\sigma_{jjt}^2 = \sigma_{ijt}^2 = \sigma_{jit}^2 = \sigma_{iit}^2 = \sum_{k=1}^{N_t} (X_{kk'l} - \bar{X}_{kk'l})^2 \qquad (9)$$

La moyenne de ces valeurs sur les T ménages permet d'estimer les périodes inter-achats au niveau de l'échantillon tout entier.

$$\overline{XX}_{ii} = \frac{1}{T} \sum_{t=1}^{T} \bar{X}_{iit} \qquad (10)$$

r_i et r_{ij} sont estimés sur l'échantillon par :

$$r_i = r_j = \frac{1}{T} \sum_{t=1}^{T} [\sigma_{kkt}^2 / X_{kkt}^{-2}] \qquad k = i,j \qquad (11)$$

w est alors estimé en remplaçant r_i et r_j dans (7), suivent alors les estimations de Sij = Sji.

FORCES DE LA METHODE

Elle permet, théoriquement, de prendre en compte (détecter) si, au sein d'un ménage, des produits ont des utilisations différentes et sont donc non concurrents.

Puisque l'index IRS est calculé à partir des intervalles inter-achats et prend en compte les taux de consommation (donc l'usage du produit), Fraser et Bradford assimilent cet index à un indice de mesure basé sur l'usage. Dans ce but, les auteurs ont comparé l'indice IRS avec sept indices prenant en compte à la fois les différentes personnes utilisatrices du produit au sein du ménage et les différentes situations d'usage du produit. Les observations étaient des reports par les consommateurs des utilisations qu'ils font des produits. Les résultats ont montré tout d'abord que deux marques ont plus tendance à être indépendantes de par la diversité des usages qui en est faite qu'à cause de l'hétérogénéité des personnes qui les consomment au sein du ménage. Ceci est un élément en faveur de l'importance de la situation d'usage lors de la mesure de la concurrence.

D'autre part, cet indice capte bien l'effet de substitution dû à l'hétérogénéité des préférences au sein du ménage ce que ne fait pas un indice construit à partir des transferts entre les marques.

FAIBLESSES DE LA METHODE

- Les auteurs (1983, 1984) définissent leur indice comme un indice de "substituabilité révélée par l'usage". Ils rendent l'index IRS symétrique en calculant la moyenne de Sij et Sji. Ceci est à notre avis réducteur et peu conforme avec la notion même de substituabilité qui, comme nous l'avons dit au premier chapitre, n'a aucune raison d'être symétrique.

- Le modèle fonctionne mal avec des produits à faible taux de pénétration, ou des produits dont la fréquence de renouvellement est faible. Pour estimer correctement les périodes inter-achats, les produits doivent être d'achat courant et les consommateurs doivent être des acheteurs fréquents. Les auteurs ont résolu ce point en étudiant des chroniques d'achats très longues ce qui ne nous parait pas très réaliste compte tenu de la rapidité avec laquelle les marchés évoluent.

- Le calcul des périodes inter-achats ne prend en compte que les seules marques i et j. Toute autre marque qui pourrait être substituable à i et j et donc modifier la valeur des périodes inter-achat n'est pas prise en compte. L'avantage de cette approche est qu'elle ne nécessite pas d'avoir effectué au préalable une définition du marché. Cependant, nous nous interrogeons sur la signification de la loi de périodes inter-achats de deux marques sorties de leur contexte concurrentiel.

Le Modèle présenté par Fraser et Bradford est intéressant par sa nouveauté, la particularité de l'indice qu'il utilise. En raison de son "*atypie*" par rapport aux modèles classiques, il a

fait l'objet de nombreuses critiques, notamment dans un article de Shocker, Zahorick et Stewart (1984). Les principales critiques portées par ces auteurs concernent :

1 L'hypothèse de constance des quantités consommées (fréquences d'achat) celles-ci pouvant en effet varier sur la période d'observation, en fonction par exemple de fluctuations saisonnières (par exemple la consommation de bière). Cependant, pour un grand nombre de produits de grande consommation, cette hypothèse nous semble justifiable, d'autant plus que, comme le soulignent Fraser & Bradford (1984) :

- les périodes inter-achats X_{ij} peuvent être pondérées par l'inverse de la quantité achetée de la marque i
- les fréquences inter-achat ne sont bien sûr pas constantes mais suivent une loi gamma, qui elle l'est !

2 Deux biens indépendants peuvent (dans certains cas) avoir les mêmes taux de consommation et donc les mêmes périodes inter-achats. Ceci est vrai mais devrait se produire très rarement si le marché a été correctement défini au préalable. Le même problème se pose d'ailleurs lors de la mesure de la concurrence à partir de données de transferts entre marques, il est possible d'observer des transferts entre des marques indépendantes, notamment si le marché a été mal défini.

3 Shocker et al reprochent à Fraser et Bradford de ne pas prendre en compte la possibilité de différents usages du produit par différents individus au sein du même ménage. Au contraire, il nous semble que l'un des aspects séducteurs du modèle est sa capacité à les prendre en compte (même s'il ne le fait pas de façon explicite). Si deux produits d'un même marché sont utilisés par différents individus pour différents usages, le modèle en tiendra compte car les produits auront des fréquences d'achat indépendantes.

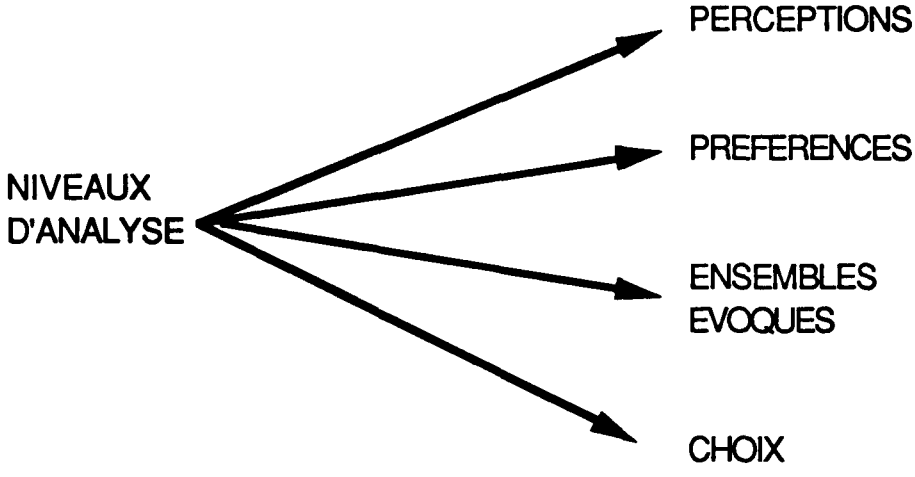
Pour conclure, nous dirons que l'un des principaux avantages de l'indice de Fraser et Bradford sur les méthodes procédant à partir de transferts entre marques est de pouvoir faire la différence (en théorie) entre deux produits achetés pour des usages différents par différents consommateurs (ou le même consommateur) au sein du même ménage. En effet, même si ces usages ne sont pas connus, les produits auront des périodes inter-achats suivant des lois indépendantes et pourront être diagnostiqués comme *"non substituts"*. Cette tentative de prise en compte de la situation d'usage à partir de données de panels mérite d'être soulignée.
Dans la section suivante, nous allons présenter les principales approches de la mesure des relations entre marques utilisant des données de jugement.

SECTION 4 ETUDE DE LA RELATION ENTRE LES MARQUES A PARTIR DE DONNEES DE JUGEMENT DES CONSOMMATEURS

Les années soixante-dix ont vu se multiplier les applications conjointes en Marketing de théories et de techniques provenant de l'étude de la psychologie du consommateur (perceptions, attitudes, processus de traitement de l'information) et celles en provenance de la psychométrie, notamment analyses factorielles, techniques de scaling. Cette double approche de l'étude du consommateur a permis d'obtenir des cartes perceptuelles des marchés sur lesquelles il était possible de représenter la *"proximité perceptuelle"* (et par extension concurrentielle) des marques. Ces approches ont obtenu un grand succès. Nous présenterons, dans cette section, les principales méthodes qui ont été proposées pour analyser les relations entre marques à partir de données de jugement du consommateur.

Pessemier (1978), notait que la meilleure manière de considérer le consommateur est de procéder à partir des trois concepts que sont "Ensembles Evoqués, Perceptions et Préférences". Ce triptyque relatif au type de concept utilisé pour mesurer la relation entre les marques nous parait constituer un découpage tout à fait cohérent pour la présentation de l'étude de la relation entre les marques à partir de données de jugement. Nous rajouterons cependant une autre approche qui procède à partir des "choix des consommateurs", tels qu'ils sont exprimés par le répondant dans des enquêtes. L'approche utilisant les données de préférences (à cause de son caractère spécifique) sera, quant à elle, présentée dans la section 5 suivante consacrée aux modèles de préférences.

FIGURE 18 NIVEAUX D'ANALYSE DE LA RELATION ENTRE MARQUES A PARTIR DE DONNEES DE JUGEMENTS



Nous passerons donc successivement en revue, dans cette section, les méthodes reposant sur les perceptions, les ensembles évoqués de marques et enfin les choix.

SS1 MESURE DE LA RELATION ENTRE LES MARQUES A PARTIR DES PERCEPTIONS

Nous distinguerons ici deux types d'approches qui reposent sur les perceptions des consommateurs :

- une approche qui procède à partir des perceptions des marques, soit en terme de similarité directe, soit en terme de perception des marques selon les attributs déterminants.
- une approche qui procède à partir de l'adéquation perçue entre les produits et les situations d'usage de ces produits.

SS1-1 MESURES PROCEDANT A PARTIR DES SIMILARITES PERCUES

Cette approche a donné lieu dans les années soixante dix à un très grand nombre de recherches et d'applications. Nous citerons Trout et Ries (1987), Wind et Robinson (1972), Assael (1971), Boss (1974) et Boss dans Bouroche (1977), Holmes (1973), Wilkes (1977), Dubois (1977 et 1979), Huber et Holbrook (1979), Pioche (1974), Somia (1974), Keon (1983), Dillon et al (1986). Dans la majorité des cas, ces approches étudient le concept de Positionnement des produits. Comme le note Dubois (1977 p10), le positionnement peut être défini selon un double point de vue :

- un point de vue étroit dans lequel il est défini en tant que *"méthode de diagnostic"* permettant de mieux connaître les perceptions et les attitudes et la position des marques selon ces perceptions et attitudes.
- un point de vue beaucoup plus large selon lequel le positionnement est considéré comme une démarche, une méthode d'action sur le marché et donc finalement une *"stratégie"*.

La littérature marketing s'est en fait beaucoup plus préoccupée du premier point de vue puisque c'est sans doute là que les aspects méthodologiques sont les plus importants. C'est ainsi que le positionnement des produits a souvent été "réduit" à l'obtention de cartes représentant les relations qu'entretiennent les marques dans l'esprit du consommateur.

Dans beaucoup de ces recherches et d'études, le positionnement se limite à étudier l'espace concurrentiel, c'est à dire la structure du marché. Nous avons d'ailleurs noté dans l'introduction générale de cette recherche, que l'analyse de la structure du marché et la segmentation se placent *"en amont"* des études de positionnement.
Selon Wind et Robinson (1972), le positionnement trouve ses origines dans les études faites par les économistes sur la structure des marchés et la substituabilité des produits.
Le concept de base sur lequel reposent ces approches est celui de *"similarité perçue entre les marques"*. Ces similarités sont collectées auprès des répondants (voir Boss dans Bouroche 1977 pour une revue des techniques de recueil et Neidell 1972, Wipple 1976, Henry et Stumpf 1975, Gupta 1984 pour une comparaison des différentes méthodes de

recueil selon un point de vue technique). Une fois agrégées, ces données acquièrent souvent le statut de mesure de la concurrence entre les marques.
Dans le chapitre précédent, nous avons remis en cause cette équivalence faite entre similarité perçue et concurrence entre les marques. De nombreux auteurs ont d'ailleurs critiqué l'assimilation qui est faite entre similarité perçue et concurrence. Nous pouvons citer à cet égard Lehmann (1972), Dubois (1977), Chandon et Strazziéri (1986).
Deux approches sont traditionnellement opposées pour procéder à la mesure et à la représentation des "distances perceptuelles" entre les marques :

- les méthodes d'analyse des similarités indirectes qui procèdent à partir de perceptions selon les attributs déterminants du produit, attributs qui sont proposés au répondant. La similarité est alors reconstituée sur les cartes des marques obtenues à partir d'analyses factorielles ou discriminantes.
- les méthodes d'analyse des similarités directes dans lesquelles les similarités perçues entre les marques.sont collectées directement

L'utilisation de ces deux types de données ont fait l'objet d'un débat qui repose sur le contenu de la mesure. Les chercheurs ou praticiens reprochent généralement aux similarités indirectes d'induire les réponses en proposant des attributs qui ne sont pas forcement ceux qu'utilise le répondant dans son processus perceptuel. Quant aux similarités directes, il leur est reproché, par les mêmes chercheurs ou praticiens, la lourdeur liée à leur collecte et la difficulté d'interprétation des cartes qu'elles permettent d'obtenir (voir Hauser et Koppelman 1979 pour un exemple de ce débat).
Nous ne nous étendrons pas d'avantage sur cette approche de la mesure de la relation entre les marques, celle-ci ayant déjà été largement diffusée et utilisée dans la recherche marketing.

SS1-2 UTILISATION DES ADEQUATIONS PERCUES ENTRE PRODUITS ET USAGE POUR L'ETUDE DES RELATIONS ENTRE LES MARQUES

Les travaux de Stefflre (1971, voir aussi dans Myers et Tauber 1977) qui, à l'origine, étaient conçus pour développer une méthodologie de création de concept nouveau, ont donné naissance à un nouveau courant de recherche en matière d'analyse de la structure des marchés. Dans cette approche, la relation entre les produits est exprimée à partir des adéquations perçues entre les produits et les situations d'usage de ces produits.
Ce nouvel axe de recherche a donné lieu à une série d'articles qui sont le fait d'un nombre restreint de chercheurs (voir Srivastava, Shocker et Day 1978, Day, Shocker et Srivastava 1979, Srivastava 1980 et 1981, Srivastava, Leone et Shocker 1981, Srivastava, Alpert et Shocker 1984, Shocker et al 1987). Curieusement, cette approche de la définition des marchés n'a suscité que peu d'intérêt de la part des chercheurs. Nous attribuons ce faible degré d'intérêt au fait que cette approche demande peu de développements statistiques, contrairement aux modèles de préférences utilisant la modélisation Logit, qui suscitent un nombre toujours croissant de recherches. Cette approche demeure pourtant très riche et prometteuse. Les auteurs cités précédemment

rappellent qu'il est important, avant toute analyse de la structure d'un marché, de choisir un point de vue qui peut être :

- un point de vue *"tactique"* accompagné d'une vision étroite du marché (actions à court terme). Dans ce cas là, les mesures de similarité traditionnelles sont suffisantes car l'analyse est menée sur un marché suffisamment homogène, avec des produits utilisables dans des situations d'usage équivalentes, pour que la similarité ait valeur de substituabilité.
- un point de vue *"stratégique"*, où la définition du marché est plus large. L'entreprise s'intéresse alors alors aux possibilités d'extension du marché dues aux lancements de nouveaux produits, aux modifications des besoins, à l'évolution de l'environnement et donc aux frontières et structures potentielles et futures du marché. Dans ce cas, il est indispensable de prendre en compte de façon explicite la variable situation d'usage lors de la mesure de la concurrence puisque cette variable explique alors une part importante de la concurrence entre les produits.

L'utilisation de la situation d'usage pour établir la substituabilité des produits s'appuie sur les constatations de Belk (1975) suivant lequel les situations d'usage sont des facteurs "écologiques" qui facilitent au consommateur la résolution de ses problèmes de choix en posant des contraintes sur l'obtention de ces solutions.

L'hypothèse de base de cette approche est que le produit n'est pas désiré pour lui même, mais pour l'ensemble des bénéfices qu'il apporte au consommateur. De la même manière, le consommateur peut être défini à partir des bénéfices qu'il désire. Selon cette perspective, produits et individus ont des caractéristiques qui sont fonction de leur environnement et notamment de la situation d'usage et du contexte de consommation. Les produits seront considérés et classés en fonction de la perception de leur caractère approprié aux situations d'usage envisagées par le consommateur.

Srivastava, Alpert et Shocker (1984), définissent un marché comme *"un ensemble de produits jugés substituables pour des situations d'usage dans lesquelles les mêmes patterns de bénéfices sont attendus par un groupe de consommateurs"*. Cette définition du marché équivaut à utiliser la *"similarité des profils d'adéquation des produits aux situations d'usage"* comme mesure de la concurrence entre les produits (ou éventuellement les marques).

De la même manière, un marché est constitué de sous-marchés qui se comportent très différemment. Ces sous-marchés doivent être définis par rapport aux situations d'usage. Or, puisqu'un même produit peut satisfaire les différents bénéfices réclamés par diverses situations d'usage, il est naturel de penser qu'un même produit pourra appartenir, à des degrés divers, à différents sous-marchés. L'utilisation d'un modèle de typologie de recouvrement permet de prendre en compte ce phénomène (voir chapitre III, voir Shepard et Arabie 1979, Arabie, Caroll, DeSarbo et Wind 1981).

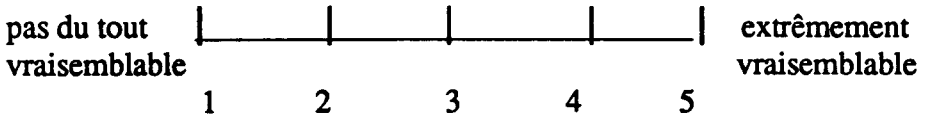
Srivastava et al (1984) proposent une méthodologie en trois phases pour définir et analyser la structure d'un marché. Day, Shocker et Srivastava dans leur article de 1981 proposent une méthodologie assez similaire bien que plus lourde à mettre en oeuvre et moins élégante, nous ne la présenterons donc pas ici.

1- Définition des situations d'usages

La procédure utilisée est analogue à celles proposées par Strefflre (1971 et 1979), Belk(1979), Day, Shocker et Srivastava (1979) et exposée au chapitre I section 5.

2- Génération de la matrice Consommateurs x Produits x Usages

Les situations d'usage obtenues dans la phase 1 sont décrites au consommateur qui doit alors évaluer les produits présentés sur des échelles de mesure de la vraisemblance de l'usage du produit aux situations. Par exemple :



Les auteurs conseillent de dédoubler l'échantillon afin de présenter à chaque demi-échantillon une description différente des diverses situations d'usages . Ceci permet d'effectuer la validation croisée des résultats obtenus sur les deux demi-échantillons.

3- Analyse de la structure du marché

Les auteurs proposent d'étudier en détail les effets principaux et les interactions liées aux trois variables que sont la Personne, l'Objet et la Situation. Ceci est effectué en utilisant l'analyse de variance à trois facteurs, la note de vraisemblance de l'usage étant la variable à expliquer.

Dans une application sur les services bancaires, les auteurs ont montré que la Situation d'usage a une bonne capacité de prédiction des préférences et qu'il existe une interaction significative entre le produit et la situation.

Les indices de substitution dans l'usage sont alors analysés avec un programme de typologie hiérarchique ou de recouvrement, les auteurs ayant montré la supériorité de ce dernier pour représenter les structures de la concurrence sur un marché. Chaque groupe ainsi caractérisé peut être considéré comme un sous marché dans lequel les produits sont en forte concurrence. Ces groupes peuvent être définis à partir des situations d'usage qui les caractérisent. Pour cela, il suffit de calculer et de comparer les vraisemblances moyennes des produits dans chaque groupe et ce, pour les différentes situations.

Outre l'analyse de la structure du marché telle qu'elle peut être perçue par les consommateurs (nous intéressons ici, rappelons le, à la "*réalité consommateur*" et non à la "*réalité du marché*"), cette approche permet d'observer :

- quelles sont les variables d'environnement les plus liées à un groupe de produits
- quelles sont les situations les plus mal desservies, ce qui peut aboutir à l'amélioration de la gamme d'un produit en desservant au mieux les différents usages sans pour autant créer de cannibalisme entre les produits. Si une entreprise intègre cette analyse au sein d'une stratégie de développement ou de diversification cela peut amener la firme à créer des produits nouveaux.
- quelles sont les situations où les produits sont perçus comme bien ou mal adaptés et ainsi effectuer le diagnostic et le suivi du positionnement des produits de la firme par rapport à la concurrence.

Malgré tous ses avantages, l'analyse de la structure des marchés à partir des substituabilités dans l'usage présente quelques limites qu'il est bon de souligner.

1 L'obtention des ensembles de situations d'usage est difficile. En particulier il est difficile d'obtenir des situations d'usage qui se situent à un *"même degré de généralité"*, degré qui de plus doit être pertinent par rapport au problème posé. Deux recherches que nous avons menées, l'une sur le riz, l'autre sur le marché de la lingerie féminine, ont montré que ces conditions sont particulièrement difficiles à vérifier.

2 La méthodologie est difficile à mettre en oeuvre lorsque l'univers de concurrence est défini de manière étroite ou, ce qui n'est pas exceptionnel, lorsque le produit n'a qu'une seule situation d'usage. Dans ces conditions, la mesure proposée par les auteurs devient inopérante. Elle est par contre beaucoup mieux adaptée pour étudier la substituabilité entre différents produits plutôt qu'entre différentes marques.
En fait, cette méthode d'analyse de la concurrence nous semble plus complémentaire qu'opposable aux autres approches.

3 Les matrices *"situations d'usage x produits"* sont en général de taille importante. Avec vingt situations et vingt produits, il y a déjà quatre cents évaluations de l'adéquation du produit à la situation à demander au répondant. Les auteurs mesurent cette adéquation à l'aide d'une échelle en cinq points. Nous pensons qu'il est plus économique et tout aussi pertinent de collecter les réponses sous une forme binaire du type adapté/non adapté. Ainsi collectées, les réponses, une fois agrégées, constituent une matrice de confusion dont le terme général nij représente le nombre de répondants ayant perçu une adéquation entre le produit i et la situation j. Une telle matrice se traite très naturellement à l'aide de l'analyse factorielle des correspondances, ce qui permet d'obtenir des cartes où figurent :

- la substituabilité entre les produits en terme -d'adéquation aux situations d'usage.
- la proximité entre les situations d'usage
- l'adéquation entre les produits et les situations d'usage.

4 Cette mesure de substituabilité accorde autant de poids au fait que deux produits (marques) soient simultanément adéquat, qu'au fait d'être simultanément non-adéquat à une situation donnée. Deux produits qui ne sont adaptés qu'à une même situation d'usage seront ainsi évalués comme étant aussi substituables que deux produits qui sont perçus comme adéquats à un grand nombre de situations d'usage.

Ceci étant, cette mesure reste très intéressante puisqu'elle mesure une réelle substituabilité entre les produits. De plus elle permet d'effectuer à la fois la définition du marché et la mesure de la substituabilité entre les marques, ce que ne font pas la majorité des autres méthodes proposées dans la littérature.

SS2 MESURE DE LA RELATION ENTRE LES MARQUES A PARTIR DES ENSEMBLES EVOQUES

Chandon et Strazziéri (1986) ont proposé un indice de mesure de la concurrence calculé à partir des fréquences de considérations simultanées de deux alternatives. Le consommateur en situation d'achat est supposé ne considérer qu'un nombre restreint de marques parmi celles qu'il connait. Cet ensemble est restreint du fait de la limite des

capacités cognitives de l'individu, ce qui le conduit à utiliser des stratégies de sélection des stimuli lors de son processus de choix. En interrogeant l'individu sur les marques qu'il considère comme un choix possible pour une situation d'usage donnée, il est possible de construire un indice de substituabilité entre les marques.
Il est possible de faire un parallèle entre ce stade du processus du choix et les modèles de choix à deux étapes dont nous avons parlé à la section 3 du chapitre I. L'ensemble évoqué, tel qu'il est défini par les auteurs, correspond bien au sous-ensemble de marques qui subsiste une fois que le consommateur a mis en oeuvre la première phase de son processus de choix à deux étapes. Ce parallèle permet de replacer l'approche de Chandon et Strazziéri dans le cadre des modèles de choix comme étant située à une étape intermédiaire, donc assez proche du comportement.
Soient :

a, le nombre d'individus considérant la marque A dans leur ensemble évoqué
b, le nombre d'individus considérant la marque B dans leur ensemble évoqué
n, le nombre total d'individus dans l'échantillon
tAB , le nombre théorique de considérations simultanées des marques A et B, attendu sous hypothèse d'indépendance des considérations (absence de concurrence entre les marques A et B).

$$t_{AB} = \frac{a \cdot b}{n}$$

t_{AB} représente le nombre de considérations simultanées attendues sous hypothèse du simple hasard.

S_{AB} est le nombre empiriquement observé de considérations simultanées des marques A et B.

On peut calculer l'écart entre les considérations simultanées observée et théoriques, soit :

$$e_{AB} = S_{AB} - t_{AB}$$

si $e_{AB} > 0$, il y a alors *"sur-concurrence"* entre deux marques.
si $e_{AB} < 0$, il y a alors *"sous-concurrence"* entre les deux marques.

L'intensité de la concurrence entre les deux marques peut être mesurée par la probabilité d'obtenir, au hasard, moins de considérations simultanées que le nombre observé. Si cette probabilité est forte et tend vers 1, cela signifie que les deux marques sont en situation de sur-concurrence. Si elle est faible (proche de 0), il y a au contraire sous-concurrence. Enfin, si elle est proche de 0.5, il y a indépendance entre les deux marques.
S_{AB} est la réalisation d'une variable aléatoire X dont les valeurs sont données par la loi hypergéométrique. L'intensité de concurrence est opérationnalisée par :

$$P(X < S_{AB}) = \sum_{x} P (X = x / a,b,n) \text{ , pour un seuil de risque } \alpha \text{ donné}$$

$P(X < S_{AB})$ est la probabilité d'obtenir au hasard, un nombre de considérations plus faible que le nombre X observé.
$x \in [\ \max(0,a+b-n),\ S_{AB}\ ]$

La loi hypergéométrique pouvant être approximée, les calculs de ces probabilités sont directs. La matrice symétrique des intensités de concurrence est ensuite traitée par une technique de MDS classique afin d'obtenir une représentation graphique des relations de concurrence et de la structuration du marché en zones de concurrence (sous-marchés). Ces champs de concurrence sont des sous-ensembles de marques en sur-concurrence entre elles et en sous-concurrence avec les autres. Il peut y avoir recouvrement entre les champs, ce qui est une hypothèse très réaliste. Du seuil de certitude α choisi dépendra la forme de ces champs de concurrence. Plus le risque est faible, plus les alternatives doivent être en sur-concurrence pour appartenir au même champ. Tout aussi simplement, la méthode permet d'évaluer les champs de concurrence, en calculant le *"volume net"* de considérations des marques y appartenant.

Faciles à collecter, ces données permettent de se *"rapprocher"* du comportement. Puisqu'il s'agit des marques *"considérées"*, leur nombre est faible (comme l'a montré Brisoux 1982, cité dans Chandon et al 1988) lorsque le répondant se trouve en situation de choix. Ceci permet au répondant, comme le notent les auteurs, d'exprimer sa prédisposition réelle à l'achat. De plus, les auteurs ont proposé de *"contrôler"* la variable situation d'usage en collectant les ensembles évoqués pour une situation d'usage bien précise qui est spécifiée lors de la collecte des ensembles évoqués. En effet, cet indice ne conserve sa valeur que si l'individu est correctement placé en situation de choix et non dans une situation artificielle. Nous retiendrons les points suivants:

- la signification de l'indépendance entre deux marques nous parait équivoque. Que signifie l'indépendance par rapport à l'absence absolue de concurrence ?
- lorsque la taille de l'échantillon augmente, le nombre de considérations simultanées attendues augmente. Les probabilités d'obtenir un nombre plus petit sous hypothèse du simple hasard tendent alors vers un ce qui rend la mesure inopérante, car peu sensible aux variations de la concurrence. Ce problème peut être facilement résolu en retenant, par exemple, le rapport S_{AB}/t_{AB} comme indice de mesure de la concurrence. Si ce rapport est supérieur à l'unité, cela signifie que les deux marques sont en situation de sur-concurrence.
- grâce au terme de comparaison t_{AB}, l'indice prend en compte l'effet de saillance des marques (qui sont citées souvent parce qu'elles viennent plus facilement à l'esprit).
- cet indice est, au plan individuel, une mesure de substituabilité soit parce que les deux marques sont similaires et apportent les mêmes bénéfices, soit parce que les deux marques sont différentes mais apportent des combinaisons de bénéfices qui sont évaluées à l'identique relativement à la situation d'usage spécifiée.

Pour conclure, nous soulignerons que cet indice est intéressant pour les raisons suivantes :

- il s'intègre dans la théorie du processus de traitement et de choix de l'individu
- il constitue une mesure de substituabilité agrégée
- il est facile à calculer et permet de prendre en compte un grand nombre de marques

- il est proche du comportement du consommateur (ce qui est intéressant si l'analyse porte sur à la concurrence) sans être pour autant influencé par les variables de situation d'achat.

Il constitue donc une voie prometteuse pour la mesure de la substituabilité entre les marques.

SS3 MESURE DE LA RELATION ENTRE LES MARQUES A PARTIR DES CHOIX DES REPONDANTS: LE MODELE PRODEGY (Urban, Johnson et Hauser 1984)

Les auteurs procèdent à partir de leur définition d'un sous-marché qui est la suivante : *"un produit appartient à un sous-marché s'il tire ses ventes, lors de son lancement, des consommateurs de ce sous-marché et non d'un autre"*. Cette définition est reformulée de façon à être rendue opérationnelle dans le modèle : *"Un marché est découpé en sous-marchés quand un produit éliminé d'un sous-marché voit ses anciens consommateurs reporter leur choix au sein du même sous-marché et ce dans une proportion supérieure à ce que ne laisserait espérer la simple proportionnalité aux parts de marché des produits qui y appartiennent"*. Dans ces conditions, le produit éliminé est considéré comme substituable aux autres produits de ce sous-marché. Il s'agit donc de tester des modèles hypothétiques (correspondant à des structures de marché hypothétiques), par comparaison à une hypothèse nulle Ho : l'absence de structure.

MODELE AGREGE DE RATIO CONSTANT (ACRM : aggregate constant ratio model)
Soient :

i et j deux marques appartenant au sous-marché S du marché total T constitué de p sous-marché, p = 1.à.P
$P_i(j)$, la part de marché de j quand la marque i est éliminée de S
$P_j(S)$, la part de marché du sous-marché S par rapport au marché total après élimination de i

Soient les valeurs empiriques suivantes, observées sur un échantillon de taille N :

n_i le nombre d'individus préférant (choisissant) la marque i
$n_i(j)$, le nombre d'individus préférant (choisissant) la marque i et choisissant une autre marque, j, appartenant au même sous-marché, lorsque i n'est plus disponible

En l'absence de structure, donc sous H0, on peut écrire :

$$P_i(j) = \frac{m_j}{\sum_{\substack{k \in S \\ k \neq i}} m_k} = \frac{m_j}{1 - m_i} \qquad (2)$$

$$P_i(S) = \sum_{\substack{j \in S \\ j \neq i}} P_i(j) = \frac{\sum_{j \in S} m_j}{1 - m_i} \qquad (3)$$

m_i, m_j, m_k sont les parts de marché de i, j, k sur le marché.
$P_i(j)$ et $P_i(S)$ sont des probabilités théoriques sous H0 (absence de structure).

(2) et (3) signifient qu'une marque j gagne proportionnellement à sa part de marché lorsqu'une marque i appartenant à son sous-marché S est supprimée. Notons que cette hypothèse revient à poser l'hypothèse d'indépendance des alternatives non-pertinentes (IIA) au sein de chaque sous-marché hypothétique et à la rejeter lorsque les marques appartiennent à deux sous-marchés différents (c'est aussi ce qui est fait dans le modèle de Hendry).

Soit $\tilde{N}i$ la probabilité $P_i(S)$ estimée sur l'échantillon :

$$\tilde{N}_i(S) = \frac{n_i(S)}{n_i} \qquad (4)$$

$n_i(S)$: nombre de choix se reportant sur le sous-marché S quand i, marque préférée, n'est plus disponible

L'hypothèse d'existence d'une structure particulière peut être validée si $\tilde{N}i(j)$ et $\tilde{N}i(S)$ sont significativement différents des valeurs théoriques correspondantes sous H0, c'est à dire quand :

$\tilde{N}_i(j)$ est différent de Pi(j)
$\tilde{N}_i(S)$ est différent de Pi(S)

PRISE EN COMPTE DE LA SITUATION D'USAGE

Le modèle permet de prendre en compte les situations d'usage en testant l'inégalité:

$\tilde{N}iu(u)$ différent de Piu(u)
u est une situation d'usage particulière.
$\tilde{N}iu(u)$ est l'estimation du nombre d'individus se reportant sur d'autres marques satisfaisant à la situation d'usage u quand leur marque préférée, i, n'est plus disponible.
Piu(u) est la part de marché des marques satisfaisant à la situation d'usage u quand la marque i est retirée.

Ceci permet de découvrir une structure définie par l'usage si chaque ensemble de produits tend à être utilisé pour un usage particulier. Deux méthodes permettent de traiter le problème :

1 - Collecter les données pour chaque usage et suivre la procédure précédente afin d'identifier une structure liée à chaque usage. On obtiendra alors autant de structures que d'usages, tous les produits figurant dans chaque structure correspondant à un (et un seul) usage particulier.

2 - Une seule structure est déterminée dans laquelle chaque produit est assigné à un groupe (sous-marché) caractérisé par la situation d'usage pour laquelle les produits de ce sous-marché sont les plus adéquats. Il existe alors une relation bi-univoque entre sous-marché et situation d'usage.

Soient :

- n_{iu}, le nombre de consommateurs utilisant le produit i pour l'usage u. La règle consiste à assigner le produit i au sous-marché qui maximise niu .
- $n_{iu}(j)$ le nombre de consommateurs qui choisissent le produit i pour l'usage u et qui choisissent j quand i est supprimé.

$$n_{iu}(u) = \sum_{j \in u} n_{iu}(j)$$

$n_{iu}(u)$ est le nombre de consommateurs qui choisissent un produit appartenant au sous marché défini par l'usage u, quand leur marque préférée, i, est éliminée.

$P_{iu}(u)$ est la probabilité d'effectuer son choix au sein du sous-marché défini par la situation d'usage u quand i, la marque préférée, est supprimée.

Si Ho est vraie, alors $\tilde{N}_{iu}(u)$ est peu différent de $P_{iu}(u)$.

$\tilde{N}_{iu}(u)$ est estimé par $n_{iu}(u)/n_{iu}$

et : $$P_{iu}(u) = \sum_{\substack{j \in u \\ j \neq i}} m_j/(1-m_i)$$

Si le sous-marché existe, on doit alors vérifier:

$\tilde{N}_{iu}(u) > P_{iu}(u)$ si i appartient à u.

$\tilde{N}_{iu}(u) < P_{iu}(u)$ si i n'appartient pas à u.

CAS DE STRUCTURES COMPOSITES

C'est le cas (le plus complexe) où différents groupes d'individus se caractérisent par différentes structures de marché. Soit la marque i appartenant au sous-marché S. On suppose dans un premier temps qu'un ensemble de structures vraisemblables ont été identifiées. Deux approches du problème sont alors possibles :

1 Assigner les individus à la structure qui maximise $P_{ci}(S)$ pour le sous marché contenant i, $P_{ci}(S)$ et S variant selon les différentes structures de concurrence hypothétiques. $P_{ci}(S)$ est la probabilité, pour le consommateur c, de choisir une marque du sous-marché S quand la marque i de S n'est pas disponible.

$$P_{ci}(S) = \sum_{\substack{i \in S \\ i \neq j}} \frac{P'_c(j)}{(1 - P'_c(j))}$$

$P_C'(j)$ est la probabilité pour le consommateur c de choisir la marque j.

2 Comparer les probabilités $P_{ci}(S)$ et $P_i(S)$ et assigner ensuite le consommateur c à la structure pour laquelle ($P_{ci}(S)$-$P_i(S)$) est maximum. $P_i(S)$ est la part de marché de la marque i au sein du sous-marché S.

TYPES DE DONNEES UTILISEES

Pour alimenter le modèle, il faut collecter les données permettant de calculer les probabilités conditionnelles de passage de la marque préférée vers une autre marque.
Les auteurs proposent quatre approches :

1 - Choix forcés

Le répondant est mis dans une situation de choix où le produit qu'il a choisi n'est pas disponible (après qu'il ait exprimé ce choix).

2 - Rangs de préférences

ni(j) est calculé à partir du nombre d'individus ayant placé les marques i et j respectivement aux premier et deuxième rangs. Il faut alors faire l'hypothèse que les préférences reflètent correctement un comportement hypothétique.

3 Intensités de préférence (estimées dans un modèle Logit).

Dans ce cas, on utilise des données de préférences à somme constante comme variables explicatives dans un modèle Logit permettant d'estimer les probabilités ACRM : $P'_C(j)$, la variable expliquée étant le "choix de la marque". Il est alors possible de calculer:

$$P_{ci}(j) = \frac{P'_C(j)}{[1 - P'_C(i)]} \qquad (10)$$

$P'_C(j) = 0$ si j n'appartient pas à E_C (l'ensemble de considérations du consommateur c)
$P'_C(j) = 1$ dans le cas inverse.

$$n_i(j) = \sum_c P_{ci}(j)\, P'_C(i) \qquad (11)$$

$$ni = \sum_c P'_C(i) \qquad (12)$$

4 Ensembles de considérations (ensemble de marques que le consommateur envisage sérieusement lors de son choix)

Soit E_C, l'ensemble de considérations du consommateur c, c a la même probabilité $P'_C(i) = 1/e_C$ (13) d'acheter chaque marque de son ensemble de considérations (e_C est le nombre de marques considérées par le répondant c). On remplace alors (13) dans (10),(11),(12).

Dans son esprit, le modèle PRODEGY est plus adapté pour fonctionner avec des données de jugement, il reste cependant utilisable avec des données de comportement, ce qui constitue un atout intéressant et une possibilité d'opérer des comparaisons.

AVANTAGES DE LA METHODE

Le principe du test découle d'une définition très naturelle de la notion de sous-marché. Il permet de tester des hypothèses formulées par des hommes de terrain et qui ont donc une certaine *"validité faciale"*.
Cette approche permet en outre de prendre en compte les situations d'usage mais aussi le cas des structures composites (bien que ce point ne soit pas assez développé par les auteurs et que la procédure paraisse lourde à mettre en oeuvre).

INCONVENIENTS DE LA METHODE

La solution de tester des hypothèses formulées au préalable nécessite de bien connaître le marché et de faire *"la"* bonne hypothèse. L'avantage lié à cet inconvénient est d'avoir une approche permettant de faire le pont entre le jugement subjectif des managers et la rigueur statistique d'un modèle. Le risque est par contre de caractériser différentes structures aussi explicatives les unes que les autres (par rapport à H0). Les auteurs ne fournissent pas de réelle solution à ce problème autrement que le choix de la structure la plus vraisemblable. La méthode (comme toutes les approches confirmatoires) peut céder une large part à l'arbitraire dans la mesure où seules les structures en lesquelles on croit seront testées et le responsable pourra considérer qu'elles sont vraies à partir du moment où il aura montré qu'elles ne sont pas fausses.

DEVELOPPEMENTS RECENTS DU MODELE PRODEGY

Novack et Stangor (1987) ont proposé une méthode d'estimation du modèle en utilisant la technique des moindres carrés pondérés. Cette méthode a l'avantage de fournir une mesure de l'adéquation des données à l'hypothèse de structure du marché qui est fixée a priori. Ceci permet ensuite d'effectuer des comparaisons statistiques concernant l'adéquation des structures, en utilisant des tests du Chi2.

Récemment, Ogawa et al (1989) ont proposé une application pratique très intéressante du modèle Prodegy à la détermination des assortiments dans les distributeurs automatiques de boissons.

TABLEAUX RECAPITULATIFS DES DIFFERENTES APPROCHES UTILISANT DES DONNEES DE COMPORTEMENT OU DE JUGEMENT

Avant d'aborder la section suivante, nous allons présenter sous forme de tableaux récapitulatifs, l'ensemble des approches que nous avons étudiées dans cette section et les précédentes. Les deux premiers tableaux (10A et 10B) regroupent toutes les approches procédant à partir de données de comportement. Dans le premier tableau (10A), nous présentons toutes les approches qui reposent sur le concept de transferts entre marques. Dans le tableau 10B nous avons regroupé les autres approches qui utilisent des données de comportement. Dans le tableau 10C, nous regroupons les modèles utilisant des données de jugement des consommateurs.
Dans ces tableaux, les approches sont décrites à partir de trois blocs de variables :

- les variables relatives à la mesure dans le modèle
- les variables relatives à la structure du modèle
- les variables précisant les particularités du modèle

DESCRIPTION DU BLOC "MESURE"

Y figurent les éléments suivants :

- le niveau de l'analyse: agrégé ou individuel
- le type de comportement de choix observé: une marque parmi N (1/N) ou un nombre quelconque de marques parmi N (K/N)
- le niveau d'analyse qui convient le mieux au modèle: marque ou produit. Ce niveau d'analyse étant en grande partie arbitraire, nous nous attacherons simplement à donner le niveau d'analyse qui convient le plus naturellement au modèle.

DESCRIPTION DU BLOC "STRUCTURE DU MODELE"

On trouve dans ce bloc les éléments suivants :

- le type de l'analyse: Confirmatoire ou Exploratoire
- le prise en compte ou non des variables du mix marketing
- le type d'attributs qui servent à construire les partitions. Selon les modèles, ces attributs pourront être physique, perçus ou de nature non-précisée par les auteurs. De plus, ces attributs peuvent être qualifiés de "discret" ou de "continu".
- le fait que le découpage du marché en sous-marchés soit conçu en termes de partitions exclusives ou bien de partitions de recouvrement (partitions floues).
- le fait qu'il existe ou non une théorie du comportement du consommateur sous-jacente au modèle proposé.
- le fait que le modèle nécessite ou non une définition préalable du marché.

DESCRIPTION DU BLOC "CARACTERISTIQUES PARTICULIERES DU MODELE"
Figurent ici, successivement :

- le fait que le modèle prenne en compte, à l'aide d'une loi statistique sur les paramètres individuels ou d'un découpage en segments, l'hétérogénéité des consommateurs.
- le fait que le modèle prenne en compte ou non les situations d'usage du produit et si oui, si cette prise en compte est explicite ou implicite (voir chapitre I section 1).
- le fait que le modèle prenne en compte ou non la recherche de variété du consommateur.
- le fait que le modèle prenne en compte la non-symétrie de la relation entre les marques.
- la facilité avec laquelle les paramètres du modèles sont estimés, sur une échelle qui va de ++ (très facile) à -- (très complexe et très coûteux).

Nous étudierons dans la section suivante les mesures de la relation entre les marques qui utilisent des modèles de préférences.
Nous avons choisi de les présenter à part, dans une section, pour les deux raisons suivantes :

- il s'agit, comme nous le verrons, d'un type particulier de mesure de la concurrence
- les modèles sont de plus en plus estimés en utilisant alternativement des données de jugement ou de comportement.

TABLEAU 10A MODELES D'ANALYSE ET DE MESURE DE LA CONCURRENCE (DONNEES DE COMPORTEMENTS)
MODELES UTILISANT DES DONNEES DE TRANSFERTS ENTRE MARQUES

AUTEURS	INDIVIDUEL AGREGE	TYPE DE CHOIX	NIVEAU D'ANALYSE	CONFIRMATOIRE/ EXPLORATOIRE	VARIABLE DU MIX	TYPE D'ATTRIBUTS	TYPE DE SOUS-MARCHE	THEORIE DU CHOIX	DEFINITION DU MARCHE	METRO DU CONSOM.	SITUATION D'USAGE	RECHERCHE DE VARIETE	NON SYMETRIE	ESTIMATION
HENRY 1970	AGREGE	L/N	MARQUE	CONFIRMATOIRE	NON	PHYSIQUE DISCRET	PARTITION	MODELE STOCHASTIQUE DE CHOIX	OUI	NON PRIS EN COMPTE	NON	NON	NON	-
LEHMANN 1972	AGREGE	L/N	MARQUE	EXPLORATOIRE	NON	INDETERMINE CONTINU	NON PRECISE	ABSENCE	OUI	NON PRIS EN COMPTE	NON	NON	NON	++
RAO & SABALAVA 1981	AGREGE	L/N	MARQUE	EXPLORATOIRE	NON	PHYSIQUE DISCRET	PARTITION HIERARCHIQUE	CHOIX INDIVIDUEL HIERARCHIQUE	OUI	NON PRIS EN COMPTE	NON	NON	NON	++
LEVINE 1979	AGREGE	L/N	MARQUE	EXPLORATOIRE	NON	INDETERMINE	NON PRECISE	ABSENCE	OUI	NON PRIS EN COMPTE	NON	NON	NON	++
McALISTER & LATTIN 1981	INDIVIDUEL	L/N	MARQUE	EXPLORATOIRE	NON	PERCU CONTINU	NON PRECISE	RECHERCHE DE VARIETE	OUI	NON PRIS EN COMPTE	NON	OUI	OUI	-
GEVON 1985	INDIVIDUEL puis AGREGE	L/N	MARQUE	CONFIRMATOIRE	NON	PHYSIQUE DISCRET	PARTITION	RECHERCHE DE VARIETE	OUI	NON PRIS EN COMPTE	NON	OUI	NON	-
CARPENTER & LEHMANN 1985	AGREGE	L/N	MARQUE	EXPLORATOIRE	OUI	PHYSIQUE DISCRET	NON PRECISE	THEORIE DE L'UTILITE	OUI	NON PRIS EN COMPTE	NON	NON	OUI	-
GROVER & RILLON 1986	INDIVIDU/ SEGMENT	L/N	MARQUE	CONFIRMATOIRE	NON	PHYSIQUE DISCRET	PARTITION FLOUE	MODELE STOCHASTIQUE	OUI	OUI PAR SEGMENT	NON	NON	NON	.
GROVER & SRINIVASAN 1987	INDIVIDU/ SEGMENT	L/N	MARQUE	CONFIRMATOIRE	NON	PHYSIQUE DISCRET	PARTITION FLOUE	MODELE STOCHASTIQUE	OUI	OUI PAR SEGMENT	NON	NON	NON	-
TOTTEN & BLOOP 1988	AGREGE	L/N	MARQUE	EXPLORATOIRE	NON	INDEFINI	NON PRECISE	ABSENCE	NON	NON PRIS EN COMPTE	NON	NON	NON	++
KUMAR & SARCAR 1989	AGREGE	L/N	MARQUE	CONFIRMATOIRE	NON	PHYSIQUE DISCRET	PARTITION	MODELE STOCHASTIQUE	OUI	NON PRIS EN COMPTE	NON	NON	NON	-
JAIN, BASS & CHEN 1990	INDIVIDU/ SEGMENT	L/N	MARQUE	CONFIRMATOIRE	NON	PHYSIQUE DISCRET	PARTITION FLOUE	MODELE STOCHASTIQUE	OUI	OUI AVEC LOI STAT.	NON	NON	NON	.

TABLEAU 10B AUTRES APPROCHES UTILISANT DES DONNEES DE COMPORTEMENT D'ACHAT

AUTEURS	INDIVIDUEL AGREGE	TYPE DE CHOIX	NIVEAU D'ANALYSE	CONCEPT DE BASE	CONFIRMATOIRE/ EXPLORATOIRE	VARIABLE DU MIX	TYPE D'ATTRIBUTS	TYPE DE SOUS-MARCHE	THEORIE DU CHOIX	DEFINITION DU MARCHE	HETERO. DU CONSOM.	SITUATION D'USAGE	RECHERCHE DE VARIETE	NON-SYMETRIE	ESTIMATION
MERUNKA & BOURGEAT 1988	INDIVIDUEL puis AGREGE	E/N	MARQUE	QUTE ACHETEE	EXPLORATOIRE	NON	INDETERMINE	NON PRECISE	ABSENCE	OUI	NON	NON	NON	NON	++
FRASER & BRADFORD 1983	INDIVIDUEL puis AGREGE	L/N	MARQUE	PERIODE INTER-ACHATS	EXPLORATOIRE	NON	INDETERMINE	NON PRECISE	MODELE STOCHASTIQUE	NON	OUI	OUI IMPLICITE	NON	NON	+
VANHONACKER 1984	AGREGE	E/N	MARQUE	ELASTICITE	EXPLORATOIRE	OUI	INDETERMINE	NON PRECISE	ABSENCE	NON	NON	NON	NON	OUI	=
COOPER 1988	AGREGE	E/N	MARQUE	ELASTICITE	EXPLORATOIRE	OUI	INDETERMINE	NON PRECISE	UTILITE STOCHASTIQUE	OUI	NON	NON	NON	OUI	=
KAMAKURA & RUSSELL 1989	INDIVIDUEL	E/N	MARQUE	ELASTICITE	EXPLORATOIRE	OUI	PHYSIQUE DISCRET	NON PRECISE	UTILITE STOCHASTIQUE	OUI	OUI	NON	NON	NON	=
ALLENBY 1988	INDIVIDUEL OU AGREGE	E/N	MARQUE	ELASTICITE	EXPLORATOIRE	OUI	PHYSIQUE DISCRET	NON PRECISE	UTILITE STOCHASTIQUE	OUI	NON	NON	NON	NON	=
SHUGAN 1987	AGREGE	E/N	MARQUE	ELASTICITE	EXPLORATOIRE	OUI	ATTRIBUT PAR DOLLAR	-	THEORIE MICRO-ECO.	OUI	NON	NON	NON	NON	-

TABLEAU 10C MODELES D'ANALYSE ET DE MESURE DE LA SUBSTITUABILITE (DONNEES DE JUGEMENT)

AUTEURS	INDIVIDUEL AGREGE	NATURE DES DONNEES	NIVEAU D'ANALYSE	CONFIRMER/ EXPLORER	VARIABLE DU MIX	TYPE D'ATTRIBUT	THEORIE	SITUATION D'USAGE	DEFINITION DU MARCHE	RECHERCHE DE VARIETE	ESTIMATION
APPROCHE DECOMPOSITIONNELLE	AGREGE	SIMILARITE DIRECTE	MARQUE OU PRODUIT	EXPLORER	NON	PERCU	PERCEPTION (COOMBS 1964)	NON	NON	NON	++
APPROCHE COMPOSITIONNELLE	AGREGE	PERCEPTION MARQUE x ATTRIBUT	MARQUE OU PRODUIT	EXPLORER	NON	PERCU	MODELE MULTI-ATTRIBUTS	NON	NON	NON	++
PRODEGY: URBAN & AL 1984	AGREGE	CHOIX FORCE	MARQUE	CONFIRMER	NON	PHYSIQUE OU PERCU	MODELE STOCHASTIQUE DE CHOIX	POSSIBLE	NON	NON	+
SRIVASTAVA, SHOCKER & AL 1978	AGREGE	ADEQUATION PRODUIT USAGE	PRODUIT	EXPLORER	NON	BENEFICE LIE AUX SITUATIONS D'USAGE	SITUATION D'USAGE	OUI, EXPLICITEMENT	OUI	NON	++
CHANDON & STRAZZIERI 1986	AGREGE	ENSEMBLE EVOQUE	MARQUE	EXPLORER	NON	PERCU	ENSEMBLE EVOQUE	OUI, PAR CONTROLE	OUI	NON	++
BUCKLIN & SRINIVASAN 1989	INDIVIDU	PERCEPTION MARQUE x ATTRIBUT	MARQUE	EXPLORER	PRIX	PERCU	THEORIE DE L'UTILITE STOCHASTIQUE	OUI	NON	NON	=

SECTION 5 MODELES DE PREFERENCES ET MESURE DE LA RELATION ENTRE LES MARQUES

Dans cette section, nous allons présenter les approches de l'étude de la concurrence à partir de modèles de préférences. Bien que toute analyse des préférences permette d'obtenir des cartes représentant une structure de relations entre marques, il est nécessaire de bien différencier cette approche des précédentes. La particularité provient du fait qu'elle utilise un *concept intermédiaire* pour étudier la relation entre deux marques: le *point idéal* (ou vecteur idéal) du consommateur (ou d'un segment particulier de consommateurs ou même de la population toute entière). L'utilisation de ce concept intermédiaire rejoint ce que nous avons noté au chapitre I, à la section 4 : la similarité entre deux marques n'est ni une condition nécessaire ni une condition suffisante de substituabilité. Deux marques peuvent être concurrentes sans être perçues comme similaires, par contre elles doivent être jugées pertinentes dans le contexte d'un achat, donc proches du point idéal.

Les modèles de préférences trouvent leur origine dans les travaux de Coombs (1964) et son modèle de dépliage selon lequel la préférence pour une marque est fonction de sa distance au point idéal de l'individu. Connaissant les préférences des consommateurs, on peut les replacer sur une carte de manière à ce que les distances de ces marques aux points idéaux des consommateurs respectent au mieux les préférences déclarées des consommateurs. C'est ainsi que deux marques proches d'un même point idéal peuvent être considérées comme substituables (ou concurrentes) alors qu'elles ne sont pas similaires. Le point idéal est donc, dans notre contexte, le concept intermédiaire permettant d'établir le degré de substituabilité (concurrence) entre deux marques.
La taxonomie classique en matière de modèles de préférences conduit à opposer (voir Bouroche 1977 et 1978):

- Modèles Internes / Externes
- Modèles à Point Idéal / Vectoriel
- Modèle utilisant des données métriques ou non métriques

Nous pensons pour notre part que le statut des modèles de préférences a considérablement évolué dans la dernière décennie et qu'il est important à l'heure actuelle d'opposer les modèles utilisant des données de jugement aux modèles utilisant des données comportements. En effet, alors que les modèles de préférences (comme leur nom l'indique) ont été conçus pour analyser des données de préférences, leur évolution récente montre qu'ils sont de plus en plus nombreux à être alimentés à partir de données de comportement (on parle de préférences révélées), ou mieux, à partir de données de préférences transformées en données de choix (voir par exemple Chapman et Staelin 1982). Nous présenterons les modèles selon cette double optique.

SS1 MODELES DE PREFERENCES PROCEDANT A PARTIR DE DONNEES DE JUGEMENT

Comme nous l'avons noté, ce type de modélisation trouve ses origines dans les travaux de Coombs (1964) avec son modèle de dépliage. Ce modèle a été généralisé et

rendu opérationnel par les travaux des statisticiens des "Bells Laboratories". L'ouvrage de Shepard, Romey et Nerlove (1972) fait une présentation complète d'un point de vue historique et technique de toutes ces méthodes.
Nous ne nous attarderons pas à décrire ces modèles qui sont maintenant largement connus et diffusés et renvoyons le lecteur à l'ouvrage précédent pour un exposé très clair et de grande qualité, ainsi qu'à l'ouvrage de Bouroche (1978) et à l'article de Shocker et Srinivasan (1979) pour une revue critique.
D'une manière générale, la préférence d'un individu i pour une marque j peut s'exprimer de deux manières :

1 Par un point idéal

On écrit alors :

$$P_{ij} = a + b \left(\sum_{t=1}^{T} w_t.(x_{jt} - I_{it})^2 \right) + e_i$$

- x_{jt} représente les coordonnées de la marque j sur la dimension t
- w_t est le poids de la dimension t dans l'évaluation par le consommateur de la marque j
- e_i est le terme d'erreur aléatoire

2 Par un vecteur de préférence idéal

La préférence pour le produit est alors exprimée par sa projection sur le vecteur idéal de l'individu :

$$P_{ij} = a + b.\left(\sum_{t=1}^{T} w_t . V_t.x_{it} \right)$$

- V_t représente la coordonnée du vecteur idéal selon la dimension t
- les autres paramètres ont la même signification que précédemment

La nature des attributs dans ces deux types de modèles est différente. Dans le modèle à point idéal, le consommateur utilise des attributs selon lesquels il peut manifester de la *satiété*. Il existe un niveau optimal au delà duquel la préférence diminue. Dans le modèle à vecteur idéal les attributs sont au contraire du type *"au plus au mieux"*.

Ces modèles de préférences ont été largement utilisés. Notamment, certains modèles de prévision de "produit nouveau" les utilisent. On peut ainsi citer le modèle Perceptor (voir Urban 1977), le modèle Linmap (voir Shocker et Srinivasan 1974). Tous ces modèles procèdent à partir de l'hypothèse selon laquelle la préférence d'un individu (ou segment d'individu) pour une marque, est une fonction inverse de la distance du point idéal de cet individu à cette marque. Les distances peuvent être ensuite décomposées dans un espace à T dimensions spécifiées à l'avance. La signification psychologique de ces dimensions diffère fondamentalement selon que le responsable marketing ou le chercheur utilise des modèles dits externes ou internes.

Dans les modèles externes, les coordonnées des marques sont considérées comme connue, elle a en général été estimée à partir de données de similarité perceptuelle. Les

points idéaux des consommateurs sont placés par rapport à ces marques. La signification des dimensions extraites est donc identique à celles qui sont obtenues dans les analyses de similarités que nous avons présentées à la section précédente. Ce sont les dimensions qui permettent au répondant d'effectuer la meilleure discrimination entre les marques. Les modèles externes reposent ainsi sur une hypothèse forte: "les perceptions des consommateurs sont homogènes et c'est par une pondération différente des dimensions qu'ils manifestent des préférences hétérogènes".

Dans les modèles internes, les points (ou vecteurs) idéaux et la location des marques sont estimés à partir des "seules" données de préférences. Les dimensions de l'espace ont alors une tout autre signification. Elles sont caractéristiques (voir Myers et Tauber 1977, p 19) de ce que le consommateur considère comme important lors de l'évaluation ou du choix des marques. Il n'y a bien sûr aucune raison pour que ces deux types de dimensions (internes et externes) soient identiques. Les dimensions perceptuelles les plus importantes ne sont pas forcément les dimensions saillantes utilisées pour opérer un choix. Les unes servent à discriminer, les autres à formaliser une préférence.

Tout ceci nous amène à considérer que les modèles externes, du point de vue de l'étude de la relation entre les marques, ne sont rien d'autre que des méthodes d'analyse des similarités. Seules les analyses internes utilisent le concept "intermédiaire" de point idéal pour établir la position des marques. Ce sont donc ces modèles internes qui seront considérés ici. D'ailleurs, Myers et Tauber (1977) préconisent d'utiliser les modèles internes lorsque l'objectif consiste en l'étude de la concurrence.

Les modèles de préférences ont, selon nous, évolué selon deux logiques. La première a consisté à introduire une dimension probabiliste dans les modèles de préférences classiques. La deuxième a consisté à introduire les modèles d'utilité (qui étaient principalement utilisés dans les modèles de choix) au sein des modèles d'analyse des préférences.

SS1-1 LES MODELES DE PREFERENCES PROBABILISTES

A l'origine de cette modélisation on trouve un ensemble de travaux dont l'objectif était de conférer une dimension probabiliste aux techniques de scaling en général et notamment au modèle d'analyse des similarités (MDS). A ce sujet, on peut citer les travaux de Ramsey avec son modèle Multiscal (1977, voir aussi dans Davies et Coxon 1981) et de MacKay et Zinnes 1982, (modèle Proscal). Takane (1981) a présenté un modèle probabiliste utilisable sur des données non-métriques.
MacKay et Zinnes ont proposé (1986) un modèle de préférences probabiliste utilisant des données de préférences exprimées sous forme de ratios de préférences. L'intérêt d'une telle approche est qu'elle permet (selon les auteurs) de prendre en compte, selon les cas, deux éléments :

- l'erreur de mesure au plan individuel, puisque la perception est un concept non observable, il faut admettre qu'elle est mesurée avec de l'erreurs.
- l'hétérogénéité des préférences, quand les analyses sont menées au plan agrégé.

Les auteurs ont ainsi présenté un modèle à point idéal dans lequel la préférence pour une marque j relativement à une marque k, pour un consommateur i est représentée par le ratio : d_{ik}/d_{ij} .

$$d^2_{ij} = \sum_{l=1}^{r} (x_{il} - x_{jl})^2$$

d_{ij} est la distance de la marque j au point idéal du consommateur i.
l=1 à r représente les dimensions de l'espace
x_{il} est la coordonnée empirique du point idéal de i sur la dimension l
x_{jl} est la coordonnée empirique de la marque j sur la dimension l

La variance de la distance du point idéal du consommateur i à la marque j, σ_{ij} s'exprime comme :

$$\sigma^2_{ij} = \sigma^2_i + \sigma^2_j$$

σ_i^2 est la variance liée au point idéal du consommateur i
σ_j^2 est la variance liée à la marque j

La loi de la distance d'une marque j au point idéal du consommateur i est probabilisée.
Le rapport $Z^2_{ij} = d^2_{ij} / \sigma_{ij}$ est distribuée suivant un chi2 non centré à r degrés de liberté, de paramètre π_{ij}.
$\pi_{ij} = D^2_{ij} / \sigma_{ij}$
D_{ij} représente la distance vraie (non observable) entre le point idéal i et la marque j

$$D^2_{ij} = \sum_{l=1}^{r} (\mu_{il} - \mu_{jl})^2$$

μ_{il} est la coordonnée "vraie" du point idéal i sur la dimension l
μ_{jl} est la coordonnée "vraie" de la marque j sur la dimension l

L'expression de la fonction de densité de Zij permet alors d'estimer tous les paramètres par la méthode du maximum de vraisemblance, en fonction des préférences exprimées par les répondants.

Ce modèle permet d'obtenir, dans un espace de préférences, la représentation des marques, représentation qui, comme l'ont montré les auteurs (1986), est plus fidèle que ce qui est obtenu à partir de méthodes déterministes En particulier, ce modèle permet de prendre en compte l'erreur de mesure au plan individuel et surtout, l'hétérogénéité des préférences, lors de l'estimation des points idéaux au niveau du segment. Ceci permet alors d'exprimer le point idéal d'un segment de consommateur comme l'espérance de la

loi dont on observe les différentes réalisations avec les individus qui appartiennent à ce segment.

SS1-2 UTILISATION DES MODELES D'UTILITE DANS LES MODELES DE PREFERENCES

Chapman et Staelin (1982) ont ouvert la voie à de nouvelles recherches en présentant un modèle d'utilité qui utilise les données de rangs de préférences. Ces données de rangs de préférences sont décomposées de manière à obtenir des données de choix qui sont ensuite introduites dans un modèle Logit d'utilité.

Cooper et Nakanishi (1983) ont présenté un modèle de préférences externe qui repose sur l'hypothèse de Coombs selon laquelle les perceptions sont homogènes mais les préférences varient (Coombs 1964). Ce modèle permet de représenter les points idéaux des individus dans un espace de marques. Nous avons noté dans l'introduction de cette section que les modèles de préférences interne ne présentent pas d'intérêt direct dans le cadre de cette section. Nous exposons cependant ce modèle car il présente un intérêt historique dans l'évolution des modèles de préférences. Dans ce modèle, la préférence relative entre deux marques j et k, pour un segment de consommateurs i est opérationnalisée à partir des fréquences relatives de choix entre paires de marques.

F_{ijk} est la fréquence relative du choix de la marque j par rapport à la marque k par le segment de consommateurs i. De même F_{ijk}^* est l'espérance de cette valeur sur le segment de consommateurs. $F_{ijk}^* + F_{ikj}^* = 1$

Les intérêts de cette approche sont multiples:

1 Elle permet de se replacer dans le cadre de la théorie de l'utilité et notamment d'utiliser la modélisation Logit. F_{ijk}^* s'exprime alors comme :

$$F_{ijk}^* = \frac{\exp(D_{ij}).\varepsilon_{ij}}{\exp(D_{ij}).\varepsilon_{ij} + \exp(D_{ik}).\varepsilon_{ik}}$$

D_{ij} est la distance "théorique" entre le point idéal i et la marque j.

ε_{ij} est un terme d'erreur relatif au segment de consommateurs i concernant la marque j. Il suit une loi log-normale dont l'intérêt est d'exprimer que la préférence ne varie plus beaucoup lorsqu'une marque est éloignée du point idéal. Cette erreur est exprimée comme dans le modèle MNL de McFadden. Les ε_{ij} suivent des lois identiques et indépendantes ($cov(\varepsilon_{ij},\varepsilon_{ik}) = 0$ pour tout j et k). Ceci signifie que le modèle repose sur l'hypothèse des IIA.

Ce modèle a été aussi présenté par les auteurs sous la forme d'un modèle vectoriel. Le choix de la forme du modèle dépend alors de la nature des attributs. Si ces derniers sont de type *"au plus au mieux"*, il est alors préférable d'utiliser la forme vectorielle.

2 Cette modélisation des préférences est aisément transposable dans le cadre de la théorie de l'utilité et des modèles probabilistes de choix. F_{ijk} représente alors la part de choix du consommateur (ou segment) i relativement aux marques j et k. F_{ijk}^* mesure la

relation de préférence exprimée par i concernant les marques j et k. Au plan agrégé, cette mesure peut être considérée comme l'expression de la relation concurrentielle entre les deux marques.

3 Cette approche, à l'inverse des approches perceptuelles, caractérise la non-symétrie de la relation de concurrence puisque F_{ijk}^* est bien sûr différent de F_{ikj}^* .

Les modèles que nous venons de présenter ont permis de créer un pont entre deux visions du comportement du consommateur qui étaient auparavant indépendantes, voire même s'affrontaient : les approches déterministes et les approches probabilistes.

Ces travaux sont l'expression de la fusion qui, à notre avis, s'opère à l'heure actuelle entre les modèles de préférences déterministes et les modèles de choix stochastiques. De cette fusion on a pu voir naître la généralisation des modèles qui utilisent les hypothèses générales des travaux de Coombs en les intégrant au sein d'une théorie de l'utilité de type probabiliste et qui utilisent pour cela des données de comportement de choix. Ceci nous amène maintenant à présenter les modèles de préférences estimés à partir de données de choix réels.

SS2 MODELES DE PREFERENCES UTILISANT DES DONNEES DE CHOIX

Avec la disponibilité croissante des données de comportement de choix, un nombre croissant de modèles de préférences conçus pour fonctionner avec ce type de données a vu le jour. Comme nous l'avons déjà noté, ce type de modèles est le résultat de la conjonction de deux courants de recherche.

• 1 Le premier courant a pour origine la théorie de l'utilité et les modèles de choix qui ont été développés en Marketing. Selon cette théorie que nous avons présentée à la section 3 du chapitre I, l'utilité d'une alternative j : v_j, est composée d'une part déterministe (u_j) et d'une part aléatoire (ε_j).

$$v_j = u_j + \varepsilon_j$$

La préférence Pj pour la marque j peut être exprimée selon la forme du modèle de Luce:

$$P_j = \frac{v_j}{\sum_{i \in S} v_i}$$

S est l'ensemble des considérations de choix

Ceci permet d'exprimer la probabilité de choix de la marque j:

$$\text{Prob}[X=j] = p_i = \exp[v_j] / \sum_i \exp[v_i] \quad (1)$$

Par rapport à un ensemble d'alternatives S, le consommateur choisit la marque (ou l'ensemble de marques) qui maximise son utilité. En général, la distribution de εj est une exponentielle double négative. Avec cette loi, lorsque les valeurs de l'utilité sont extrêmes, les probabilités de choix sont peu sensibles aux variations de l'utilité.
D'autre part, en reprenant la forme du modèle de Lancaster (voir Lancaster 1966), la part déterministe de l'utilité s'exprime comme une fonction des caractéristiques objectives de la marque.

$$v_j = A'j.W \quad (2)$$

A_j est le vecteur *"technologique"* (r,1) décrivant les caractéristiques objectives de la marque j selon r attributs.
W est un vecteur de poids des attributs physiques dans la détermination de l'utilité totale.

En intégrant (2) dans (1) on obtient la forme générale des modèles de choix.
Une des premières applications de ce type de modèle en marketing est à celle qui a été proposée par Batsell et Lodish (1981) qui ont modélisé les parts de choix de marques de Soft Drink à partir de leurs attributs physiques.

2 Le deuxième courant est issu de la recherche en psychologie mathématique. Ce courant possède deux apports essentiels.
D'une part les modèles multi-attributs selon lesquels les attitudes sont le résultat d'un processus de traitement de l'information de type compositionnel. Le modèle le plus connu étant celui de Fishbein (voir chapitre I section 3).

$$A_j = \Sigma\, a_{jr}.w_j \quad (3)$$

A_j: attitude par rapport à la marque j
a_{jr}: niveau de croyance relatif à la possession par la marque j de l'attribut r
w_j: poids de l'attribut r dans la formation de l'attitude (attrait de cette caractéristique)

D'autre part, le modèle de Coombs (voir plus haut) qui exprime la préférence pour une marque comme étant une fonction de l'inverse de la distance de cette marque au point idéal.

$$P_j = a + b.\sum_r w_r.(a_{jr} - I_r)^2 \quad (4)$$

w_r est le poids de la dimension r dans la formation de la préférence.
a_{jr} est la position de la marque j sur la dimension r
I_r est la position du point idéal sur la dimension r

Depuis le début des années 1980, notamment avec le développement des modèles logit et probit (voir Maddala 1983, Gourrieroux 1984, Amemiya 1981, Hausman 1980, ainsi que les articles de Doyle 1977 et Malhotra 1984 pour les applications en Marketing), les

approches intégrant la forme du modèle d'utilité (1) avec les formes (2) ou (3) ou (4) se sont multipliées.
On obtient alors des modèles permettant de localiser des marques et éventuellement des points idéaux, dans des espaces d'attributs physiques ou perçus. Ces modèles sont en fait des modèles explicatifs des parts de choix des marques, parts de choix qui proviennent de plus en plus de données de panels. Les modèles sont alors utilisés non seulement pour expliquer et prévoir les choix des marques (ce qui correspond aux formes (1) + (2) ou (1) + (3)) mais aussi localiser les positions des marques et des points idéaux des consommateurs (ce qui correspond à la forme (1) + (4)). Dans cette forme (1+4), les modèles ont la particularité d'inférer des "cartes perceptuelles" à partir de la mesure de choix réels, ce qui reste tout à fait défendable dans la mesure ou le comportement est le résultat d'un processus perceptuel de traitement de l'information.
L'extrême souplesse de ce cadre général d'analyse et la qualité des données disponibles, expliquent sans doute la multiplication des approches proposées tout au long des dix dernières années. En effet, à partir de ces différentes modélisation, il est possible de considérer alternativement que:

- les utilités sont individuelles ou agrégées,
- les poids des axes sont idiosyncratiques ou globaux,
- les positions des marques sont données (modèle externe) ou à estimer dans le modèle (modèle interne),
- les préférences (telles qu'elles se manifestent au travers des choix) peuvent être exprimées sous la forme de vecteurs ou de points idéaux,
- les erreurs aléatoires (eij) peuvent être indépendantes (modèle MNL qui repose sur l'hypothèse IIA) ou liées (modèle probit généralisé).
- le choix peut être unique (une marque parmi n) ou multiple (un nombre k, fixé ou quelconque parmi n). Nous remarquerons à ce niveau que le modèle proposé par DeSarbo et Hoffman (1987) ou DeSarbo et Cho (1989) est un des rares à avoir modélisé les choix multiples.
- l'agrégation sur les individus peut être effectuée en utilisant une distribution de probabilité, qui peut porter sur la position du point idéal (voir Kamakura et Srivastava 1986), sur les poids des attributs des choix (comme c'est le cas dans Choice Map de Elrod,1988), ou plus généralement sur toute combinaison de ces éléments (voir MacKay et Zinnes 1986). Ce type d'agrégation permet de prendre en compte l'hétérogénéité des consommateurs.
- les attributs du choix peuvent être indépendants ou liés, c'est une des possibilités du modèle Proscal de MacKay et Zinnes.

Comme nous pouvons le constater, toutes ces possibilités offrent un cadre pratiquement infini de modélisation. Le principal obstacle auquel se heurtent ces modèles est l'augmentation considérable du nombre de paramètres à estimer au fur et à mesure que les modèles reposent sur des hypothèses moins contraignantes. Un exemple frappant est celui de l'hypothèse des IIA qui ne tient plus lorsque les alternatives présentent des niveaux de substituabilité différents. Pour lever cette hypothèse, certains auteurs ont proposé des modèles qui nécessitent un tel nombre de paramètres à estimer que le nombre de marques étudiées simultanément se réduit alors à trois dans l'exemple proposé ! Y a- t-il encore intérêt à lever l'hypothèse si l'univers concurrentiel est autant réduit?

Nous ne pouvons finalement que constater la convergence entre les modèles probabilistes de choix et les modèles de préférences qui ont été probabilisés. Certains modèles de choix utilisent des préférences, comme le modèle proposé par Chapman et Staelin (1982) qui utilise des données de rangs de préférences dans le modèle de choix de Luce pour effectuer des prévisions de choix. L'avantage est qu'avec ce type de données il est possible de prévoir des parts de choix qui ne sont pas influencées par les variables de situation d'achat, ce qui revient à construire des modèles de choix "toutes conditions de ventes égales par ailleurs". Ceci est un atout intéressant, notamment lorsqu'une l'entreprise souhaite effectuer des prévisions pour un produit nouveau.
Inversement, d'autres approches permettent d'inférer des cartes de préférences à partir de données de choix réel, ce qui constitue une mesure intéressante de la concurrence entre les marques, dans laquelle il parait normal de prendre en considération les variables de situation de vente, puisque l'entreprise souhaite pouvoir répondre à la question "par qui suis-je réellement concurrencé ?".

Nous présenterons maintenant certains modèles qui nous paraissent les plus représentatifs et nous reportons le lecteur au tableau 8 pour un panorama plus large des modèles de préférences et de choix. Nous présenterons le modèle de DeSarbo et Hoffman qui a la particularité de pouvoir modéliser des choix multiples. Contrairement au modèle Logit qui contraint la somme des probabilités pi de choix des marques à être égale à l'unité, dans ce modèle, la somme des pi peut dépasser l'unité. Ceci laisse la possibilité au consommateur d'effectuer plusieurs choix. Cette hypothèse nous parait particulièrement réaliste sur des marchés comme le Yaourt ou les Soft Drink. Nous présenterons le modèle de Kamakura et Srivastava dont la caractéristique principale est qu'il permet de lever l'hypothèse des IIA. Enfin, nous présenterons le modèle Choice Map qui nous semble très prometteur car il offre beaucoup de souplesse.

SS2-1 MODELISATION DES CHOIX MULTIPLES

DeSarbo et Hoffman (1987) ont présenté un modèle probabiliste de préférences, calibré sur des données binaires qui peuvent correspondre à trois types de choix :

- le choix d'une marque parmi n, à chaque occasion d'achat
- le choix d'un nombre k de marques parmi n, à chaque occasion d'achat
- le choix d'un nombre quelconque de marques parmi n, à chaque occasion d'achat

Ce modèle présente deux volets A et B :

- le volet A permet d'obtenir une représentation simultanée des individus (ou segments) et des marques
- le volet B permet de reparamétrer des informations de provenance externe sur les cartes, notamment :
  - les caractéristiques des individus
  - les attributs physiques du produit
  - un "concept de produit" dont l'analyste connait les caractéristiques

Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

PRESENTATION DU MODELE

Le modèle repose sur l'hypothèse classique du modèle de dépliage de Coombs (1964) dans lequel l'utilité du produit pour un consommateur (variable latente non observable) est représentée comme une fonction inverse de la distance du point idéal de ce consommateur au produit. On définit ainsi la variable latente de *"désutilité"* Dij :

$$D_{ij} = \sum_{t=1}^{T} w_{it}.(a_{it} - b_{jt})^2 + c_i + \partial_{ij} \quad (1)$$

D_{ij} est la désutilité du produit j pour l'individu i
T est le nombre de dimensions "saillantes" prises en compte dans le choix
w_{it} est la saillance de la dimension perceptuelle t chez l'individu i
a_{it} est la t ième coordonnée du point idéal reflétant les attentes du consommateur i
b_{jt} est la t ième coordonnée du produit j dans l'espace commun des perceptions
c_i est une constante additive spécifique au consommateur i
∂_{ij} est un terme d'erreur aléatoire, qui procure au modèle sa dimension probabiliste. La distance D_{ij} du produit j au consommateur i devient ainsi un indicateur de la probabilité de choix de ce produit par le consommateur.

L'originalité du modèle est d'exprimer que pour que le produit j soit choisi par un consommateur, il faut que sa désutilité soit inférieure à un seuil "d_i*", spécifique à ce consommateur. Une telle vision du comportement de choix permet de prendre en compte le cas des choix multiples où toutes les marques dont la désutilité ∂ij est inférieure au seuil di* sont choisies par le consommateur.
La probabilité de choix de la marque j par le consommateur i peut alors s'écrire:

$$P(y_{ij} = 1) = P(D_{ij} \leq d_i^*) \quad (2)$$

(1) et (2) permettent d'écrire la probabilité de choix d'une marque j par un consommateur i comme étant la probabilité pour que la valeur du niveau de désutilité de cette marque soit inférieure au seuil admis par cet individu.

$$P(Y_{ij}=1) = P(\partial_{ij} < -\sum w_{it}.(a_{it}-b_{it})^2 - C_i^*)$$

$C_i^* = c_i - d_i^*$ exprime la sensibilité du consommateur i dans son processus de choix. Plus cette valeur est forte et moins le consommateur est sensible dans l'expression de ses préférences.
$Y_{ij} = 1$ si j est choisie par i
$Y_{ij} = 0$ autrement
Y_{ij} suit une loi binomiale ß(1,p_{ij}) où p_{ij} est la probabilité que la marque j soit choisie par le consommateur i sur la période d'observation .
Y_{ij} doit respecter deux hypothèses :

- l'indépendance du choix entre les consommateurs, c'est une hypothèse réaliste
- l'indépendance du choix entre les produits, qui est une hypothèse nettement moins réaliste (elle correspond à faire l'hypothèse classique des IIA relative au modèle de choix de Luce)

En posant: $f_{ij} = -\sum wit.(a_{it}-b_{it})^2 - C_i^*$, on peut alors écrire :

$$P(Y_{ij}=1) = 1 / (1 + \exp^{fij}) = P_{ij}$$

$$P(Y_{ij}=0) = \exp^{fij} / (1 + \exp^{fij}) = 1 - P_{ij}$$

En estimant a_{it}, w_{it}, C_i^*, b_{jt}, on obtient une représentation spatiale des marques et des points idéaux. Cette représentation peut être interprétée en termes de processus de choix à l'aide des poids wit mais aussi en terme d'espace concurrentiel des marques étudiées. En effet, le volet B du modèle permet de représenter des variables "attribut du produit" après avoir reparamétré les coordonnées des produits comme fonction linéaire de ces variables "attribut" par l'équation $B = X.\Gamma$.

B matrice des coordonnées des produits
X matrice des coordonnées des "attributs objectifs du produit"
Γ matrice des liens entre les caractéristiques physiques des produits et la nature des dimensions de l'espace produit.

Le même type de reparamétrage peut être effectué à partir des caractéristiques objectives des individus.
Cette modélisation de la concurrence mérite les commentaires suivants :

1 Nous l'avons déjà souligné, le modèle permet de prendre en compte des choix de type multiple. C'est un des rares modèles proposé en marketing qui permette de prendre en compte de tels choix. Nous noterons cependant que le modèle de Wierenga (1983) permet lui aussi de prendre en compte ce type de choix, mais dans un contexte de choix plus général.

2 C'est une approche souple dans la mesure où de l'information externe au modèle de base.peut être intégrée. Pouvoir intégrer de l'information concernant les caractéristiques objectives des produits ou des consommateurs est tout à fait adapté à l'information disponible dans les panels de consommateurs qui comportent, en plus des historiques d'achats des consommateurs, de l'information objective concernant les produits vendus en magasin et les panélistes (données socio-démographiques). Cependant, il faut noter qu'un tel reparamétrage n'est pas spécifique au modèle, il a déjà été proposé dans les approches classiques de modèles de préférences sur données de jugement sous l'appellation de "fixation de vecteurs de propriété" .

3 C'est une approche individuelle, ce qui comporte deux inconvénients majeurs dans le cadre de l'étude de la concurrence. D'abord, il est difficile d'étudier beaucoup de consommateurs et donc obtenir des cartes fiables (représentatives) du marché. Ensuite, le

nombre de paramètres à estimer est très important. L'utilisation des cartes du marché devient alors trompeuse.

Il est a noter que DeSarbo et Cho (1989) ont proposé une nouvelle version de ce modèle. Dans sa conception générale le modèle reste inchangé. La méthode d'estimation a par contre évolué, il s'agit du modèle Probit non linéaire.

SS2-2 MODELE DE PREFERENCES NE FAISANT PAS L'HYPOTHESE DES IIA

L'obtention de modèles d'analyse de la concurrence ne reposant pas sur l'hypothèse des IIA fait l'objet d'un enjeu statistique qui a donné lieu à plusieurs types de solutions (voir chapitre I section 3). Nous ne présenterons ici que les approches s'inscrivant dans le cadre des modèles de préférences.

Currim (1982), a proposé un modèle de choix permettant de se passer de l'hypothèse des IIA. C'est un modèle Probit généralisé.

Kamakura et Srivastava (1984) ont proposé une version du modèle de choix Probit multivarié. Par rapport au modèle de Currim, qui nécessite l'estimation des paramètres de la matrice de covariance des erreurs εij pour chaque ensemble d'alternative, ici la matrice des covariances entre les termes d'erreur du modèle d'utilité est estimée à partir des similarités perçues entre les marques.

$$cor(\varepsilon_i, \varepsilon_j) = K.\exp(-\alpha d_{ij})$$

K et α sont des constantes à estimer qui transforment les distances perceptuelles dij en corrélations entre les termes d'erreur : $cor(\varepsilon_i, \varepsilon_j)$
dij est la distance perceptuelle entre les marques i et j.

$$d_{ij} = \sum_k [w_k . (Z_{ik} - Z_{jk})^2]^{1/2}$$

w_k est le poids de l'attribut k dans la fonction d'utilité
Z_{jk} est la quantité d'attribut k percue dans la marque j

Comme wk est estimé dans le modèle multi-attributs classique $u_j = w_k.Z_j + \varepsilon_j$, ce modèle nécessite l'estimation d'un faible nombre de paramètres.

Reprenant leurs travaux sur les modèles de choix, Kamakura et Srivastava (1986) ont proposé un modèle de préférences de type probabiliste et agrégé, dans lequel la préférence pour une marque s'exprime comme une fonction inverse de sa distance au point idéal (hypothèse classique de Coombs) mais aussi de la distance aux autres marques. Ce modèle permet de rejeter l'hypothèse des IIA .
Le modèle est donc tout à fait compatible avec les conditions de la concurrence telle qu'elle s'exprime réellement sur le marché. Le modèle est, notamment, mieux à même que les modèles faisant l'hypothèse des IIA, de prendre en compte l'effet du lancement d'un produit nouveau sur la structure de la concurrence.

La préférence de l'individu i pour une marque j s'exprime sous la forme d'une distance aléatoire Dij :

$$D_{ij} = \sum_{k} [w_k.(x_{jk}-I_{ik})^2 \ ] + \mu_{ij}$$

x_{jk} est la quantité d'attribut k possédée par la marque j lors d'une occasion d'achat particulière :
k = 1 à K
j = 1 à J
I_{ik} est la position du point idéal du consommateur i sur l'attribut k
w_k est le poids de l'attribut k
μ_{ij} sont les perturbations aléatoires suivant des lois normales $N(0,\sigma)$ indépendantes et identiques. μ_{ij} prend en compte la mauvaise spécification du modèle, l'omission de variables, l'aléa dans le comportement qui n'est pas lié aux attributs du produit (notamment, tout ce qui a trait à la situation d'achat).

Pour passer au niveau agrégé tout en prenant en compte l'hétérogénéité des consommateurs, les coordonnées des points idéaux sont supposées varier d'un consommateur à un autre, les poids des différents attributs restant eux identiques .
Le point idéal du consommateur i sur la dimension k s'écrit:

$$I_{ik} = a_k + e_{ik}$$

e_{ik} est une composante aléatoire liée à l'individu i relativement à la dimension k. Elle est distribuée selon la loi normale multivariée à k dimensions : $N(0,\Sigma)$

$$\Sigma = \begin{pmatrix} \sigma^2_{11} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \\ \sigma^2_{k1} & \cdots\cdots & & \sigma^2_{kk} \end{pmatrix}$$

Une telle matrice de variance covariance des eik signifie que les préférences selon certains attributs peuvent être corrélées (par exemple quand les consommateurs aiment les boissons sucrées ils aiment aussi les boissons gazeuses).
Ainsi, Pa(S), la probabilité pour que la marque a soit choisie par le consommateur i parmi l'ensemble S, peut être définie comme: .

$$P_{ai}(S) = \text{Prob } [D_{ai} < D_{bi} \quad \text{pour toute marque b de S}$$

Nous ferons les remarques suivantes concernant ce modèle de préférence.

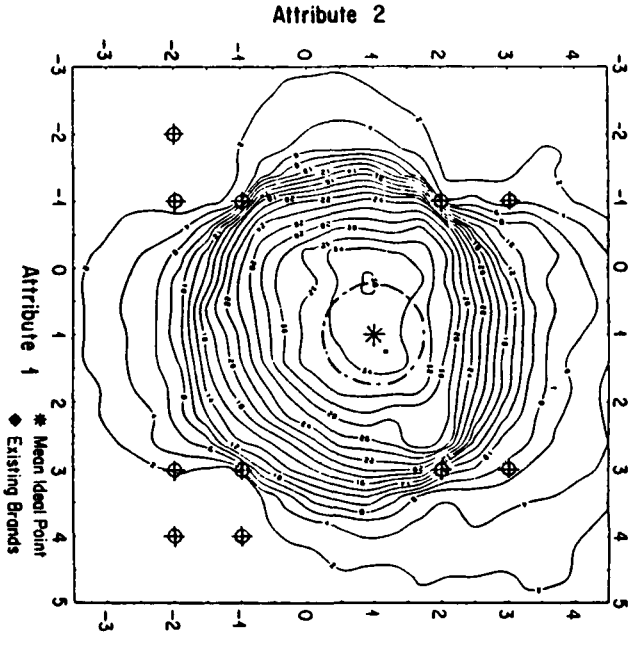
Le point essentiel qui le caractérise est que l'hétérogénéité des préférences (à travers la distribution des points idéaux) permet de prendre en compte l'interdépendance entre les marques qui prévaut lors du choix d'une alternative par un consommateur.

Cette interdépendance est prise en compte par le biais de la matrice de covariance des composantes aléatoires de l'utilité des différentes alternatives de choix. Une telle mesure permet de mieux cerner la réalité du phénomène concurrentiel sur un marché. Par exemple, elle permet de prendre en compte le cannibalisme existant entre les marques d'un même producteur.

La part de marché prévisionnelle d'une marque n'est pas uniquement fonction de la distance de cette marque au point idéal. Elle prend aussi en compte l'état de la concurrence existant, c'est à dire la proximité des autres marques. Ceci est tout à fait cohérent avec les remarques que nous avons faites au chapitre I quand nous avons comparé les concepts de Concurrence, Similarité perçue et Substituabilité.

Une marque, selon le modèle de Kamakura et Srivastava, peut avoir intérêt à s'éloigner du point idéal si certaines marques qui lui sont similaires en sont proches. Les courbes d'iso-parts de marché ne sont plus forcément des formes géométriques simples (cercles, ellipses) centrées sur le point idéal comme dans les modèles classiques de préférence. Elles peuvent devenir des courbes dont le tracé s'adapte aux zones de forte ou faible concurrence (voir figure suivante).

FIGURE 19 REPRESENTATION DES COURBES D'ISO-PREFERENCES



"Extrait de Kamakura et Srivastava 1986 p 206"

Enfin, nous soulignerons un autre avantage lié à une modélisation qui ne repose pas sur l'hypothèse des IIA. L'hypothèse des IIA, pour être "réaliste" contraint le modèle à fonctionner pour un ensemble de marques fortement similaires. En effet, dans de telles conditions, les niveaux de similarité entre les marques étant comparables, l'hypothèse reste "tenable". Ceci restreint alors les modèles de préférence probabilistes à n'étudier que

des univers de concurrence très restreints (c'est d'ailleurs une hypothèse de base du modèle Choice Map présenté plus loin). A l'inverse, le rejet de l'hypothèse IIA permet au modèle d'analyser des marques appartenant à un univers concurrentiel plus large, au sein duquel les niveaux de substituabilité perçue par le consommateur peuvent être très variables. Le modèle n'est plus alors un simple outil d'analyse au sein d'un sous-marché particulier mais il permet d'analyser la concurrence entre des marques appartenant à différents sous-marchés.

Ce modèle présente par contre une faiblesse qui mérite d'être relevée. Il fait l'hypothèse d'un point idéal unique. Ceci signifie que, même si ce point est probabilisé et permet de prendre en compte une forme d'hétérogénéité des consommateurs, la structure des préférences des consommateurs est supposée unique.

Ceci contraint le modèle à être utilisé :

- soit sur segment homogène de consommateurs
- soit sur un ensemble homogène de marques, ce qui retire alors au modèle tous les avantages vus précédemment et qui sont liés au rejet de l'hypothèse IIA.

Au vu de cette limite, le modèle Choice Map (voir plus bas) nous semble, d'un point de vue théorique, plus satisfaisant car il permet lui aussi de prendre en compte le degré de similarité entre les marques, même si cette prise en compte ne se fait qu'au plan agrégé.

SS2-3 LE MODELE CHOICE MAP

Elrod (1988) a proposé un modèle d'analyse des préférences qui permet de représenter la carte concurrentielle des marques et de connaître la distribution des préférences des consommateurs. C'est un modèle interne, vectoriel, agrégé et probabiliste Il repose sur trois considérations fondamentales :

- la structure des relations concurrentielles au sein d'un sous-marché est défini au préalable .
- les marques, au sein d'un sous-marché, se différencient selon des "attributs intangibles" car les attributs "physiques" sont rapidement copiés.
- les consommateurs ne sont pas toujours capables d'expliquer les différenciations qu'ils opèrent entre les marques car elles sont relatives à des attributs intangibles.

PRESENTATION DU MODELE

La modélisation (très élégante) s'effectue en deux étapes, au plan individuel puis agrégé.

1 MODELISATION AU PLAN INDIVIDUEL

Chaque individu est défini par son vecteur Y de fréquences d'achats observé concernant les marques: $Y = (y_1, y_2, y_i...y_J)'$

y_i est le nombre de fois que la marque i est choisie sur la période d'observation.

Par hypothèse, une seule marque est choisie à chaque occasion d'achat.
La marque i est achetée si son utilité Ui est maximum.

$$U_i = \max (U_1, U_2, .U_i...U_J) \quad (1)$$

U_i est l'utilité associée à la marque i .

L'utilité d'une marque i est la somme d'une composante déterministe v_i et d'une composante stochastique σ_i de moyenne nulle .

$$U_i = v_i + \sigma_i \quad (2)$$

vi est la résultante des attributs invariants (objectifs) de la marque i ainsi que de l'influence moyenne des variables de son mix marketing.
σ_i représente l'effet des variations dues aux actions de marketing à court terme de la firme (promotion) ainsi que l'effet dû à la situation d'achat du consommateur (variable point de vente). σ_i suit une distribution exponentielle double et négative : $Prob[\sigma_i \leq x] = \exp [-\exp(-x)]$
Les σ sont indépendants entre eux et suivent la même distribution.

La probabilité d'achat d'une marque i en n'importe quelle occasion peut alors s'écrire selon la forme classique du modèle MNL:

$$Prob[X=i] = p_i = \exp[v_i] / \sum_{j=1}^{J} \exp[v_j] \quad (4)$$

Ainsi, un consommateur avec le vecteur de probabilité d'achat des marques

$$P = [p_1, p_2 ...p_J],$$

observé sur T occasions d'achats aura le vecteur de fréquences d'achats suivant :

$$Y = [y_1, y_2, ...y_J]$$

affecté de probabilités qui sont données par la loi multinomiale

$$\text{Prob}[y/P,T] = \frac{T!}{\prod_{i=1}^{J} y_i!} \cdot \prod_{i=1}^{J} P_i^{y_i} \quad (5)$$

En considérant que la composante déterministe de l'utilité, v, est une fonction des attributs invariants de la marque, il est possible d'écrire :

$$v = A.W \quad (6)$$

A est la matrice des coordonnées des J marques en M dimensions
W est la matrice des poids des M dimensions

De par la forme de (6), nous pouvons constater qu'il s'agit d'un modèle de préférence de type vectoriel.

(4) + (5) + (6) permet d'écrire :

$$\text{Prob}\langle Y/W,A,T\rangle = \left(\frac{T!}{\prod_{i=1}^{J} yi!}\right) \cdot \prod_{i=1}^{J} \left(\frac{\exp(ai.w)}{\sum_{i=1}^{J} aj.w}\right)^{yi}$$

2 MODELE AGREGE

L'agrégation est effectuée en prenant en compte l'hétérogénéité des consommateurs. Celle-ci s'exprime au travers du poids qu'accordent les individus aux différentes dimensions.

L'hétérogénéité des poids qu'acccordent les individus aux attributs, (qui s'accompagne de l'hypothèse d'homogénéité des perceptions des produits), est une hypothèse qui a souvent été évoquée dans la littérature sur les processus de traitement de l'information. Elle est une des bases théoriques sur laquelle repose le modèle Indscal de Caroll et Chang (voir les chapitres 13,14 et 15 de Davies et Coxon 1981 et notamment le modèle Pindis de Borg et Lingoes au chapitre 14 du même ouvrage). D'autre part, plusieurs recherches ont montré que les perceptions des consommateurs sont stables et que ce sont les poids qu'ils accordent aux attributs qui varient, notamment en fonction de la situation d'usage du produit (voir Srivastava 1980 et 1981).

Ainsi, le vecteur w des poids, suit une loi normale à M dimensions, N(∂, Id), où Id figure la matrice identité à M dimensions. L'expression agrégée qui est obtenue prend en compte l'hétérogénéité des consommateurs. Cette hétérogénéité ne s'exprime pas, dans ce modèle, par des différences dans les perceptions mais par des différences dans les poids qui sont accordés aux dimensions évaluatives.

Les trois hypothèses faites sur les σ (les σ sont indépendants les uns des autres, ont des variances identiques et sont distribués selon une double exponentielle) équivalent à considérer que les choix individuels des marques se manifestent indépendamment des contextes concurrentiels de ces marques, comme c'est le cas dans le modèle de Luce. Le

modèle repose donc, au plan individuel sur l'hypothèse IIA. Bass et al (1984) ont montré que ces hypothèses sont valables pour 75% des ménages, notamment lorsque les produits sont d'un achat fréquent et en phase de maturité .
Une propriété intéressante de ce modèle est qu'au plan agrégé, les trois conditions précédentes ne sont plus nécessaires, ceci grâce à l'hypothèse qui est formulée quant à l'hétérogénéité des poids des dimensions évaluatives des consommateurs (les w). Ainsi, Choice Map est à même de prendre en compte la substituabilité entre les marques dans le processus de choix, au plan agrégé, ce que ne fait pas le modèle MNL classique.

Pour terminer cette section, nous proposons un tableau qui présente les principaux modèles de choix et de préférences issus de la littérature Marketing et qui permettent d'étudier la relation de concurrence des marques sur un marché particulier. Ce tableau présente les principales caractéristiques de ces modèles à savoir :

- s'il s'agit d'un modèle de choix ou de préférences
- si le modèle est interne ou externe
- si le modèle est à point idéal ou vectoriel (quand il s'agit d'un modèle de préférence)
- si le modèle repose ou non sur l'hypothèse des IIA
- quel type de choix est pris en compte, une marque parmi N (1/N) ou un nombre quelconque parmi N (k/n)
- de quelle nature sont les attributs du choix, physique ou perçue
- quelle est la variable dépendante (expliquée) dans le modèle
- le niveau d'analyse du modèle: agrégé ou individuel
- dans le cas où le modèle procède à partir de données individuelles, si l'hétérogénéité des consommateurs est prise en compte dans l'agrégation des fonctions d'utilité individuelles et si oui selon quel paramètre. Nous ferons remarquer que cette caractéristique est très importante, car c'est de la qualité de la manière avec laquelle cette l'hétérogénéité est prise en compte que dépend la qualité des résultats du modèle agrégé.

TABLEAU 11 PRINCIPAUX MODELES DE CHOIX ET DE PREFERENCES

AUTEURS	TYPE DE MODELE	INTERNE EXTERNE	POINT IDEAL/ VECTEUR	TYPE DE CHOIX	TYPE D'ATTRIBUT	HYPOTHESE IIA	VARIABLE DEPENDANTE	AGREGE/ INDIVIDUEL	AGREGATION SUR LES INDIVIDUS
BATSELL &LODISH 1981	MODELE DE CHOIX	-	-	1/N	PHYSIQUE	oui	PARTS DE CHOIX	INDIVIDUEL	-
CHAPMANN & LODISH 1982	MODELE DE CHOIX	-	-	1/N	PERCU	oui	RANG DE PREFERENCE	INDIVIDUEL	-
CURRIM 1982	MODELE DE CHOIX	-	-	1/N	PERCU DISCRET*	non	CHOIX	AGREGE	-
GUADAGNI & LITTLE 1983	MODELE DE CHOIX	-	-	1/N	PHYSIQUE	oui	CHOIX (1/0)	INDIVIDUEL	-
COOPER & NAKANISHI 1983	MODELE DE PREFERENCE	EXTERNE	POINT IDEAL OU VECTEUR	1/N	PERCU	oui	PREFERENCE	AGREGE	-
KAMAKURA & SRIVASTRAV A 1984	MODELE DE CHOIX	-	-	1/N	PERCU	non	CHOIX	AGREGE	-
KAMAKURA & SRIVASTRAV A 1986	MODELE DE PREFERENCE	EXTERNE	POINT IDEAL	1/N	PERCU	non	FREQUENCE DE CHOIX	AGREGE	DISTRIBUTION DES POINTS IDEAUX
DeSARBO & HOFFMAN 1986	MODELE DE PREFERENCE	EXTERNE OU INTERNE	POINT IDEAL	K/N	PERCU + vecteurs de proprietes	oui	CHOIX (1/0)	INDIVIDUEL	-
MacKEY & ZINNES 1986	MODELE DE PREFERENCE	EXTERNE OU INTERNE	POINT IDEAL	1/N	PERCU	oui	PREFERENCE DE RATIO	INDIVIDUEL	-
ELROD 1988	MODELE DE PREFERENCE	EXTERNE OU INTERNE	VECTEUR	1/N	PERCU + vecteurs de propriétés	oui (au plan individuel)	PARTS DE CHOIX	AGREGE	DISTRIBUTION DES POIDS DES ATTRIBUTS

* Les attributs sont perceptuels, mais de type discret. Les niveaux sont établis à partir d'interview de groupes.

CONCLUSION

Nous avons vu dans ce chapitre qu'un grand nombre d'approches différentes ont été proposées pour étudier les relations entre marques et leur représentation sous forme de cartes ou de graphes. Nous distinguons deux grands types de méthodes ou plutôt deux tendances dans la manière de modéliser la concurrence sans qu'il soit possible d'établir une frontière nette entre ces deux tendances.

Le premier type regroupe un ensemble de modèles dont l'objet est avant tout de découper le marché en sous-marchés et d'expliquer la logique de ce découpage. Ce découpage est le plus souvent fondé sur les modalités discrètes des attributs physiques du produit, comme par exemple l'opposition entre café moulu et café en grain. Le marché est ainsi conçu comme un vaste ensemble, compartimenté en sous-marchés. Selon ce point de vue, deux types de relations sont analysés, des relations intra ou inter-partitions. A l'extrême, comme dans le modèle de Hendry ou celui proposé par Grover et Srinivasan, toutes les marques au sein d'une même partition présentent le même niveau de relation qui s'exprime par la loi de "proportionnalité des transferts aux parts de marché des marques". Enfin, ce type d'approche se caractérise le plus souvent par une démarche de type confirmatoire, où le cherche/responsable marketing fait des hypothèses sur une structure pour tester ensuite si cette structure est cohérente avec l'observation des comportements sur le marché.

Parallèlement, coexiste un deuxième type de modèles dont l'objectif n'est pas tant le découpage du marché en sous-marchés, mais plutôt l'analyse des relations existant entre les marques et leur représentation sous forme de cartes. Il s'agit de méthodes qui raisonnent le plus souvent au sein d'un univers concurrentiel qui a été défini de manière étroite. Cet univers concurrentiel n'est pas conçu comme un tout se découpant en partitions (sous-marché) mais plutôt comme un espace au sein duquel il existe des zones de concurrence plus ou moins forte.

Les dimensions dans cet espace ont un caractère continu (par opposition à discret) mais aussi perceptuel. Appartiennent à cette catégorie toutes les méthodes dites de "positionnement des marques" qui ont connu un grand succès et qui utilisent essentiellement des données de type perceptuel. Y appartiennent encore tous les modèles de choix et de préférences, de conception plus récente et qui eux aussi ont pour objet l'obtention de cartes représentant le champ concurrentiel. Ces méthodes adoptent le plus souvent une définition du champ concurrentiel assez étroite. Notamment, l'hypothèse des IIA nécessite, pour être tenable, d'avoir défini un champ concurrentiel étroit. A l'inverse, les modèles qui ne font plus cette hypothèse se heurtent à des contraintes statistiques telles (lors de l'estimation des paramètres) que le nombre de marques analysables simultanément contraint l'analyste à fixer un champ concurrentiel encore plus étroit (voir chapitre I section 3). Le plus souvent, la démarche retenue est de type exploratoire, la structure n'est pas fixée *"a priori"*.

Il est bien sûr tentant de vouloir intégrer ces deux types d'approches au sein d'un même processus qui serait celui qu'utilise le consommateur pour effectuer ses choix. C'est ce que proposent Shocker et al (1987), comme nous l'avons décrit à la section 3 du chapitre I.

Dans une première phase, le processus de choix du consommateur est extensif, il raisonne au sein d'un univers de marques étendu. Ce processus est hiérarchique (intra-attribut), ce qui permet au consommateur de ne pas avoir à traiter trop d'informations

pour aboutir à la sélection d'un sous-ensemble de marques. Ce sous-ensemble de marques qu'il considère sérieusement dans son choix constitue son "ensemble évoqué", au sens de Howard (1977 voir dans Chandon et Strazziéri 1986). Au sein de ce sous-ensemble, le consommateur met en oeuvre une deuxième phase d'évaluation qui s'accompagne d'un traitement de l'information de type compensatoire, inter-attributs. Cette phase, plus coûteuse en information à traiter est possible parce que l'ensemble des marques à été préalablement restreint. Elle correspond à une vision spatiale de la concurrence.

La figure suivante représente le parallèle possible entre les deux types de modèles d'analyse de la concurrence et la modélisation du choix du consommateur en deux étapes. Le premier type de modèle d'analyse de la concurrence correspond à l'étude de la première phase du processus et le deuxième type de modèle, à la deuxième phase de ce processus.

Ce parallèle, bien que séduisant, doit cependant être considéré avec beaucoup de précautions, ceci pour deux raisons essentielles. Tout d'abord, il est difficile, comme nous l'avons déjà noté, de faire le lien entre le processus de choix d'un consommateur à un instant donné et les choix effectués par une population de consommateurs sur une période de temps. Comment un processus individuel et instantané se refléterait-il au niveau d'une population, sur une période de temps beaucoup plus longue ?
D'autre part, aucune étude empirique n'a, à notre connaissance, tenté d'établir un tel lien. L'établissement de ce lien constitue pourtant, selon nous, une voie de recherche intéressante.

Dans le chapitre suivant, nous allons aborder les problèmes statistiques relatifs à l'analyse de la structure contenue dans une matrice d'indices de relation entre marques (concurrence ou substituabilité) et à sa représentation graphique sous forme de carte ou de graphe.

TABLEAU 12 PARALLELE ENTRE MODELE D'ANALYSE DE LA CONCURRENCE ET PROCESSUS DE CHOIX DU CONSOMMATEUR

COMPORTEMENT DU CONSOMMATEUR	ANALYSE DE LA STRUCTURE DU MARCHE	EXEMPLES DE MODELES
PHASE 1	MODELE DE TYPE 1	
•UNIVERS ETENDU •TRAITEMENT HIERARCHIQUE (INTRA-ATTRIBUT) •ATTRIBUT PHYSIQUE ET DISCRET	•"STRUCTURATION DU MARCHE •ATTRIBUTS PHYSIQUES ET DISCRETS •DEMARCHE CONFIRMATOIRE •NON SPATIAL	•MODELE HENDRY •GROVER ET AL 1987 •JAIN ET AL 1990 •CARPENTER ET AL 1985 •MODELE PRODEGY
"ENSEMBLE EVOQUE"	"PARTITION DE MARCHE"	CHANDON ET AL 1986
PHASE 2	MODELE DE TYPE II •ETUDE DES RELATIONS ENTRE MARQUES	•PERCEPTOR •MODELES DE PREFERENCE ET DE CHOIX -CHOICE MAP -COMPETITOR -DeSARBO ET HOFFMAN 1986
•UNIVERS ETROIT •TRAITEMENT CONTINU (INTER-ATTRIBUTS) •ATTRIBUT PERCU	•UNIVERS ETROIT •REPRESENTATION DANS UN ESPACE CONTINU •ATTRIBUTS PERCUS •INDICE DE MESURE DE LA RELATION ENTRE MARQUES	
"CHOIX FINAL"	"CARTOGRAPHIE DU MARCHE"	

CHAPITRE TROISIEME: ANALYSE DE LA STRUCTURE DES INDICES: ASPECTS METHODOLOGIQUES LIES A LA REPRESENTATION DE LA STRUCTURE DES INDICES

Dans l'introduction générale de cette recherche, nous avons montré qu'il est très important, d'un point de vue méthodologique, de bien distinguer trois étapes dans l'analyse de la structure du marché :

- définir correctement le marché
- définir un indice de mesure des relations entre marques (substituabilité ou concurrence)
- analyser la structure de ces indices et la représenter sous forme de graphiques (cartes de concurrence, arbre hiérarchique...)

La définition du marché a été étudiée à la section cinq du chapitre premier, la mesure de la concurrence constitue le coeur de cette recherche et a fait l'objet du chapitre deuxième. Nous allons dans ce chapitre, étudier plusieurs points méthodologiques liés à la mise en oeuvre de la troisième phase: l'analyse de la structure contenue dans une matrice d'indices de concurrence. La spécificité de cette troisième phase tient à son fort contenu en statistique. Notre objectif n'est pas d'approfondir l'étude des méthodes permettant de caractériser une structure concurrentielle, non parce que nous pensons qu'il s'agit d'un problème secondaire, mais parce qu'il ne relève pas du cadre de cette recherche qui se veut essentiellement consacrée à l'étude de la structure des marchés d'un point de vue marketing. Nous aborderons successivement les points suivants :

-l'esprit des méthodes d'analyse, qui peut être confirmatoire ou exploratoire
-la prise en compte de la non-symétrie des indices de concurrence
-l'analyse de données de concurrence "ternaires", lorsque plus de deux dimensions sont traitées simultanément.
-les récents développements en matière de constitution de partitions et de traitement des données qualitatives.

SECTION 1 METHODES CONFIRMATOIRES ET EXPLORATOIRES

Il s'agit là d'une des distinctions majeure au sein des méthodes d'analyse des indices de relation entre marques. Nous distinguons en effet deux grandes approches : les approches confirmatoires et les approches exploratoires.

SS1 APPROCHES CONFIRMATOIRES

Selon cette optique, les indices calculés au plan empirique sont analysés afin de confirmer (ou d'infirmer) des hypothèses qui ont été faites sur la structure du marché. Il existe, au sein du courant des méthodes confirmatoires, deux types d'approches:

-des approches confirmatoires au niveau des indices

-des approches confirmatoires au niveau de la structure des indices

SS1-1 CONFIRMATION D'INDICES

Nous trouvons là les approches classiques de l'analyse de la structure des marchés, comme celles qui sont utilisées dans les modèles Hendry ou Prodegy. Le principe retenu consiste à émettre une hypothèse de structure du marché et à calculer, en fonction de cette hypothèse, des indices dits *"théoriques"*. Ces indices théoriques sont par la suite comparés aux indices empiriques correspondants. Une structure est considérée comme "valide" quand les indices théoriques sont proches des indices empiriques.

Cette méthode ne fait pas appel à des méthodes d'analyse multivariées qui permettent d'analyser simultanément l'ensemble des indices, elle se limite à l'utilisation de tests de comparaison des valeurs théoriques et empiriques (tests de moyennes, test de chi2).

SS1-2 CONFIRMATION DE STRUCTURES D'INDICES

La confirmation se situe ici au niveau de la structure des indices. Le principe consiste à émettre, comme précédemment, des hypothèses sur la structure et à essayer ensuite de les confirmer à partir des indices empiriques calculés. La méthode est plus puissante que dans le cas précédent puisqu'elle permet d'effectuer une validation au niveau de la structure toute entière. Elle a été rendue possible par le développement de méthodes dites *"confirmatoires"* (pour trouver un exposé complet de ce développement, voir l'ouvrage de Fornell 1982).

D'une manière générale, les méthodes confirmatoires ne consistent pas en de *"nouvelles"* techniques statistiques, mais plutôt en une nouvelle manière de les utiliser et d'estimer les paramètres. En particulier, les méthodes les plus classiques en "analyses produits" peuvent être utilisées de manière confirmatoire. C'est le cas de l'analyse factorielle (voir Joreskog 1978, Bentler 1980), de l'analyse discriminante (voir Kumar et Dillon 1987), de l'analyse des similarités et des préférences (voir MacKay et Zinnes 1986, Fornell et Denison dans Fornell 1982). Les avantages de telles approches sont nombreux, ils permettent notamment :

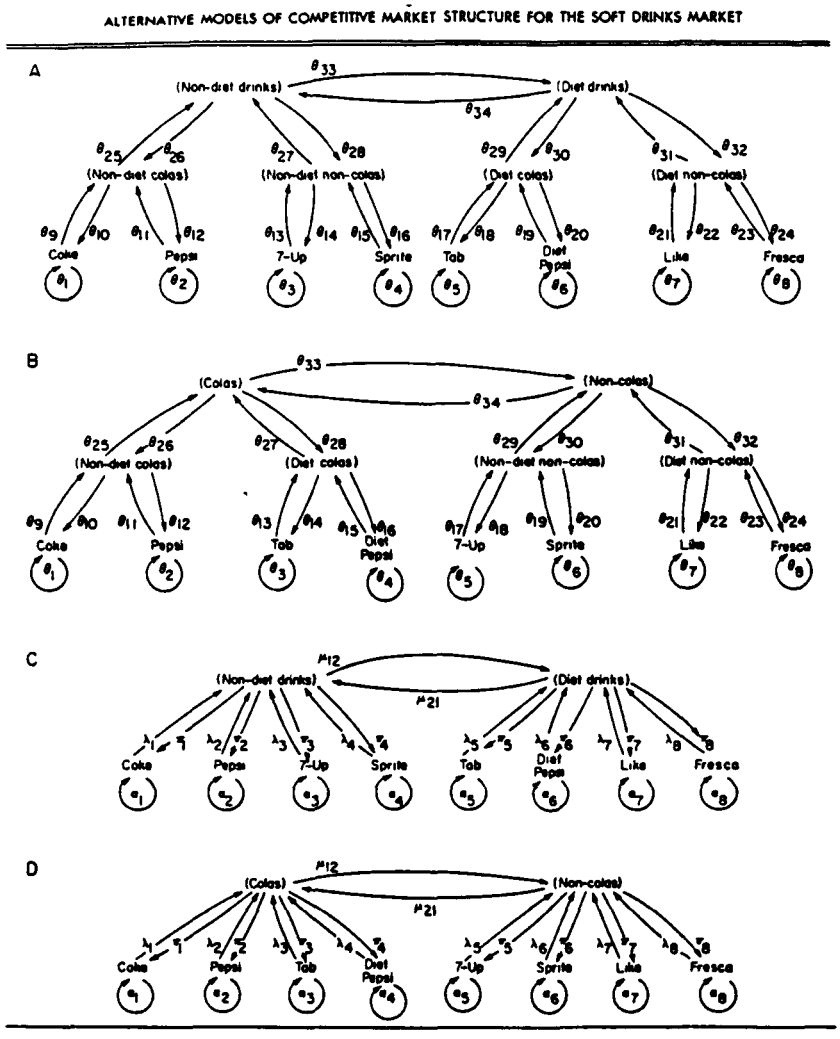
-de comparer et de tester statistiquement la différence entre deux structures

-de tester la dimension de la structure
-de tester l'orthogonalité de la structure, la position de certains points.
-etc ..

D'autres méthodes confirmatoires existent et ont été récemment proposées pour analyser la structure d'un marché. Citons en particulier l'analyse en classes latentes pour laquelle Clogg a développé un programme (MLLSA, voir Clogg 1979) qui a été utilisé dans plusieurs modèles (voir Grover et Dillon 1985, Grover et Srinivasan 1987, Parry et Gentler 1988).

Les modèles log-linéaires offrent aussi un cadre d'analyse naturel pour l'analyse des structures entre marques, puisqu'ils analysent des données de fréquences (parts de marchés, proportion d'acheteurs, proportion de ménages effectuant des transferts entre deux marques ..). On trouve une présentation de ces modèles dans DeSarbo et Hildebrand 1980, Konig, Nerlove et Oudiz 1979, Evrard et Laurent 1982. Ces modèles log linéaires se prêtent bien à une utilisation confirmatoire. Kumar et Sashi (1989) ont ainsi présenté une utilisation de cette technique dans un modèle très proche, dans sa conception, du modèle Hendry. La figure suivante illustre le type de résultats obtenu par les auteurs, sur le marché des soft drink. Notons que comme dans le modèle de Hendry, la méthode permet d'obtenir une estimation des coefficients reliant deux sous-marchés entre eux. Cette information ne figure pas dans la majorité des modèles d'analyse de la structure des marchés proposés dans la littérature. Elle est cependant, selon nous, très intéressante puisqu'elle permet d'établir le niveau de concurrence existant entre deux sous-marchés. Elle permet notamment au responsable d'une gamme de produits présents sur les différents sous-marchés, d'avoir un indicateur de la proximité concurrentielle de ses marques et donc du degré selon lequel peuvent se cannibaliser ses marques qui sont présentes sur les différents sous-marchés. Ceci peut permettre de détecter des opportunités de lancement d'une nouvelle marque au sein d'un sous-marché qui serait en faible concurrence avec les autres sous-marchés sur lesquels l'entreprise est déjà présente. Enfin, une telle représentation permet d'analyser globalement la structure concurrentielle, aux niveaux des différents sous-marchés et à l'intérieur de ces sous-marchés. Elle permet aussi de visualiser le degré de fidélité caractérisant chaque marque (sur le graphique, ceci peut se faire à l'aide des coefficients Téta1 à Téta8 ou Alpha1 à Alpha8).

FIGURE 20 REPRESENTATION DE LA STRUCTURE CONCURRENTIELLE A L'AIDE DU MODELE LOG-LINEAIRE



"Extrait de Kumar et Sashi 1989 p 450"

Pour conclure ce paragraphe consacré aux modèles confirmatoires, nous insisterons sur leur caractère managérial. En effet, ces modèles nécessitent de formuler des hypothèses sur les structures plausibles du marché et donc nécessitent un dialogue entre le chercheur ou le statisticien et le praticien du marché.

C'est en effet, selon nous, du praticien que doivent provenir les suggestions en matière d'hypothèses de structures à tester, ce qui constitue un gage de validité faciale de la structure qui pourra être caractérisée.

Le risque, contrepartie de cette démarche, est de ne tester *"que"* les hypothèses auxquelles le praticien a songé, ce qui peut être dangereux dans le cas d'une évolution rapide des structures du marché.

SS2 APPROCHES EXPLORATOIRES

Dans leur esprit, elles sont très différentes des approches précédentes. Une fois qu'une matrice d'indices de relations entre marques a été constituée, l'utilisation d'une méthode d'analyse multivariée permet de caractériser la structure contenue dans les données. Il n'est pas nécessaire d'avoir une expérience avec le marché ou d'avoir réfléchi à des structures hypothétiques plausibles. A la limite, n'importe quel indice peut être analysé et fournir des résultats interprétables . Notamment, à la différence des modèles confirmatoires, il n'est pas nécessaire que ces indices soient calculés à partir d'un modèle intégrant une théorie du choix ou de la préférence (comme par exemple le modèle d'utilité de Luce, 1959).

De nombreuses techniques sont utilisables. Nous citerons les plus connues qui sont les analyses factorielles, discriminantes, typologiques, les techniques de scaling.

Les inconvénients liés à ces approches sont les suivants :

-il n'y a pas de réflexion *"a priori"* sur les structures de marché hypothétiques

-les données analysées sont censées contenir "la vérité", ce qui peut être dangereux si les données ne sont pas fiables

-différentes méthodes et/ou différents jeux de données peuvent conduire à l'obtention de différentes structures (voir notamment Hauser et Koppelman 1979).

Par contre, ces méthodes sont faciles à mettre en oeuvre et largement diffusées dans les bibliothèques de programmes. Leur interprétation est en général simple et "souple".

Pour notre part, nous pensons que dans l'avenir, les méthodes confirmatoires connaîtront de plus en plus de succès et devront être de plus en plus utilisées, tant dans le domaine de la recherche que dans celui du métier des études.

SECTION 2 PRISE EN COMPTE DE LA NON-SYMETRIE DES RELATIONS ENTRE MARQUES

A la section quatre du chapitre premier, nous avons insisté sur le caractère non-symétrique des relations entre marques, qu'il s'agisse des concepts de Substituabilité ou de Concurrence.

Certains auteurs, notamment ceux qui utilisent des mesures de la concurrence à partir de l'élasticité croisée aux actions marketing ont retenu ce point et l'ont traité avec un succès variable (voir Vanhonacker 1984, Cooper 1988).

Dans les modèles d'analyse de la concurrence, le problème de la non-symétrie apparait en général sous les deux formes suivantes :

-soit, de par la forme théorique du modèle, les indices sont par définition symétriques, le problème ne se pose alors plus. C'est le cas, notamment, des indices théoriques dans le modèle Hendry.

-soit, les indices calculés au plan empirique ne sont pas symétriques et alors, cette non-symétrie est attribuée à des fluctuations d'échantillonnage et surtout, au fait que le marché n'est pas en équilibre. Lorsqu'un marché est en équilibre, la loi de proportionnalité aux parts de marché (voir Bass 1974, Rubinson et al 1980) équivaut à considérer les relations de concurrence, au sein d'un sous-marché, comme étant symétriques.

En fait, selon nous, la non-symétrie ne doit pas être considérée comme une déviation par rapport à un état d'équilibre, mais comme un état normal des relations de concurrence. Pour qu'un marché soit en équilibre, il faut que toutes les actions marketing des différents concurrents s'équilibrent et soient stables dans le temps. Or, l'observation des conditions sur le marché contredit sérieusement et en permanence cette hypothèse.

Par exemple, les actions de promotion à court-terme viennent constamment perturber cet état d'équilibre, avec des diminutions de prix, des mises en avant, des jeux, etc ...

D'autre part, de nouveaux produits sont introduits, les gammes existantes sont élargies, rajeunies, les packaging sont redessinés. Toutes ces actions indiquent que *le jeu concurrentiel passe, au contraire, par le "déséquilibre" du marché.* Un état d'équilibre ne peut être observé que sur une période très courte or, comme nous l'avons noté au premier chapitre, la concurrence doit être conçue en terme de diagnostic, sur un intervalle de temps.

A côté des motifs théoriques qui conduisent à concevoir la concurrence entre deux marques comme un concept symétrique, des raisons méthodologiques peuvent expliquer la tendance à considérer que les indices de concurrence ou de substituabilité sont symétriques. En effet, il est bien plus délicat d'analyser une matrice d'indices quand ces indices ne sont pas symétriques.

Pour cette raison, dans de nombreuses approches (citons Fraser et Bradford, 1983, qui, utilisant l'analyse factorielle, n'est capable d'analyser que des indices symétriques), les indices sont rendus symétriques en utilisant une fonction mathématique ou en effectuant tout simplement la moyenne entre les termes S_{ij} et S_{ji}. Ceci conduit, à notre avis, à un appauvrissement considérable des concepts de concurrence ou de substituabilité.

Nous allons présenter, dans cette section, les propositions qui ont été faites, en matière statistique afin d'analyser des indices non-symétriques.

SS1 TRAITEMENT DE LA NON-SYMETRIE PAR SEPARATION DES LIGNES ET COLONNES DE LA MATRICE D'INDICES EN DEUX ENTITES DISTINCTES

Cette méthode consiste à prendre en compte la non-symétrie des indices en considérant que les lignes et les colonnes d'une matrice d'indices correspondent à deux entités distinctes qui sont traitées et représentées de façon séparée. Ceci permet de considérer la relation entre deux marques i et j selon un double point de vue : entre i et j ou entre j et i.

Ainsi, Gower (1978) a proposé d'utiliser les modèles de dépliage (par exemple Prefmap) pour analyser une matrice d'indices non-symétriques. En effet, dans les modèles de préférences classiques comme Prefmap, chaque ligne de la matrice correspond à la préférence d'un individu (son point idéal), les colonnes correspondent aux marques par rapport auxquelles sont exprimées ces préférences. Les matrices analysées sont rectangulaires. L'utilisation de ces modèles avec des matrices (carrées) d'indices de similarité non-symétriques permet d'obtenir une représentation différente pour les lignes et les colonnes, ce qui est une manière de prendre en compte la non-symétrie. L'inconvénient de la méthode réside dans l'interprétation qui en est difficile.

Cooper (1988), a présenté une approche très novatrice et très riche d'un point de vue managérial. L'auteur propose de décomposer les indices d'élasticité croisée entre marques en deux effets: la puissance et la réceptivité de la marque (voir chapitre II, section 1).

Soit, pour la période t, la matrice d'élasticité E_t, de terme général $\{e_{ij}\}$.

	1	2		j	N
1				e_{1j}	
				e_{2j}	
				.	
i	e_{i1}	e_{i2}	e_{ij}		e_{iN}
.				.	
.				.	
N				e_{iN}	

Dans la ligne i, e_{i1} à e_{iN} représentent la manière dont les ventes de la marque i sont affectées par les actions (en terme de prix) des marques 1 à N.

De même, dans la colonne j, figurent les éléments e_{ij} à e_{Nj} qui caractérisent dans quelle mesure la marque j affecte les ventes des marques 1 à N.

L'élasticité E_{ij} de la marque i par rapport aux actions de la marque j peut donc s'écrire comme le produit (produit scalaire) de la réceptivité de la marque i (sa propension à être affectée par les autres) par la puissance de la marque j (sa propension à influencer les ventes des autres marques). La similarité entre deux lignes i et i' mesure la similarité des deux marques en terme de réceptivité.

La similarité entre deux colonnes j et j', mesure la similarité de ces deux marques en terme de puissance sur les autres marques.

Ce double point de vue (réceptivité, puissance) permet de mener des analyses sur des matrices d'indices non-symétriques. De plus, cette décomposition n'est pas un simple

artifice statistique visant à rendre symétrique ce qui ne l'est pas, elle est tout à fait cohérente avec le principe de concurrence, qui doit être conçu comme une relation à double sens. Dans la deuxième partie, nous développerons un indice construit à partir de données de transferts, en utilisant cette double logique.

SS2 ANALYSE SPATIALE DES MATRICES ASSYMETRIQUES

Ce type d'approche consiste à utiliser une fonction mathématique permettant de passer d'un ensemble de dissimilarités non-symétriques à un ensemble de distances euclidiennes, qui elles, ont toutes les propriétés pour être traitées par les méthodes classiques.

Une des premières approches proposée est le modèle Asymscal de Young (1975, cette version est intégrée dans le modèle Alscal disponible dans les bibliothèques SAS et SPSS). Dans la modélisation de Young, la non-symétrie serait due à une non-symétrie des poids qui permettent de pondérer les axes dans le modèle de scaling individuel Indscal.

$$\partial_{ij} = [\sum_{t=1}^{T} v_{it}.w_{jt}.(x_{it} - x_{jt})^2]^{1/2}$$

∂_{ij} est la dissimilarité empirique (observée) entre i et j
v_{it} est le poids de la dimension t dans l'espace des lignes de la matrice
w_{jt} est le poids de la dimension t dans l'espace des colonnes. On retrouve la forme du modèle Indscal. lorsque v_{it} est égal à w_{jt}
x_{it} et x_{jt} sont les coordonnées des marques i et j sur la dimension t

Cette procédure, si elle est capable de prendre en compte la non-symétrie du point de vue statistique, n'apporte pas, à notre avis, une solution théorique intéressante au problème.

Weeks et Bentler (1982) ont proposé un modèle de scaling dans lequel les dissimilarités ∂ij (non-symétriques) sont décomposées en une distance dij (symétrique) et la partie dissymétrique de la dissimilarité.

$$\partial_{ij} = b.d_{ij}^2 + k + c_i - c_j + e_{ij}$$

b et k sont les coefficients d'estimation de l'équation
c_i et c_j correspondent à la part non-symétrique de la dissimilarité
e_{ij} est un terme d'erreur aléatoire

Les termes de distance dij sont ensuite introduits dans un modèle classique d'analyse des similarités. Cette alternative met bieb n valeur la difficulté liée au traitement des données assymétriques puisque celles-ci sont rendues symétriques par un traitement statistique préalable.

Harshman, Green, Wind et Lundy (1982) ont proposé un modèle (Dedicom) qui permet de décomposer les similarités. Dans ce modèle, l'assymétrie serait due à l'effet de facteurs sous-jacents non observables, de même nature que ceux qui sont extraits dans une analyse factorielle. A la différence de l'analyse factorielle, ces facteurs seraient liés entre eux de manière non-symétrique. C'est cette non-symétrie entre les facteurs non observables qui, en se reflétant sur les objets, rendrait la relation entre les objets elle aussi "non-symétrique".

Dans une étude sur le marché des Shampooings, les auteurs montrent que les deux dimensions sous-jacentes, Vigueur et Abondance de la chevelure, sont deux concepts non observables liés de façon assymétrique. La vigueur évoque très souvent l'abondance alors que l'inverse est moins vrai. De cette relation assymétrique naîtraient les relations assymétriques entre les qualificatifs associés aux marques de shampooings. Leur modèle s'exprime sous la forme :

$$X = A.R.A' + E$$

X : matrice (n,n) de similarités
A : matrice (n,p) des loading des n objets sur les p dimensions non observables, A' en est la transposée.
R : matrice des covariances entre les p dimensions sous-jacentes
E : matrice (n,n) des termes d'erreur

Cette modélisation est en fait un cas particulier d'analyse factorielle où la matrice R n'est pas symétrique (quand R est la matrice identité, le modèle se ramène à celui de l'analyse factorielle orthogonale). Les auteurs ont montré que la méthode permet d'obtenir un ajustement plus précis que les techniques classiques, qui ne prennent pas en compte l'assymétrie.

DeSarbo et Manrai (1988) ont proposé une approche dont la logique est plus théorique que statistique, ce qui constitue un progrès par rapport aux approches précédentes. Les auteurs ont opérationnalisé le modèle de similarité psychologique de Krumhansl (1978). Ce modèle (voir chapitre I section 4) a pour caractéristique de considérer la perception comme étant fondamentalement assymétrique. La dissimilarité s'exprime comme :

$$\partial_{ij} = f^{-1}(d\sim_{ij})$$

$d\sim_{ij} = d_{ij} + a.\Gamma(i) + b.\Gamma(j) + e_{ij}$
$d\sim_{ij}$ est la distance "modifiée" entre i et j
d_{ij} est la distance euclidienne classique entre i et j
$\Gamma(i)$ et $\Gamma(j)$ représentent la densité des voisinages des objets i et j. C'est cette densité du voisinage qui permet de prendre en compte la non-symétrie des similarités.
a et b sont des coefficients positifs
eij est un terme d'erreur

Toutes les techniques précédentes procèdent de manière à obtenir une distance euclidienne qui permet ensuite d'effectuer des représentations spatiales. Un autre traitement des données assymétriques consiste à abandonner la notion d'espace euclidien dans lequel les distances sont obligatoirement symétriques.

SS3 ANALYSES NON SPATIALES DES MATRICES ASSYMETRIQUES

Les représentations non-spatiales utilisent la théorie des graphes et plus particulièrement la théorie des arbres additifs. Un arbre additif est un graphe dans lequel il n'y a pas de bouclage des relations (le chemin à partir d'un point quelconque ne permet jamais de revenir à ce point).

Ces méthodes constituent en fait, un cadre d'analyse plus général et moins contraignant que celui de l'analyse typologique hiérarchique (Johnson 1967). Dans une analyse hiérarchique, les objets (marques) regroupés doivent vérifier (au sein de chaque groupe) l'inégalité ultramétrique, à savoir:

$$d(i,j) \leq \max \{d(i,k) , d(j,k)\} \text{ , pour tout k n'appartenant pas au groupe}$$

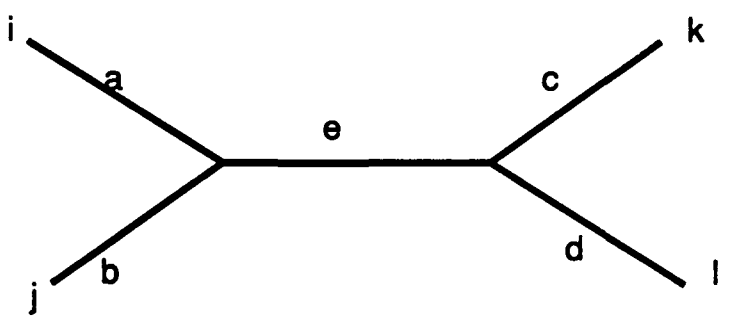
Deux objets groupés doivent toujours être plus proches entre eux qu'avec n'importe quel autre.

Tversky et Sattath (1977), ont proposé un modèle moins contraignant, qui repose sur l'inégalité des quatre points. Soient deux groupes (i,j) et (k,l). Il faut alors vérifier l'inégalité :

$$d(i,j) + d(k,l) \leq \max \{ d(i,k) + d(j,l) , d(i,l) + d(j,k)\}$$

Cette contrainte d'inégalité permet de construire des arbres additifs dans lesquels les objets (marques) constituent les sommets et les arcs reliant ces objets constituent des distances entre ces objets. La distance entre deux objets est égale à la somme la plus courte des arcs reliant ces deux sommets.

Exemple :



Sur le graphique précédent, la distance entre i et j, d(i,j) est égale à (a + b).

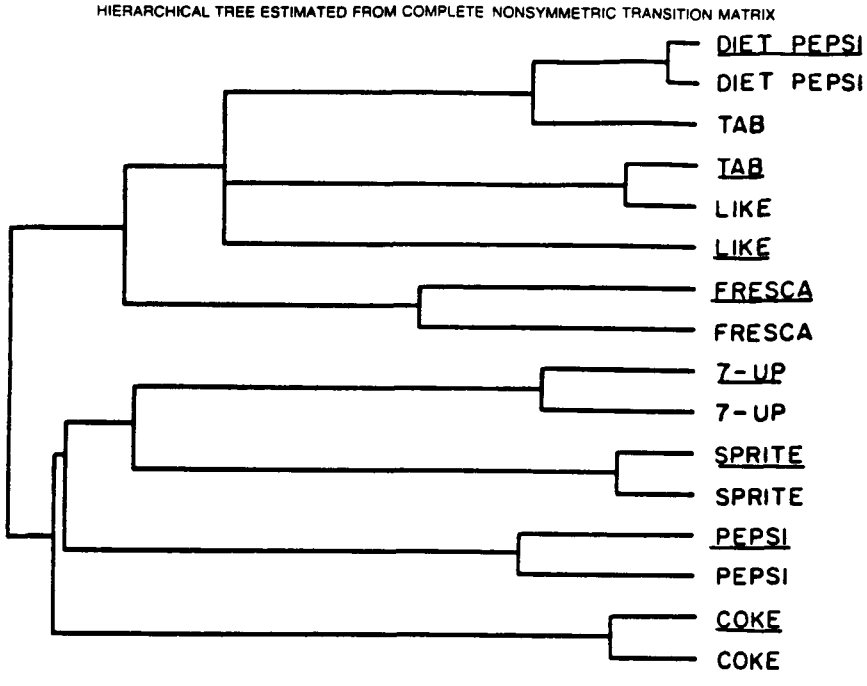
Les auteurs ont proposé un modèle (Addtree) qui procède en deux étapes :

1 - La recherche de la forme de l'arbre
2 - L'estimation des longueurs des arcs dans cet arbre

L'inconvénient de la méthode est lié à la lourdeur de la procédure d'estimation.

DeSarbo (1982) a proposé un modèle (Gennclus) assez similaire mais plus général et plus performant, qui permet notamment de prendre en compte la non-symétrie (les principes de ce modèle sont exposés à la section quatre). DeSarbo et de Soete (1984), ont proposé une application du modèle Gennclus sur les données de Rao et Sabavala (1981). Cette application permet d'effectuer l'analyse de la structure du marché à partir des données de transferts originelles qui sont assymétriques. Les auteurs ont montré qu'une analyse typologique hiérarchique sur le triangle inférieur de la matrice des indices (ce que font Rao et Sabavala) ou sur le triangle supérieur, ou encore sur la moyenne des deux triangles, aboutit à des typologies substantiellement différentes. La structure obtenue avec le modèle d'analyse non-symétrique figure dans le schéma suivant.

FIGURE 21 REPRESENTATION DE LA HIERARCHIE DU MARCHE A L'AIDE DU MODELE GENNCLUS

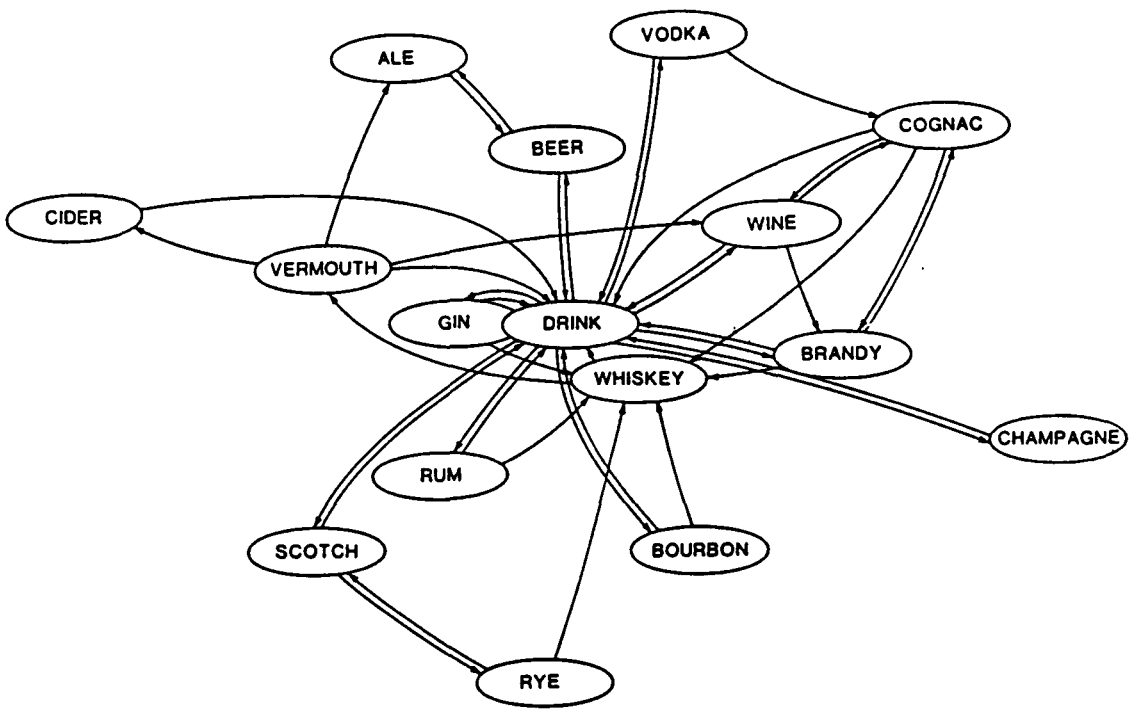


"Extrait de DeSarbo et De Soete 1984 p 607"

Hutchinson (1989) a proposé un modèle dans lequel les dissimilarités assymétriques sont modélisées à l'aide de la théorie des graphes. A partir d'une série de théorèmes sur les propriétés des graphes, le modèle permet de déterminer la forme du graphe et la longueur des arcs qui en relient les sommets. Dans ce type d'analyse, la logique de la "typologie" disparait au profit de celle de "réseau de relations". Cette approche nous semble très prometteuse, notamment lorsqu'elle est appliquée à des données de type "psychologique".

Une des caractéristiques distinctive de cette approche par rapport aux précédentes est que les arcs sont directionnels et les réseaux peuvent "boucler". Ce type de modèle offre donc un cadre d'analyse encore plus général que celui des modèles additifs ou hiérarchiques. L'exemple suivant, sur le marché des boissons alcoolisées, permet d'apprécier la spécificité et l'intérêt de la méthode.

FIGURE 22 RESEAU DE RELATIONS ENTRE BOISSONS ALCOOLISEES



"Extrait de Hutchinson 1990"

Concernant les trois types d'approches (SS1, SS2 et SS3) que nous venons de présenter et qui permettent de traiter des données non-symétriques, nous ferons les commentaires suivants.

La première, notamment telle que l'a formalisée Cooper (1988), nous semble particulièrement bien adaptée à l'étude de la réalité du marché. De même, la deuxième présente l'avantage d'améliorer sensiblement la qualité de l'ajustement statistique, comme l'ont montré DeSarbo et Manrai (1988), par rapport à des méthodes qui ne prennent pas

en compte la non-symétrie. Ces deux types de méthodes permettent d'obtenir des représentations spatiales du champ concurrentiel, cartes qui restent un instrument privilégié de diagnostic pour les responsables.

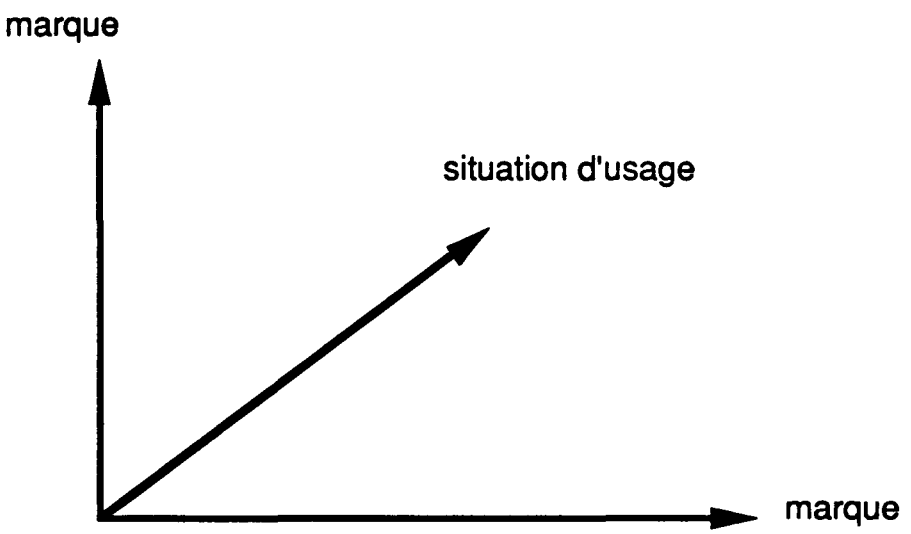
La troisième abandonne l'idée de représentation spatiale et nous semble particulièrement bien adaptée à l'étude de "la réalité du consommateur" dont, comme de plus en plus de chercheurs le reconnaissent, les processus de traitement de l'information ne sont pas forcément compatibles avec les espaces euclidiens des statisticiens. Pour cette raison, il nous semble que la théorie des graphes offre un cadre d'analyse plus naturel pour l'étude des structures cognitives de l'individu. Les études à venir sur la substituabilité et la similarité perçue entre les marques ont donc, nous semble-t-il, tout intérêt à utiliser et approfondir ce type de modélisations.

SECTION 3 ANALYSE DES MATRICES TERNAIRES D'INDICES

Nous avons insisté dans le premier chapitre sur l'importance de la prise en compte de concepts comme la situation d'usage lors de l'étude des relations entre marques. Nous avons aussi noté que le temps joue un rôle important dans cette étude et qu'il est important d'étudier l'évolution des structures concurrentielles au cours du temps, de manière à contrôler l'efficacité des actions menées par *"nous et la concurrence"*.

D'autre part, il est souvent important de tenir compte de l'hétérogénéité des consommateurs, ce qui conduit à l'obtention de structures de concurrence différentes pour un ensemble de segments de consommateurs.

En termes statistiques, ces trois points (situation d'usage, temps, hétérogénéité des consommateurs), représentent chacun une dimension supplémentaire à rajouter aux deux dimensions de la matrice de données de concurrence entre marques, comme le montre le graphique suivant.



La prise en compte de ce type de variables, si le chercheur souhaite avoir une approche globale (être capable d'analyser simultanément l'ensemble des matrices de relations entre indices correspondant, par exemple, à différentes situations d'usage) nécessite d'utiliser des méthodes capables de traiter des données dites *"ternaires"*.

Les approches permettant d'analyser des données en trois dimensions existent mais sont, à notre avis, sous-utilisées dans le domaine de l'analyse de la concurrence entre les marques.

Elles sont pourtant, pour certaines, disponibles depuis longtemps. Par exemple, Tucker a proposé un modèle d'analyse en composantes principales en trois dimensions en 1963. Kroonenberg et Leew (1980) ont proposé une estimation de ce modèle qui est fiable et rapide, à l'aide des moindres carrés alternatifs.

Dans le domaine de l'analyse des similarités, le modèle Indscal qui permet de traiter des données ternaires (voir Carroll et Chang dans Davies et al 1981) date de 1970. De la même époque, le modèle Kyst permet d'envisager une matrice de données comportant jusqu'à sept dimensions !

Récemment, Jedidi et DeSarbo (1988) ont proposé un modèle d'analyse des similarités permettant de traiter des données binaires (oui/non, achat/non achat) selon trois dimensions, par exemple, le fait qu'un individu i perçoive la marque j comme étant adaptée à la situation k. Ce modèle est en fait la généralisation du modèle de DeSarbo et Hoffman (1986) présenté au chapitre II, section 5, au cas où la matrice des données comporte trois dimensions.

Ce modèle est présenté par les auteurs comme permettant de prendre en compte la situation d'usage du produit. Les points (ou vecteurs) idéaux, des individus ont des positions spécifiques en fonction des situations d'usage alors que les marques ont des positions fixes. Les poids des dimensions d'évaluation des alternatives peuvent varier d'une situation d'usage à une autre. Pour le modèle à point idéal, l'utilité de la marque j pour l'individu i, dans la situation d'usage k (U_{ijk}) s'écrit :

$$U_{ijk} = \sum_{t=1}^{T} w_{kt}.(x_{jt} - y_{it})^2 + c_i$$

w_{kt} : poids de la dimension t dans la situation d'usage k
x_{jt} : coordonnée de la marque j selon la caractéristique t
y_{jt} : coordonnée du point idéal de la marque j sur la caractéristique t
c_i : constante

Les auteurs ont utilisé le concept de point idéal (ou vecteur idéal) "flottant" dans l'espace des situations d'usage, l'espace des objets étant fixé. Cette modélisation est cohérente avec les résultats de Miller et Ginter (1979) selon lesquels, les situations d'usage affectent les poids accordés aux attributs des produits, les perceptions restant, elles, identiques (voir chapitre I section1).

De la même manière, les méthodes d'analyse des correspondances ont été généralisées au cas de tableaux de contingence en trois dimensions (voir Lebart, Morineau et Warwick 1984).

Carroll (1968) a généralisé l'analyse canonique au cas où il y a plus de deux ensembles de variables. Ténenhaus (1983), a étendu le champ d'application de l'analyse canonique généralisée au cas où les données sont de n'importe quel type (nominales, ordinales, numériques). La méthode permet d'analyser simultanément p matrices rectangles individus x variables. Elle est le cadre général permettant de retrouver (comme l'ont montré Tenenhaus et Young 1985) un grand nombre de méthodes, telles que l'analyse des correspondances, l'analyse en composantes principales, la régression multiple, l'analyse des correspondances. Les programmes ACGT et ACPT (voir Tenenhaus 1987) permettent respectivement de traiter le cas de l'analyse canonique généralisée et de l'analyse en composantes principales sur des données ternaires.

Pour conclure, nous insisterons sur le fait qu'il existe un grand nombre de méthodes disponibles pour traiter les données ternaires mais que ces méthodes sont peu utilisées. Nous citerons comme application dans le domaine de l'analyse de la structure des marchés, Jedidi et DeSarbo (1988) qui ont généralisé le modèle de DeSarbo et Hoffman au cas à trois dimensions en introduisant comme troisième dimension la situation d'usage. Cooper (1988) a, lui, traité la dimension du temps en analysant des matrices d'élasticités

croisées entre marques calculées sur plusieurs périodes ceci, en utilisant l'analyse factorielle à trois dimensions.

Des recherches plus nombreuses pourraient approfondir ces approches. Les exemples d'applications possibles peuvent être les suivants : analyser différentes structures de concurrence selon différentes situations d'usage, différents segments de consommateurs, différentes périodes du temps, différents points de ventes ...

SECTION 4 DEVELOPPEMENTS RECENTS EN MATIERE DE CONSTRUCTION DE PARTITIONS ET DE TRAITEMENT DES DONNEES QUALITATIVES: APPLICATIONS A L'ANALYSE DE LA STRUCTURE DES MARCHES

SS1 NOUVELLES METHODES D'ETUDE DE LA PARTITION DES MARCHES

De nombreux modèles d'analyse de la structure des marchés considèrent que le découpage d'un marché en sous-marchés s'effectue sous la forme de partitions et que chaque marque est affectée à un sous-marché et un seul. A coté des raisons théoriques qui peuvent être invoquées pour justifier ce procédé, ceci peut être pour partie attribué au fait que les méthodes classiques d'analyse typologique hiérarchique procèdent par découpage (ou par regroupement) successifs des marques, en fonction des indices de similarité entre ces marques.

Un nouveau type d'analyse typologique existe pourtant depuis maintenant plus de dix ans. Il permet de concevoir la structure du marché sous la forme de *partitions de recouvrement* où chaque marque peut appartenir à *"différents"* sous-marchés ou bien sous la forme de partitions floues où chaque marque appartient, à des *degrés divers,* à l'ensemble des sous-marchés.

Shepard et Arabie (1979), Shepard (1980), ont présenté un modèle de typologie de recouvrement appelé Addclus. Dans ce modèle, la similarité entre deux objets est exprimée comme la somme des poids des caractéristiques qu'ils ont en commun.

$$S_{ij} = \sum_{k=1}^{K} w_k \cdot p_{ik} \cdot p_{jk}$$

$p_{ik} = 1$ si i possède la caractéristique k
$p_{ik} = 0$ autrement.
Comme le notent les auteurs, quand les pik ne sont plus binaires (0 ou 1) le modèle est équivalent à un modèle d'analyse factorielle.
wk est le poids de cette caractéristique dans la relation de similarité

Ce modèle a été généralisé par DeSarbo (1982) qui a présenté le modèle Gennclus, modèle qui permet d'effectuer des analyses de typologie hiérarchique de recouvrement, permet de traiter la non-symétrie des indices et enfin, permet d'analyser des données qui peuvent être d'intervalle ou de ratio. Pour n objets et m partitions, la matrice S~ des similarités estimées s'exprime sous la forme:

$$S\sim = P.W.Q + C$$

$S\sim$: matrice (n,l) de similarités estimées, de terme général $s\sim_{ij}$
S : matrice (n,l) de similarités empiriques observées, de terme général s_{ij}
P : matrice (n,m) décrivant l'appartenance des objets correspondant aux n lignes de la matrice S, aux m partitions (les valeurs sont égales à 1 s'il y a appartenance, 0 dans le cas contraire).
Q: matrice (l,m) décrivant l'appartenance des objets correspondant aux l colonnes de la matrice S aux m partitions
W: matrice (m,m) décrivant l'association entre les m groupes
C: matrice (n,l) d'éléments c, constantes additives

Quand la matrice de similarité est symétrique, alors P = Q.
Pour une partition de recouvrement, la contrainte d'ordre suivante s'impose:

$$\sum_{k=1}^{m} p_{ik} \geq 1$$, pour tout objet i (chaque objet appartient au moins à une partition)

Pour une partition forte, la contrainte d'égalité suivante s'impose:

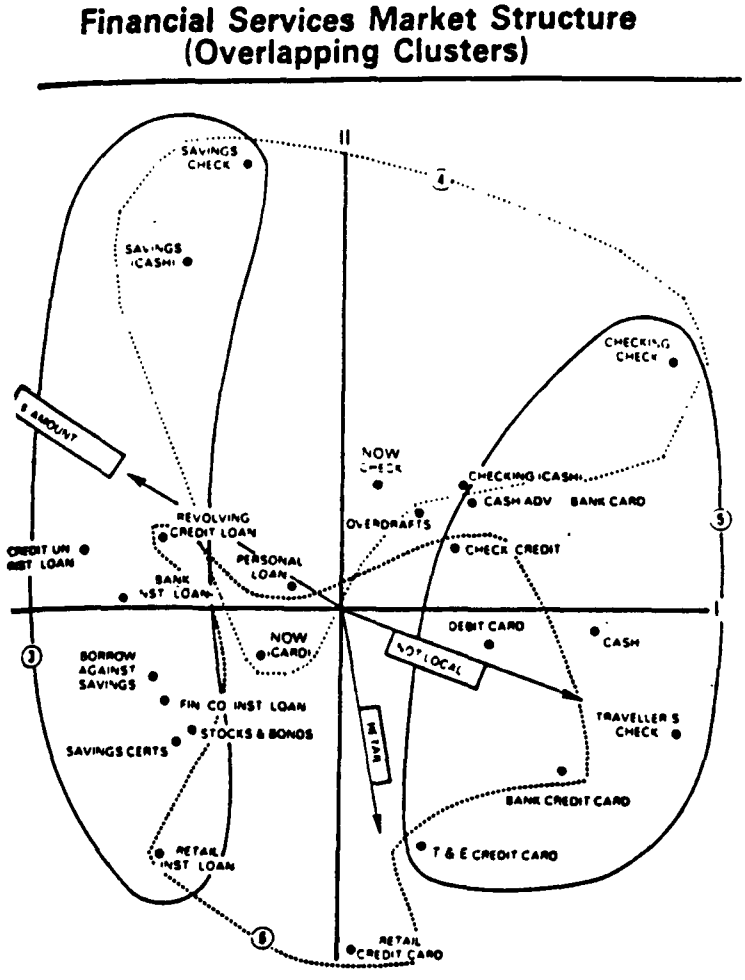
$$\sum_{k=1}^{m} p_{ik} = q_{jk} = 1$$, pour tous objets i, j

L'algorithme utilisé minimise le stress :

$$\min \sum_{i=1}^{n} . \sum_{j=1}^{l} [s_{ij} - s\sim_{ij}]^2$$

Arabie et al (1981) ont proposé une application du modèle Addclus pour représenter la structure du marché des aliments pour le petit déjeuner, à partir de données de similarité perçues. Les auteurs ont insisté sur le caractère plus *"réaliste"* de ce type de représentations. Il existe une autre application du modèle Addclus dans Srivastava, Alpert et Shocker (1984). Les auteurs utilisent la méthode pour représenter la structure du marché des produits financiers, à partir de données de substituabilité perçue dans l'usage (voir Figure 23 suivante).

FIGURE 23 EXEMPLE DE PARTITION DE RECOUVREMENT



"Extrait de Srivastava et al 1984 p 41"

Plus récemment, Hruschka (1986) a proposé l'utilisation de la typologie floue pour définir et analyser la structure d'un marché, mais aussi pour effectuer la segmentation des consommateurs.

Dans une partition floue, chaque marque appartient à l'ensemble des partitions, mais avec un degré variable. Ceci conduit à calculer des probabilités d'appartenance de chaque marque aux différents sous-marchés. Chaque marque est finalement affectée au sous-marché dans lequel sa probabilité d'appartenance est la plus forte.

Reprenant le critère général en typologie de minimisation de la variance intra-groupe, la méthode consiste à minimiser l'expression suivante :

$$\sum_{i=1}^{n} \cdot \sum_{k=1}^{c} u(i,k).d^2(i,k)$$

k = 1 à c classes
i = 1 à n marques
$d^2(i,k)$ est le carré de la distance de la marque i à la classe k

u(i,k) constitue le critère d'appartenance de la marque i à la classe k
u(i,k) = 0 ou 1 pour une partition forte
u(i,k) ε [0,1] pour une partition floue
$\sum_k u(i,k) = 1$ pour tout i
u(i,k) est une fonction de la distance de la marque i à la classe i

Face à l'intérêt de méthodes de ce type, qui sont du reste disponibles dans les bibliothèques comme SAS ou SPSS, nous ne pouvons que regretter leur faible degré d'utilisation en Marketing, notamment dans le cadre de l'analyse de la structure de concurrence entre les marques, qui pourtant se prête très naturellement à un tel type d'utilisation.

SS2 LE TRAITEMENT DES DONNEES QUALITATIVES

Pour terminer cette section, nous présenterons quelques développements récents en matière d'analyse des données de type qualitatif (voir Green et al 1977 pour une revue de l'analyse des données qualitatives en marketing).

Nous avons déjà insisté sur la croissance de l'utilisation des données de panels. Ces données se caractérisent essentiellement par leur caractère qualitatif. En effet, y figurent le plus souvent :

-des historiques décrits en terme d'achat (1)/ non-achat (0)
-les caractéristiques des marques selon des attributs physiques et discrets
-les caractéristiques des individus selon leurs caractéristiques socio-démographiques qui sont encore de type discret

L'essentiel de l'information se présente donc sous la forme de données discrètes.

Levine (1979) a été un des premiers auteurs à proposer d'utiliser les données brutes d'achat ou de non-achat des marques pour représenter la structure d'un marché. Dans la matrice analysée figurent en ligne les individus puis les marques, de même en colonnes figurent les individus (qui ont ou non choisi les marques) et les marques (qui ont été ou non achetées par les individus).

La matrice se présente de la manière suivante :

	Individus		Marque	
	1 2	N	N+1	N+M
1 2 . . . . N	0			
N+1 . . . . N+M			0	

Au croisement individu x individu, on trouve une sous-matrice nulle, de même pour le croisement marque x marque. Les deux autres sous-matrices sont constituées de 1 (la marque a été achetée) ou de 0 (dans le cas inverse).

Cette matrice est analysée par la méthode des centroïdes qui consiste à placer chaque individu au barycentre des marques qu'il a achetées et chaque marque au barycentre des individus qui ont choisi cette marque, ceci jusqu'à l'obtention de la convergence de la représentation des individus et des marques. Cette méthode est en fait un cas particulier d'analyse des correspondances.

Parallèlement au développement de l'analyse des correspondances et de l'analyse des correspondances multiples (voir Lebart et al 1984) d'autres méthodes se sont développées pour analyser les données de type qualitatif. Nous citerons bien sûr les modèles Logit et Probit qui ont connu un succès immense, essentiellement dû à la souplesse des cas qu'ils sont capables de traiter. Nous avons déjà parlé de ces méthodes et cité les principales références en la matière (voir section 3 du chapitre I). Ces modèles sont en fait un cas particulier d'une catégorie plus large de modèles appelés modèles log linéaires (voir Konig et al 1979, Blattberg et Dolan 1981, Evrard et Laurent 1983, DeSarbo et Hildebrand 1980) qui sont eux aussi très utiles pour analyser des données discrètes (voir section précédente).

En matière de modèle de préférences, Takane (1983) a développé un modèle à point idéal qui n'utilise que des données de choix de type binaire. Nous citerons encore le modèle de DeSarbo et Hoffman (1986) dont nous avons déjà parlé. Enfin, il nous parait important de souligner les travaux de Tenenhaus qui, en utilisant des procédures de codage optimal, a généralisé plusieurs méthodes auparavant limitées à l'analyse de

données quantitatives (voir Tenenhaus 1977, 1979, 1983), notamment l'analyse en composantes principales, la régression et l'analyse canonique généralisée.

Pour conclure, nous insisterons sur le fait que les statisticiens ont développé un ensemble de méthodes qui constitue un corps de plus en plus souple et puissant. Il serait profitable que les recherches sur l'analyse de la structure des marchés puisent dans ce potentiel de méthodes puisque, nous l'avons souligné, l'analyse de la structure contenue dans une matrice d'indices de mesure de relations entre marques relève, pour une bonne partie, du domaine de la statistique. Cependant, l'élément essentiel, qu'il ne faut pas perdre de vue, est de savoir mesurer de manière fiable et valide la concurrence (ou la substituabilité) entre les marques. Ce point sera l'objet de notre deuxième partie, partie qui se place résolument au niveau de l'étude de la concurrence entre les marques à partir de données de panels.

CONCLUSION DE LA PREMIERE PARTIE

Dans l'introduction générale, nous avons situé l'analyse de la structure des marchés par rapport aux autres grands concepts de l'analyse *"produit"* que sont la segmentation et le positionnement des produits (voir figure 1, p 5). Sur cette figure apparait la place fondamentale qu'occupe l'analyse de la structure des marchés. La *"définition du marché"* apparait comme une étape préalable et nécessaire à toute analyse. L'étude de la *"structure du marché"* apparait, elle, comme une étape qui doit être menée conjointement avec la segmentation des consommateurs. Nous avons d'ailleurs présenté à ce sujet plusieurs modèles qui effectuent segmentation et analyse de la structure du marché de façon simultanée (voir les modèles de Grover et Dillon 1985, Grover et Srinivasan 1987).

De la qualité de ces analyses, dépend la qualité des politiques de produit que l'entreprise saura mettre en place par la suite, notamment celle de positionnement des produits.

La méthodologie d'analyse de la structure d'un marché que nous avons proposée dans l'introduction générale se décompose, rappelons le, de la façon suivante. Avant toute chose, il est indispensable de définir l'objet de l'analyse, définition qui doit permettre de répondre à des questions du type : *"pour quelles raisons le responsable marketing désire-t-il analyser la structure du marché ?", "quel type de décision devra permettre d'opérer cette analyse ?"*. De la qualité de ce diagnostic concernant la nature du besoin et les objectifs de l'analyse dépendra la valeur des choix qui seront faits dans les trois phases suivantes.

PHASE 1 DEFINIR CORRECTEMENT LE MARCHE

La définition préalable des besoins revêt ici une importance capitale.

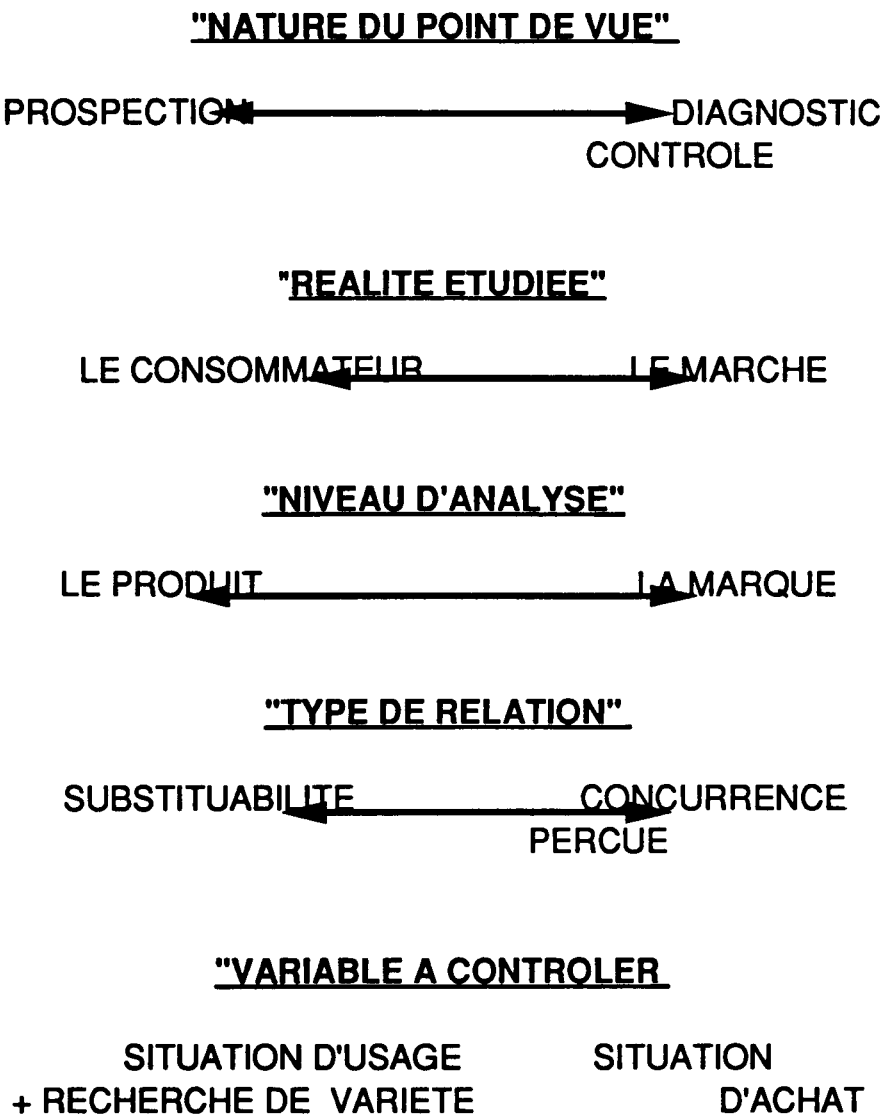
Si l'objet de l'étude est de nature prospective, comme par exemple étudier la possibilité du lancement d'un nouveau produit ou l'élargissement d'une gamme de produits existante, il est important d'adopter une vision *large* du marché. Cette vision devra s'appuyer sur l'étude de la réalité du consommateur et utiliser des méthodes semblables à celles qui ont été proposées par Srivastava et al (1978). La situation d'usage du produit est dans ce cas une variable dont il est indispensable de tenir compte. D'autre part, le produit parait mieux adapté que la marque en tant qu'unité d'analyse, il est, en effet, nécessaire d'adapter au mieux le degré de généralité de l'unité d'analyse au contexte de l'étude. Le marché devra être défini en terme d'unités *"substituables"* dans l'esprit du consommateur (au sens où nous l'avons défini au premier chapitre), pour une situation ou un groupe de situations d'usage particulières.

A l'autre extrême, si l'objet final de l'analyse est le diagnostic de la position qu'occupe une marque sur le marché ou l'analyse de l'efficacité des actions qui ont été menées sur le terrain (promotion, mise en avant ..), le marché pourra être défini de manière plus *"étroite"*, en choisissant une unité d'analyse fine (la marque ou même une référence particulière de la marque). L'objet de l'analyse est alors la *"réalité du marché"*. Le marché est défini en terme de marques concurrentes (effectivement concurrentes, quelles qu'en soient les raisons).

Ces deux points de vue peuvent être représentés comme étant deux extrêmes entre lesquels il est important de bien savoir se placer.

Ces deux extrêmes sont séparés en deux parties bien distinctes qui correspondent au fait que c'est soit la *"réalité consommateur"* soit la *"réalité du marché"* qui est étudiée (voir figure 24 suivante). La réalité consommateur place résolument l'analyse dans le domaine des perceptions et la principale variable à contrôler est alors la situation d'usage (éventuellement la recherche de variété). Dans l'étude de la réalité du marché, où ce sont les comportements effectifs qui sont analysés, un nouveau type de facteur intervient: ce sont les variables situationnelles qui recouvrent en particulier toutes les actions en magasin.

FIGURE 24 CONTINUUM DES NIVEAUX D'ANALYSE SUR UN MARCHE



PHASE 2 LE CHOIX D'UNE MESURE DE LA RELATION ENTRE LES MARQUES

Une fois les objectifs définis avec précision, la réalité étudiée fixée (marché ou consommateur), les frontières du marché définies, l'unité d'analyse choisie (le produit ou la marque), le marché défini, il est alors nécessaire de faire le choix d'une mesure de la relation entre les marques. Si les éléments précédents ont été correctement fixés, ce choix doit être relativement aisé.

Comme en segmentation des consommateurs il n'existe pas "*un type*"de variable de segmentation universel, en analyse de la structure des marchés il n'existe pas "*une mesure*" de la relation entre les marques "*universelle*". C'est un point fondamental sur lequel nous nous permettons d'insister.

Deux grands types de mesures sont utilisables selon que l'on étudie la réalité consommateur (substituabilité) ou la réalité du marché (concurrence). Au sein de chacune de ces deux catégories coexistent encore de nombreuses mesures différentes et il appartient au responsable de l'étude de choisir la mesure la mieux appropriée au problème tel qu'il a été défini. Le responsable marketing tiendra compte, dans ce choix, de problèmes pratiques tels que :

-le nombre d'unités analysées : si celui-ci est élevé, il faudra éviter d'utiliser des méthodes reposant sur des hypothèses fortes, comme les IIA ou bien des méthodes qui nécessitent l'estimation d'un nombre de paramètres trop important.

-la certitude avec laquelle le marché a été défini pourra guider le choix vers des méthodes qui ne sont pas trop sensibles à cette définition lors du calcul des indices. Cependant, comme nous l'avons constaté au chapitre II, ces mesures sont rares, d'où l'importance qu'il y a à définir correctement le marché. Notamment, lors de l'étude de la "réalité de marché" à partir de données de comportement passé, la qualité de cette définition a une influence très importante sur le calcul des indices de concurrence. Ce problème est moins aigu lorsque c'est la "réalité du consommateur" qui est étudiée. Par exemple, pour des indices construits à partir de données de transferts, cette définition influe directement sur le calcul des indices.

PHASE 3 ANALYSE DE LA STRUCTURE DES INDICES

Il est important, dans une méthodologie d'analyse de la structure d'un marché, de bien distinguer les phases deux et trois de manière à être conscient du type de mesure qui est utilisé et afin de ne pas mélanger le problème de la mesure avec celui du traitement de cette mesure. Nous avons mis l'accent au chapitre III sur les points méthodologiques qu'il s'agit de traiter avec soin. En particulier, faut-il utiliser une méthode confirmatoire ou exploratoire?

A la limite, une démarche purement confirmatoire peut être trop rigide et conservatrice. A l'inverse, les démarches purement exploratoires souffrent d'un manque de rigueur et de trop d'arbitraire. Ainsi, il peut être intéressant de développer une méthode "*mixte*" permettant d'allier la créativité des méthodes exploratoires à la rigueur des méthodes confirmatoires. Dans un premier temps les structures possibles seront caractérisées à l'aide d'une démarche exploratoire, structure que l'on essayera de confirmer dans un deuxième temps.

D'autre part, nous avons insisté sur la nécessité de prendre en compte la non-symétrie de la relation entre marques. Certaines recherches ont montré que la non-symétrie peut expliquer la moitié de la variance des relations entre marques (DeSarbo et De Soete 1984), ne pas tenir compte de la non-symétrie équivaut à perdre une grande part du phénomène concurrentiel.

Une autre possibilité qui doit être considéré avec attention dans cette phase est l'analyse de données ternaires. Quand les données disponibles permettent de prendre en compte plus de deux dimensions (par exemple différentes matrices d'indices correspondant à différentes situations d'usage), plutôt que de procéder à des analyses séparées ou à l'agrégation sur la troisième dimension, il est plus instructif de traiter simultanément ces trois dimensions. Cette troisième dimension peut permettre, alternativement, de traiter le cas de l'hétérogénéité des consommateurs (qui appartiennent à différent segments), de l'influence de la situation d'usage, de l'effet du temps ou du point de vente, etc...

D'autre part, le responsable marketing pourra avoir intérêt à utiliser des méthodes non spatiales (par opposition aux méthodes spatiales). Parmi ces méthodes, de nouvelles techniques utilisant la théorie des graphes permettent d'effectuer la représentation de la structure des indices sous forme de réseau. De telles analyses peuvent venir en complément, ou se substituer, aux analyses spatiales.

Pour finir, nous observerons que la recherche sur l'analyse de la structure des marchés peut se développer et progresser sur les trois fronts que constituent les trois étapes de la méthodologie d'analyse de la structure des marchés. Ces trois points peuvent donner lieu à des recherches indépendantes.

Nous avons insisté sur la nécessaire qualité des mesures de concurrence qui sont utilisées pour étudier la structure d'un marché. Face à la diversité des modèles proposés, le chercheur doit s'interroger sur la convergence de ces modèles en matière de résultats utilisables par le responsable marketing. Par exemple, au sein même des modèles de mesure de la concurrence utilisant des données de transferts entre marques, de multiples modèles coexistent et utilisent des mesures de la concurrence différentes. Ces mesures sont elles fiables, convergent-elles vers des structures identiques ? A quel degré ces mesures sont-elles dépendantes des caractéristiques du ménage consommateur, du point de vente, de la période d'observation ?

Nous nous proposons d'étudier ces points dans la deuxième partie de cette recherche.

*
* *

PARTIE II

ETUDE DE LA VALIDITE DU CONCEPT DE CONCURRENCE ENTRE LES MARQUES ET DES VARIABLES QUI INFLUENCENT CE CONCEPT

INTRODUCTION DE LA DEUXIEME PARTIE

Cette recherche sera consacrée à l'étude de la mesure de la concurrence entre les marques à partir de données de comportement observées sur le marché.

LE CHOIX D'UN DOMAINE DE RECHERCHE

Nous avons choisi, dans le cadre de cette recherche, d'étudier l'analyse de la structure des marchés à partir de données de comportement. Nous allons successivement justifier les deux composantes de notre choix à savoir :

- l'étude de la mesure des niveaux des relations entre marques
- l'analyse de la structure des marchés à partir de données de panels

1 -ETUDE DE LA MESURE DE LA RELATION ENTRE LES MARQUES

Dans la première partie, nous avons vu que la méthodologie de l'analyse de la structure des marchés peut être découpée en trois grandes phases :

1 - La définition de l'univers de concurrence
2 - La mesure de la relation entre les marques
3 - La représentation et l'interprétation de la structure concurrentielle

La phase 1 consiste à définir l'univers de concurrence à étudier, c'est à dire à fixer les frontières du marché. Selon nous, il n'y a pas de place pour une approche formalisée de ce problème lorsque le chercheur ou le praticien utilise des données de comportement. La définition des frontières du marché doit plutôt s'appuyer sur deux types d'approches :

1 - Une réflexion managériale concernant les buts poursuivis dans l'étude de la structure du marché. La définition d'un univers de concurrence est, en effet, en partie spécifique au type d'étude que l'entreprise souhaite mener (voir Srivastava et al 1981). Nous distinguerons les approches tactiques (à court terme) des approches stratégiques (à plus long terme).

Dans les approches tactiques, l'objectif consiste en l'étude d'un sous-marché particulier et des relations de concurrence directes sur ce sous-marché. L'unité d'observation peut aller jusqu'à la référence d'une marque dans son niveau de décomposition le plus fin (par exemple Coca Cola Standard pet 150 cl).

Dans les approches stratégiques, l'entreprise s'intéresse plutôt à la structure qui unit les sous-marchés entre eux et à leur possible évolution. L'unité d'observation peut aller jusqu'à la catégorie de produit, le marché des

produits pour le Bain et la Douche est à cet égard un bon exemple. Plusieurs marchés cloisonnés et correspondant à des situations d'usage différentes ont fini par évoluer vers un grand marché constitué des différents marchés antérieurs. Dans ce cas, l'observation de données sur les comportements passés ne peut conduire qu'à constater et non anticiper les évolutions du marché liées au lancement de produits par la concurrence.

2 - Des études utilisant des données de jugement (notamment toutes les analyses à partir de matrices Produits x Usages du type *procédure Stefflre*). De telles études présentent un intérêt évident si l'entreprise souhaite étudier de possibles liens de substitution entre des produits qui ne sont pas directement concurrents et quand elle s'intéresse aux évolutions possibles d'un marché pour l'avenir.

La phase 2 consiste à mesurer les niveaux des relations entre marques. Nous avons passé en revue dans la première partie les différentes approches présentes dans la littérature. A l'inverse du point précédent (la définition du marché), la mesure des niveaux des relations entre marques se prête très bien à une aproche formalisée.
C'est sur ce terrain que nous avons choisi de développer notre recherche .

La phase 3 consiste à effectuer la représentation de la concurrence entre les marques. Une fois obtenue une mesure de la concurrence entre les marques qui soit *"valide"* il s'agit alors d'en extraire le maximum d'information à savoir :

- obtenir une représentation du marché sous forme de carte ou de graphe
- avoir les positions précises des marques les unes par rapport aux autres, leur proximité étant fonction de l'intensité de leurs relations.
- caractériser les facteurs qui structurent le marché, c'est à dire en connaître la nature et l'importance relative.

Nous considérons, qu'une fois obtenues des données valides lors de la phase précédente, cette phase se réduit à un problème d'analyse des données qui est plutôt du ressort de la Statistique que du Marketing. C'est pourquoi, cette recherche sera consacrée à l'étude de la mesure de la concurrence entre les marques. Nous ne traiterons pas des problèmes relatifs aux phases 1 et 3 de la méthodologie d'analyse de la structure des marchés .

2 - ETUDE DE LA STRUCTURE DES MARCHES A PARTIR DE DONNEES DE COMPORTEMENTS

Nous avons vu dans la première partie qu'il existe deux grandes approches de l'analyse de la structure des marchés qui ont chacune, d'un point de vue théorique et pratique, des forces et des faiblesses :

- les approches utilisant des données de jugement
- les approches procédant à partir de données de comportement

La mesure des niveaux de relations entre marques, à partir de données de comportement, s'inscrit à notre avis dans un domaine bien précis: celui des études de type tactique .

En effet, mesurer la concurrence entre deux marques i et j à partir de seules données de comportements d'achats n'est possible que lorsque ces deux marques i et j sont directement concurrentes. En dehors de ce cas, établir que deux marques sont concurrentes parce qu'elles sont achetées par le même panéliste et que celui-ci a, par exemple, changé de l'une pour l'autre, peut se révéler hasardeux. Dès que l'analyse s'écarte des problèmes de concurrence directe, deux marques peuvent être achetées (simultanément ou successivement) parce qu'elles sont concurrentes, mais aussi pour d'autres motifs (le hasard des processus d'achats de la ménagère, ...).

Notre choix s'est porté sur les approches procédant à partir de données de comportement, nous étudierons ainsi, dans cette recherche, la mesure de la concurrence entre les marques. Notre choix repose sur les considérations suivantes:

- au plan général, l'accumulation actuelle d'informations relatives aux comportements et aux caractéristiques de l'offre en magasin correspondant à ces comportements représente pour le chercheur et le décideur en marketing un terrain d'investigation immense qui reste encore assez peu exploité. Il est donc naturel que les recherches actuelles se tournent vers ce domaine .
- l'analyse de la structure des marchés à partir de données de jugement apparait comme étant un courant de recherche plus ancien et plus mature qui mérite moins d'effort de validation .
- à l'inverse, les recherches portant sur l'analyse de la structure des marchés à partir de données de panels se sont multipliées au cours des dix dernières années. Il est nécessaire, face à cette multiplication, d'effectuer réflexions et validations concernant les outils qui sont proposés.
- d'un point de vue pratique, les bases de données de panels comme SCAN 5000 constituent un réservoir d'information à la fois riche et fiable qu'il est motivant d'analyser .

LES OBJECTIFS DE LA RECHERCHE

Nous avons choisi dans cette recherche d'étudier le problème de *"la mesure de la concurrence entre les marques à partir de données de panels"*. Cet objectif recouvre les ambitions suivantes :

1 - APPLIQUER LA THEORIE DE LA MESURE A L'ETUDE DE LA VALIDITE DU CONCEPT DE CONCURRENCE

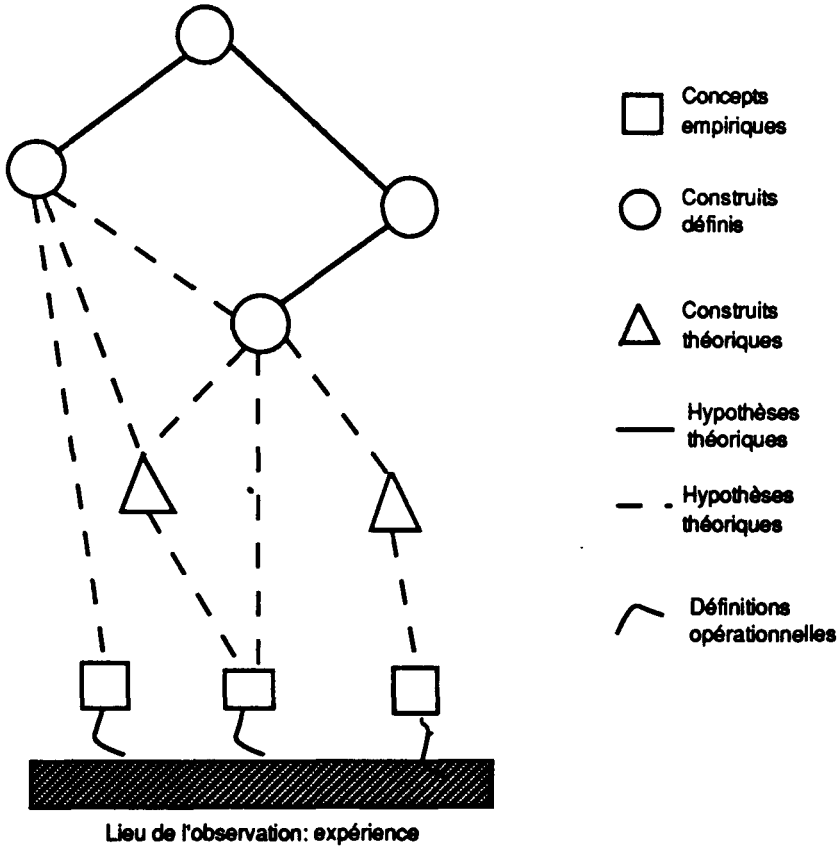
Face à la multiplication des méthodologies d'analyse de la structure des marchés qui sont proposées, il est bon de confronter et d'évaluer ces différentes approches. La théorie de la mesure offre un cadre d'analyse tout à fait propice à ces objectifs. Cette théorie recouvre un corps théorique et un ensemble de techniques qui sont destinés à apprécier la validité des mesures effectuées en psychologie.

Depuis une dizaine d'années (notamment depuis la publication en février 1979 par la revue Journal of Marketing Research d'un numéro spécial consacré à la validité des mesures), les chercheurs en marketing sont plus nombreux à étudier la fiabilité et la validité des instruments de mesure qu'ils proposent.

Ce besoin de validation des mesures est selon nous à la jonction de deux courants de pensée qui se sont développés de manière très dynamique ces dernières années en marketing.

Le premier courant, qualifié de *"conceptualiste"*, par opposition au courant classique dit *"opérationnaliste"*, procède à partir du constat de base suivant:"*la majeure partie des phénomènes étudiés en Science Sociale (en particulier en marketing) n'existent pas pour autant qu'ils sont observés ou mesurés"*. Ils n'existent que par les mots qui ont servi à les définir. On les appelle des concepts théoriques .Ces concepts naissent d'abord au niveau théorique et peuvent ensuite être observés au plan empirique.

FIGURE 1 STRUCTURE DE LA THEORIE



" Adapté de Bagozzi 1980 p 65 "

Ainsi, les mesures que nous pouvons faire ne sont que le reflet du concept tel qu'il a été défini au plan théorique. Les concepts d'Attitude, de Satisfaction du consommateur en sont des exemples bien connus. Ce faisant, les reflets que nous mesurons ne sont pas exempts d'erreur de mesure. Il est donc nécessaire de prendre en compte et d'évaluer l'importance de ces erreurs afin de statuer sur la qualité des mesures qui ont été pratiquées. L'ouvrage de Bagozzi (Bagozzi 1980) est à notre avis une des meilleures manifestations de ce courant de pensée en marketing.

Le deuxième courant de pensée peut être caractérisé par sa manière de traiter l'information collectée. C'est le courant des *"approches confirmatoires"* du traitement de l'information (par opposition aux approches empiriques). L'esprit d'un tel courant peut être résumé de la manière suivante : *"les théories (si elles sont valides) existent avant les observations qui en sont la manifestation"*. Les

observations qui sont collectées servent à confirmer ou infirmer les théories mais elles ne font pas les théories. Un tel point de vue par rapport au traitement de l'information est en filiation avec les théories développées par K. Popper (Popper 1963). Selon cet auteur, une théorie en Science Sociale ne peut être *"démontrée"* (comme un théorème mathématique), elle ne peut être que rejetée ou confirmée. On parle ainsi de *"falsificationnisme"* dans la mesure où une théorie (ou une hypothèse) est tenue pour vraie tant qu'elle n'a pas été invalidée. Plus le nombre d'observations ou d'expériences qui ont confirmé une théorie est important, plus cette théorie prend de la valeur au sein du corps scientifique (théorie de l'accumulation), mais personne ne sera jamais certain que la théorie est *"vraie"*.
Une telle vision de la démarche scientifique a donné naissance à une nouvelle génération de méthodes qui sont dites *"confirmatoires"*, confirmatoires parce que leur point de départ ne réside pas dans les données qu'elles analysent mais dans la théorie qui (peut-être) est à l'origine des données observées. Ce courant de recherche est fort bien développé dans l'ouvrage de Fornell (Fornell 1982). Ici encore, une telle approche débouche sur le problème de la qualité de la mesure puisque c'est d'elle que dépend la validation des théories. Notamment (voir Nunnally 1978), l'erreur de mesure diminue le lien observé (empirique) entre deux concepts. Une bonne théorie peut être invalidée par des mesures entachées d'erreur et inversement une mauvaise théorie peut être validée par des mesures de mauvaise qualité.

Ces deux courants (Conceptualiste et Confirmatoire) représentent à nos yeux une transformation importante quant à l'esprit et à la forme de la recherche en marketing . Il est difficile de nier la valeur de la démarche et des techniques qui sont issues de ces deux courants.

La recherche dans le domaine de l'analyse de la structure des marchés est, nous l'avons souligné, très dynamique depuis une dizaine d'années. Cependant, nous pouvons constater que cette recherche a peu intégré les méthodes et les démarches dont nous venons de parler. Beaucoup d'approches ont été proposées pour analyser la structure d'un marché, mais peu de recherches ont été menées dans le but d'établir la validité de ces approches. La concurrence entre deux marques est pourtant un concept théorique dont on ne mesure que les reflets (multiples) à travers les mesures qui en sont proposées. Ces mesures procèdent à partir de définitions du concept de concurrence qui peuvent être très différentes. Il sera donc intéressant de confronter ces différentes approches, de voir dans quelle mesure elles convergent et d'évaluer leurs qualités respectives.

La théorie de la mesure a été développée dans le domaine de la *"Psychométrie"* (mesure des phénomènes psychologiques). La concurrence telle qu'elle est mesurée à partir de données de panels est bien sûr hors de ce champ. Cependant, nous verrons (et c'est un des objectifs de cette recherche) que les outils de la théorie de la mesure s'appliquent très bien à de telles mesures des comportements.

2 - ETUDIER L'INFLUENCE DU CONCEPT DE CONCURRENCE A LA MANIPULATION DE PLUSIEURS VARIABLES

Le deuxième objectif de cette recherche sera d'étudier l'influence de variables classiques (comme le Revenu, la CSP) sur les mesures de la concurrence entre les marques et d'analyser le degré de sensibilité des mesures de la concurrence à la manipulation de ces variables.

Il est important de savoir si la structure d'un marché établie à partir de données de comportement est une caractéristique stable de ce marché ou bien si au contraire elle est spécifique à des segments particuliers de consommateurs. Plusieurs recherches à partir de données de jugement ont montré que les espaces perceptuels sont stables et ne varient pas d'un segment de consommateur à un autre (voir première partie, chapitre I). Par contre, les préférences pour les marques, au sein de ces mêmes espaces, varient de façon beaucoup plus forte. Il sera intéressant de mener le même type d'analyses concernant les cartes de marché obtenues à partir de données de comportement.

PLAN DE LA RECHERCHE

Le chapitre IV sera consacré à la description de la recherche et sa mise en oeuvre. Nous présenterons le *"terrain"* sur lequel est établie cette recherche, à savoir le panel SCAN 5000 de la société Nielsen. Nous décrirons les choix qui ont été opérés en matière de marchés, de marques ainsi que les indices qui ont été étudiés. Enfin, nous détaillerons les calculs qui ont été effectués, leur mise en forme en vue de leur analyse dans les chapitres suivants.

Le chapitre V sera consacré à l'étude de la validité des indices et à leur comparaison. Après avoir exposé les éléments fondamentaux de la théorie de la mesure, nous en utiliserons les outils afin d'apprécier le niveau de validité des indices étudiés. Nous effectuerons en outre un ensemble d'analyses visant à comparer les indices et à étudier leur regroupement en *"Familles d'Indices"*.

Le chapitre VI sera consacré à l'analyse de l'influence de six variables sur la mesure de la concurrence entre les marques: le Revenu, la Quantité Consommée, la CSP, la Taille du Foyer, le type de Point de Vente et le Temps. Nous analyserons successivement l'influence globale de ces variables sur la mesure de la concurrence puis leur influence sur chacun des indices. Nous terminerons cette étude par la comparaison des effets respectifs de l'instrument de mesure (l'indice) d'une part et de l'objet mesuré (les comportements de choix) d'autre part sur la mesure elle même.

CHAPITRE IV PRESENTATION DE LA RECHERCHE

Nous allons exposer ici la manière avec laquelle nous avons conduit cette recherche, ainsi que les choix que nous avons été amené à effectuer. Dans la première section, nous décrirons les caractéristiques du terrain sur lequel nous avons mené nos investigations, à savoir le panel SCAN 5000 de la société Nielsen.

Etant donné que nous étudions la validité d'instruments de mesure et que nous ne disposons pas de référence sûre de comparaison (un mètre étalon en quelque sorte), l'esprit de notre investigation sera *"comparatif"*. Ces comparaisons seront effectuées à trois niveaux : le marché, les variables étudiées et l'indice de mesure.

Dans la deuxième section, nous présenterons les marchés qui ont été étudiés, les marques qui ont été retenues sur ces marchés et enfin, la façon dont nous avons manipulé les variables étudiées.

Dans une troisième section, nous présenterons les indices qui ont été retenus dans cette étude, sachant qu'ils ont tous déjà fait l'objet d'une présentation théorique dans la première partie.

Enfin, la dernière section décrira la mise en oeuvre de la recherche, à savoir les calculs effectués et la mise en forme de ces calculs permettant leur traitement dans les chapitres V et VI.

SECTION 1 LE PANEL SCAN 5000

Cette section est consacrée à la description du panel SCAN 5000 (aujourd'hui devenu SCAN 7000) qui constitue le *"terrain"* à partir duquel nous avons pu mener cette recherche. Nous décrirons successivement l'esprit et l'environnement de ce panel puis son contenu.

SS1 ESPRIT ET ENVIRONNEMENT DU PANEL SCAN 5000

SCAN 5000 est un panel de consommateurs développé par la société NIELSEN (tout ce que nous dirons sur ce panel est aussi valable pour SCAN 7000, son successeur, sauf remarque expresse de notre part). Bien qu'à lui seul cet outil soit une source de recherche et d'aide à la décision immense, il n'existe pas de façon isolée mais s'intègre au sein d'un projet plus large, qui est le concept de "Source Unique" (Single Source). Le concept de Source Unique à pour objet de regrouper tout un ensemble d'informations relatives :

- au marché : ventes, parts de marché, taux d'essai et taux de ré-achat, transferts entre marques
- aux médias : Investissements publicitaires, parts de voix, couverture, répétition, duplication, pige, etc
- au point de vente : caractéristiques de l'offre en magasin : prix, promotions, linéaires, PLV

Le constat à partir duquel la société Nielsen a bâti sa *"Proposition de Single Source"* est que les agrégats classiques observés sur les marchés sont constitués, au niveau élémentaire, de comportements individuels. Pour étudier les comportements à la source, il est nécessaire d'utiliser des modèles de choix du consommateur de type individuel. Or, pour prétendre étudier de tels modèles, il est impératif que "toute" l'information (les intrants) alimentant ces modèles soit collectée à un même niveau d'agrégation, le niveau le plus élémentaire qui est le consommateur (en fait c'est le ménage qui est étudié). Parmi ces intrants permettant d'expliquer le comportement, figurent l'observation des comportements passés, les attitudes, mais aussi l'ensemble des variables de marketing mix auquel est exposé le consommateur. Ces variables comprennent à la fois les éléments du mix des marques achetées par le consommateur, mais aussi les éléments du mix des marques concurrentes qu'il n'a pas achetées, mais auxquels il a été exposé. L'esprit du Système de Source Unique est d'observer au mieux les variables explicatives du comportement afin de *"substituer une observation intelligente à la classique expérimentation"* qui, en cherchant à manipuler ou contrôler des variables d'influence, ne parvient pas à obtenir des résultats transposables dans la réalité. Les informations recueillies au plan individuel peuvent être ensuite agrégées au niveau souhaité, en fonction de contraintes spécifiques à l'étude: un segment particulier de consommateurs, un sous-marché ou encore le marché total. Etant donné que l'information a été collectée sous sa forme la plus élémentaire, elle peut être agrégée à n'importe quel niveau tout en restant pertinente.

La figure suivante présente la correspondance entre les niveaux individuel et agrégé (Source : documentation Nielsen 1989).

FIGURE 2A CORRESPONDANCE ENTRE NIVEAU AGREGE ET ELEMENTAIRE

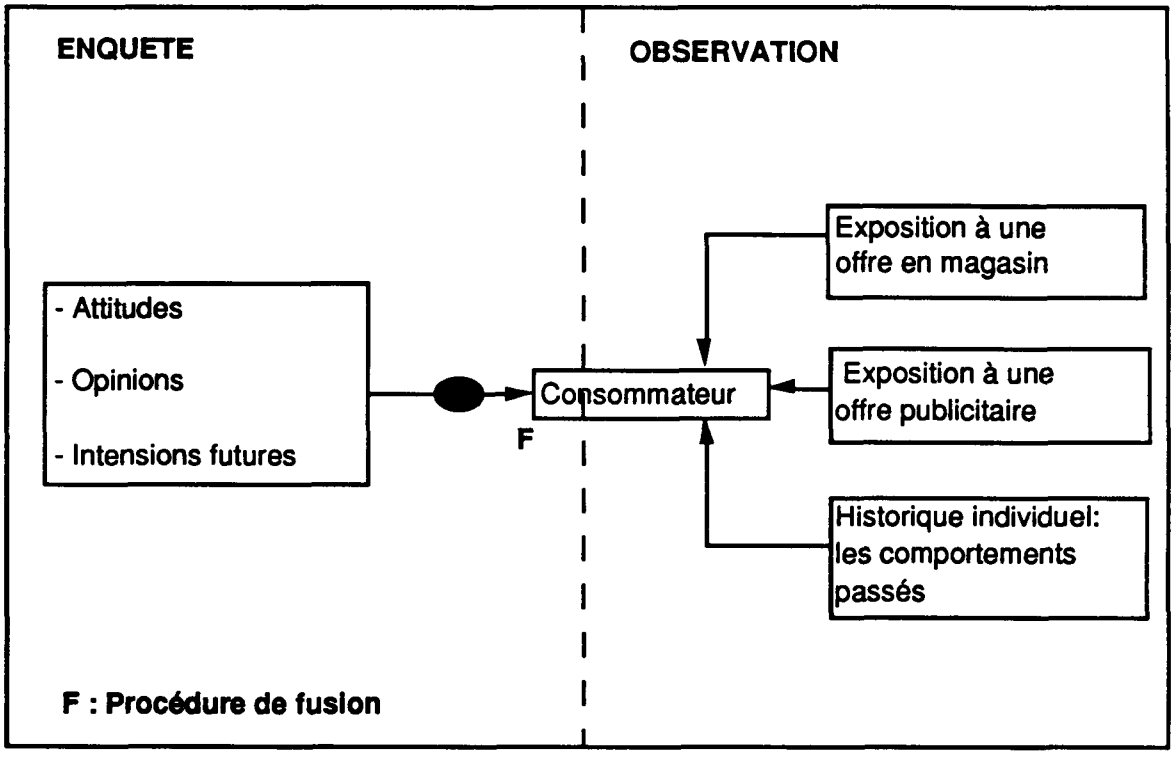
Niveau élémentaire "consommateur"	Niveau agrégés
Exposition à une offre en magasin au moment de l'achat	Informations relatives aux actions promotionnelles en magasin, les prix moyens, etc ...
Exposition à la communication d'un ensemble de marques concurrentes (offre publicitaire)	Informations relatives aux audiences, aux pressions publicitaires (pige publicitaire), ainsi qu'aux "effets" de la publicité
Comportements passés	Informations relatives au marché: parts de marché, etc ...

L'objectif du Système à Source Unique et donc de modéliser (afin d'expliquer, prévoir ou simuler) le comportement du consommateur individuel, à partir des principales variables d'influence, à savoir :

1 Ses Attitudes et Intentions
2 La Pression Publicitaire (au sens large) à laquelle il est confronté à son domicile : Télé, Radio, Presse Ecrite.
3 Ses Comportements passés
4 L'Offre à laquelle il est confronté en magasin

La figure suivante illustre ce principe (Source: documentation interne Nielsen 1989).

FIGURE 2B SCHEMA DU SYSTEME A SOURCE UNIQUE COMPLET



Nous allons commenter successivement chacun de ces blocs de variables d'influence :

1 LA MESURE DES ATTITUDES

L'objectif est de connaître la perception qu'a le consommateur du marché et l'image qu'il a des marques appartenant à ce marché. L'étude des attitudes repose sur la connaissance des éléments classiques suivants :

- connaissance des scénarios d'achat et de consommation des produits la catégorie étudiée
- importance des attributs perceptuels dans le choix
- évaluation des produits selon ces attributs
- choix passés et intentions futures

Le recueil de l'information s'effectue sous la forme d'un nombre constant d'interviews tous les jours. L'utilisation d'une telle procédure permet de suivre en continu l'évolution des attitudes, notamment :

- avant / après une campagne de communication
- après deux campagnes, "nous et la concurrence"

2 OFFRE PUBLICITAIRE

Le deuxième bloc d'influence est constitué par toutes les variables composant *"l'offre publicitaire"*. La Société Nielsen développe un système d'audimétrie installé sur les télévisions de ses panélistes. L'utilisation du système de reconnaissance "Monitor Plus"

fournit très rapidement l'identification nécessaire des messages reçus. Pour chaque foyer il est ainsi possible de connaître l'ensemble des messages publicitaires auxquels il a été exposé à la télévision. La connaissance de l'exposition aux autres médias (Presse, Radio et Cinéma) se fait sur la base de déclarations d'habitudes de fréquentation qui sont suffisamment stables pour ne pas avoir à être répétées de manière trop fréquente.
Les deux blocs d'influence suivants (3 Comportements Passés et 4 Offre Totale) sont les éléments qui sont fournis par SCAN 5000.

3 LE COMPORTEMENT PASSE

C'est une variable privilégiée pour l'explication des comportements présents. A partir des comportements passés, sont établis (au plan individuel ou agrégé):

- des indicateurs classiques de ventes : taux de pénétration, taux d'essai, taux de ré-achat, cycle d'achat, QA/NA etc ..
- des indicateurs dynamiques du processus de choix: fidélité à la marque, transferts entre marques (brand switching) permettant notamment de calculer la probabilité d'achat d'une marque i sachant que le consommateur a acheté la marque j lors de son dernier achat dans la catégorie de produits étudiée (Pi/j).

4 ETUDE DE L'OFFRE TOTALE EN MAGASIN

Lorsqu'une marque i est choisie par un consommateur en magasin, elle est vendue à un certain prix, dispose d'un certain linéaire, fait l'objet ou non de promotion, de PLV, etc ..
Il est important de prendre en compte tous ces éléments pour expliquer le choix du consommateur, mais ceci est encore insuffisant. En effet, le choix final du consommateur est en réalité le résultat du choix d'une marque (voire plusieurs) parmi N marques concurrentes. Il est donc indispensable de connaître aussi les éléments de l'offre des concurrents qui étaient présents en magasin lors du choix de la marque i. On parlera alors de concept *"d'Offre Totale"*.
L'Offre Totale permet d'entrer en rupture avec les modèles de prévision de type monadique, c'est à dire qui ne sont basés que sur les caractéristiques de la marque pour laquelle est effectuée la prévision (le modèle de ce type le plus connu est BASES, développé par la société Burke). Ainsi que nous l'avons déjà dit, la concurrence est par essence relative, il est donc nécessaire de considérer les offres *"relatives"* des marques achetées, relativement aux marques non-achetées. Nous avons d'ailleurs pu constater dans la première partie de la recherche, que les modèles de prévision de parts de marché les plus récents prennent en compte tous les effets croisés (interactions) de chaque élément du mix d'une marque avec celui de ses concurrentes et éventuellement d'autres éléments du mix (voir par exemple les modèles proposés par Carpenter et al 1985, Kamakura et Srivastava 1984, Batsell et Lodish 1981, Batsell et Polking 1985). Les modèles de prévision acquièrent ainsi une dimension explicative qu'ils n'avaient pas avant.

Les quatre types d'information que nous venons de passer en revue correspondent aux deux types d'approches que notre revue de littérature a opposés dans la première partie : les approches jugements (correspondant au bloc 1) et les approches comportements (correspondant aux trois autres blocs). Un des atouts majeurs du Système à Source

Unique est de ne pas opposer les deux approches (comme c'est le cas dans la littérature Marketing!) mais au contraire d'essayer de les intégrer au sein d'une même démarche. Le point névralgique du système est alors la *"fusion"* (voir figure 2B). En effet, les informations relatives au bloc 1 d'une part et 2, 3 et 4 d'autre part, ne peuvent pas pour des raisons évidentes de pollution, être collectées auprès des mêmes répondants. Ainsi, coexistent séparément deux échantillons: un premier échantillon constitué de répondants (pour le bloc 1) et un deuxième échantillon constitué de panélistes (pour les blocs 2, 3 et 4). Il est nécessaire d'opérer une fusion entre les informations en provenance de ces deux échantillons. La fusion consiste à attribuer des valeurs d'attitudes aux panélistes sur la base de leur ressemblance avec des répondants ayant ces valeurs d'attitudes. La fusion s'opère à partir des éléments communs aux deux échantillons :

- comportements passés des répondants et des panélistes
- signalétiques et éventuellement styles de vie des individus appartenant à ces deux échantillons

SS2 DESCRIPTION DU CONTENU DE SCAN 5000

Il existe trois types de panels de consommateurs.

Les panels qui sont constitués pour une catégorie de produits particulière et qui étudient un groupe de magasins distributeurs de cette catégorie de produits. Les panels qui regroupent les principaux points de vente d'une ville représentative du marché national, c'est le cas du panel SCANEL, qui est développé par la société Sécodip. L'objectif de ces panels est d'observer un marché *"instrumental"*, de manière à pouvoir opérer des expérimentations en vraie grandeur (test de packaging, promotion ..). Le troisième type de panel est constitué d'un échantillon représentatif de l'appareil de grande distribution alimentaire national, son objectif est l'obtention et l'estimation de valeurs généralisables au niveau du marché tout entier. SCAN 5000 appartient à cette troisième catégorie.

SCAN 5000 est un panel de 5000 ménages divisé en dix groupes de 500 environ chacun, chaque groupe étant fidèle à un magasin particulier (magasins appartenant à l'échantillon du panel de détaillant SCAN TRACK). La valeur de l'étude d'un marché à partir de l'observation des comportements d'achats des ménages dans un seul magasin repose essentiellement sur une hypothèse, la fidélité du consommateur au magasin dans lequel sont observés ses achats. Cette hypothèse a été validée par plusieurs études qui ont montré notamment que (Source: Documentation interne Nielsen 1989) :

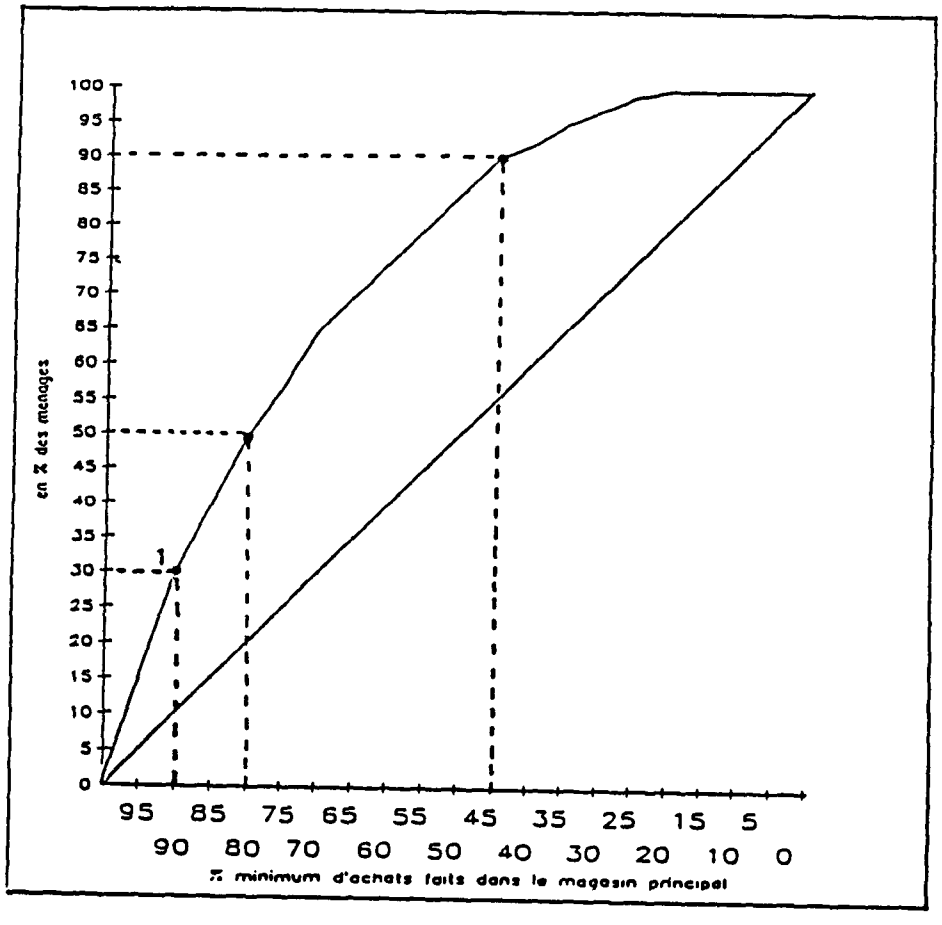
- 94 % des foyers français fréquentent au moins une grande surface .
- un ménage fréquente en moyenne 2.9 GMS dans l'année
- la fréquentation en GMS s'établit comme suit :

exclusivement une GMS	10 %
une GMS principale et une GMS secondaire	20 %
une GMS principale et une GMS secondaire et une GMS occasionnelle	38 %
une GMS principale et une GMS secondaire et deux GMS occasionnelles	32 %

Le graphique suivant exprime la relation entre le pourcentage cumulé d'acheteurs et le pourcentage minimum des achats effectués dans le magasin principal (Source : documentation interne Nielsen 1989).

GRAPHIQUE 1

La distribution de la proportion des quantités achetées dans le magasin principal est résumée dans la courbe de concentration suivante :



Exemple de lecture du point 1 : 30 % des menages font au moins 90 % de leurs achats dans leur GMS principal.

A partir de ce graphique, nous pouvons constater, par exemple, que 50 % des ménages effectuent au moins 80 % de leurs achats dans le même magasin principal.

D'autre part, différentes études complémentaires ont montré :

- que les ménages fidèles à un magasin ont des profils socio-économiques tout à fait voisins de ceux de l'ensemble de la population.
- que leur comportement à l'égard des marques (fidélité, mixité, échanges entre marques, sensibilité au prix ..) est tout à fait comparable à celui de la population

totale. Au regard de ces études, la société Nielsen est en droit d'espérer avoir avec SCAN 5000 un outil fiable pour étudier le lien entre offre en magasin et comportement du consommateur.

Les informations fournies par SCAN 5000 se présentent sous la forme de trois bases de données :

- une base "Actes d'Achats"
- une base "Caractéristiques Panélistes"
- une base "Offre Magasin"

Dans la base *"Actes d'Achats"*, figure, pour chaque individu, l'historique de tous ses achats sur la période durant laquelle il a appartenu au panel. Ces historiques comprennent notamment la date de l'achat, les références achetées (jusqu'au degré le plus fin de la nomenclature), les quantités et le nombre de produits achetés, la forme du produit (par exemple boîte ou bouteille), le magasin et le numéro d'identification du panéliste.

Dans la base *"Caractéristiques Panéliste"* figurent toutes les variables décrivant les caractéristiques principales du panéliste (essentiellement des variables de type socio-démographique).

Dans la base *"Offre Magasin"*, figure, pour chaque jour ouvré, l'offre totale des magasins observés concernant la catégorie de produits étudiée, à savoir :

- la présence (ou non) en rayon
- le linéaire attribué à chaque référence
- le type de promotion présente pour chaque produit :
    - mise en avant
    - offre spéciale
    - type de réduction de prix
- les prix

SECTION2 MARCHES MARQUES ET FACTEURS MANIPULES

Dans cette section, nous allons décrire et justifier successivement :

- le choix des marchés
- le choix des marques
- le choix des facteurs et leur manipulation

SS1 LE CHOIX DES MARCHES

Dans la première partie (théorique) de cette recherche, nous avons mis l'accent sur le fait que la situation d'usage du produit ainsi que la recherche de variété du consommateur concernant la catégorie de produits étudiée sont deux variables qu'il faut considérer lors de l'étude de la concurrence entre les marques. En particulier, nous avons noté que la majorité des indices procédant à partir de données de comportement, indices qui sont étudiés ici, ne permettent pas de les prendre en compte explicitement (particulièrement en ce qui concerne la situation d'usage). D'autre part, nous avons aussi noté que l'implication (l' interaction entre la personne et l'objet) est aussi une variable qui peut avoir de l'importance sur la manière de mesurer la concurrence. Ne pouvant pas, bien sûr, manipuler *"ex ante"* ces variables puisque nous ne disposons que de données de comportements passés, nous pouvions par contre essayer de construire, *"ex post"*, un pseudo-plan d'expérience et choisir les marchés de façon à mettre ce plan d'expérience en oeuvre. Le plan d'expérience aurait alors été le suivant :

GROUPE1 PRODUITS IMPLIQUANTS

		SITUATION D'USAGE	
		Unique	Multiples
RECHERCHE DE VARIETE	Absence	LESSIVE	CAFE
	Présence	BAIN DOUCHE	SHAMPOOINGS

GROUPE 2 PRODUITS NON-IMPLIQUANTS

RECHERCHE DE VARIETE	SITUATION D'USAGE : Unique	SITUATION D'USAGE : Multiples
Absence	CAMENBERT	HUILE
Présence	BISCUITS	BRSA

Les tableaux précédents, présentent les croisements des deux variables situation d'usage et recherche de variété , qui par simplicité, ne présentent que deux niveaux : présence présumée ou absence présumée de l'influence de la variable. Ces deux tableaux représentent un pseudo-plan d'expérience dans lequel sont contrôlées (ex post) les trois variables Implication, Recherche de Variété et Situation d'usage du produit. L'affectation des marchés dans ce plan à été effectuée à partir d'avis d'experts de ces marchés.

Les marchés suivant furent donc sélectionnés afin de couvrir au mieux (sans que les coûts de traitement ne deviennent exorbitants), les cases des deux tableaux précédents :

- Lessives
- Café
- Shampooings
- Boissons rafraîchissantes sans alcool (BRSA)
- Huile

Une partie des autres marchés a dû être abandonnée pour des considérations techniques (nombre de marques trop important, données peu fiables ..). Enfin, suite à une erreur de manipulation, deux bases de données ont été détruites et n'ont pu être reconstituées à temps : le Café et la Lessive . Ceci est d'autant plus dommageable qu'une partie des calculs avait été effectuée sur le marché des lessives . Nos analyses finales porteront finalement sur quatre marchés :

- le marché des BRSA, qui a été dédoublé en deux marchés :
  le marché des Boissons Gazeuses, structuré selon les marques . Ce marché sera appelé BOISSON .
  le marché des Boissons aux Fruits Gazeuses, structuré selon les parfums. Ce marché sera appelé par la suite PARFUM .
- le marché des Huiles de Table qui sera appelé HUILE
- le marché des Shampooings, qui sera appelé SHAMP .

SS2 LE CHOIX DES MARQUES (Unité d'Analyse)

Il fallait tenir compte, pour sélectionner les marques, de plusieurs types de contraintes:

SS2-1 CONTRAINTES D'ORDRE THEORIQUE

Le choix des marques est très important puisqu'il conditionne deux éléments fondamentaux :

1 L'univers de l'étude c'est à dire la définition du marché étudié. Nous avons insisté dans la première partie sur les problèmes liés à une définition arbitraire du marché puisque, pour la majeure partie des indices, cette définition a ensuite une influence sur les calculs et donc les résultats. Par exemple, le calcul des transferts (switching) entre deux marques i et j n'est pas le même suivant l'univers de concurrence (les marques concurrentes) dans lequel sont placées les deux marques. Seuls quelques rares indices ne sont pas sensibles à cette définition (l'indice de Fraser et Bradford et l'indice de Merunka et Bourgeat, voir section suivante).
2 L'unité d'analyse : Le degré de précision avec lequel sont définies les références (Lesieur / Lesieur Tournesol / Lesieur Tournesol 300 cl) lors de la mesure de la concurrence entre les marques doit être cohérent avec les objectifs visés par l'analyse de la structure du marché. Par exemple, sur le marché de la lessive, si la concurrence se joue essentiellement en termes de formes du produit (le consommateur hésite entre une lessive en poudre ou une lessive liquide) qu'en termes de marques (le consommateur hésite entre Ariel et Dash) alors, l'omission de la forme du paquet qui résulte de l'agrégation sous la même marque des deux formes du produit (par exemple SKIP Liquide et SKIP en poudre agrégés en SKIP) risque de gommer *"l'essentiel du jeu concurrentiel"*.

SS2-2 CONTRAINTES TECHNIQUES

Si nous retenons dans l'analyse toutes les références d'un marché (même en s'en tenant aux marques les plus achetées), nous arrivons pour un marché comme celui de la lessive qui ne compte qu'un nombre limité de marques, à près de 280 références, soit une matrice de relations entre marques à 280 x 280 cases. Les temps et les coûts de calculs deviennent rapidement prohibitifs et les interprétations très difficiles. Nous préconisons ici une limite maximale de l'ordre de 25 unités d'analyse.

SS2-3 CONTRAINTES DE FIABILITE STATISTIQUE

Chaque indice doit avoir été estimé à partir d'un nombre suffisant d'actes d'achats et (ou) d'individus et ce, sur une période qui ne soit pas trop longue. En effet, sur de longues périodes d'observation, apparaît le risque de prendre en compte une évolution de la structure même du marché. Ce peut être le cas, par exemple, sur le marché des Colas, lorsque Coca-Cola introduit Coca Light qui vient modifier la structure de la concurrence en y ajoutant une dimension nouvelle : calorique / non calorique. De plus, sur de longues périodes, intervient à coup sûr une modification du champ concurrentiel: certaines

marques (ou références) apparaissent, d'autres disparaissent, ce qui est moins important que le phénomène précédent, mais peut quand même être source de confusions. Pour une période d'un an , avec plus de 20000 actes d'achats et 5000 panéliste , avec 25 marques, la matrice des relations entre marques comprend 625 cases. Il y a alors en moyenne 32 actes d'achats et 8 individus pour estimer le lien entre deux marques. C'est là une limitation sérieuse au nombre de marques analysables de façon simultanée.

SS2-4 CONTRAINTES PRATIQUES

L'analyse d'un trop grand nombre de marques rend difficile la lecture, l'interprétation et en fin de compte la prise de décision. Il est préférable de raisonner sur des unités d'observations agrégées et peu nombreuses pour procéder à des analyses sur un marché défini de manière large.

Le choix des marques sur chacun des marchés a été effectué en tenant compte de ces contraintes et avec l'aide d'experts, Responsables de Clientèle Nielsen. Les petites marques ont été en général agrégées sous l'appellation *"Autres"*. Certaines références d'une même marque ont été agrégées lorsqu'elles sont considérées comme ne représentant pas des *"entités concurrentielles"* réellement perçues et différenciées par le consommateur (ceci, toujours suivant l'avis des experts qui fondent leurs opinions sur les études consommateur qu'ils ont menées par le passé). La description du choix des marques sur chaque marché figure à l'Annexe 1.

SS3 LES VARIABLES MANIPULEES

L'étude de l'influence de variables comme le Revenu, la CSP du chef de famille ou toute autre variable classique (dite aussi de première génération) est à notre avis très importante, ceci pour les raisons suivantes :

- il n'y a pas eu, à notre connaissance, d'études systématiques de l'effet de ces variables sur la mesure de la concurrence à partir de données de panels.
- si ces variables ont une influence forte sur les relations entre marques, donc sur la physionomie générale d'un marché, il serait bon d'en tenir compte, notamment lors des procédures d'échantillonnage. Il s'agit là d'un enjeu méthodologique.
- il est important de savoir de quelles variables dépend l'instrument de mesure qui permet au responsable marketing ou au chercheur d'analyser le jeu concurrentiel. De plus, puisque nous étudions simultanément plusieurs marchés et plusieurs indices de mesure, il est important de savoir si l'influence relative des variables est la même sur tous les marchés ou bien si elle varie d'un marché à l'autre. Il s'agit là d'un objectif théorique et pratique très important.
- enfin, une grande partie de la littérature sur les modèles de choix de type stochastique repose sur une hypothèse souvent critiquée mais peu étudiée au plan empirique: la stabilité des préférences des consommateurs au cours de la période d'étude (voir à la première partie, chapitre I, section 3). Notamment, la stabilité dans le temps des transferts entre marques est une question très importante. Il est donc intéressant d'étudier l'influence de la variable Temps (période d'observation).

Les variables que nous avons étudiées se devaient évidemment d'être disponibles dans les bases de données SCAN 5000 . Nous avons finalement manipulé les variables suivantes :

- Le Revenu du Chef de Famille
- La CSP du Chef de Famille
- La Taille du Foyer
- La Quantité Consommée sur la période d'observation
- Le Lieu de l'Achat (Point de Vente)
- Le Temps, à partir de la période d'observation

Pour étudier l'effet de ces variables, nous les avons "manipulées" de la manière suivante. Chaque variable a été scindée en trois modalités. Par exemple pour le Revenu, nous avons distingué les Gros, Moyens et Petits Revenus. L'échantillon total des actes d'achats est alors découpé en trois sous-échantillons, chaque sous-échantillon correspondant à l'une des modalités de la variable (nous parlerons par la suite de "facteur manipulé"). A chaque modalité d'un facteur manipulé correspond ainsi un sous-échantillon d'actes d'achats. Nous appellerons par la suite "univers", un ensemble d'actes d'achats particulier. Chaque univers est repéré par le nom de la modalité du facteur qui en est l'origine. Par exemple, l'univers constitué par les actes d'achats des panélistes à gros revenus sera appelé REVG. L'univers total des actes d'achats (aucun facteur n'est manipulé) sera appelé GLOBAL. Le tableau suivant présente l'ensemble des modalités des facteurs manipulés ainsi que les noms des univers correspondants.

TABLEAU 1 CODIFICATION DES FACTEURS MANIPULES ET DES UNIVERS CORRESPONDANTS

FACTEUR MANIPULE	CODE	MODALITE	CODE UNIVERS	N°
AUCUN	GLOBAL			1
REVENU	REV	GROS	REVG	2
		MOYEN	REVM	3
		PETIT	REVP	4
CSP	CSP	SUPERIEUR	CSPS	5
		EMPLOYE	CSP6	6
		RETRAITE	CSP7	7
MAGASIN	MAG	HYPER	MAGG	8
		SUPER	MAGM	9
		SUPERETTE	MAGP	10
TAILLE FOYER	TAYF	GROS	TAYG	11
		MOYEN	TAYM	12
		PETIT	TAYP	13
QUANTITE CONSOMMEE	CONS	GROS	CONSG	14
		MOYEN	CONSM	15
		PETIT	CONSP	16
DATE D'ACHAT	DATE	SEMESTRE1	DAT1	17
		SEMESTRE2	DAT2	18

SS3-1 CODIFICATION DES VARIABLES QUANTITATIVES

Se posait alors un problème délicat , celui de la règle de codification de ces variables. Par exemple, la variable Revenu est codifiée en trois modalités: Gros, Moyen et Petit Revenu. Cependant, à partir de quel critère allons nous établir les bornes permettant d'effectuer ce découpage? Ce découpage pouvait en fait se faire selon trois critères :

1- Un découpage à partir de bornes fournies arbitrairement
2- Un découpage à partir de bornes définissant des intervalles de largeur égale
3- Un découpage à partir de bornes définissant des univers contenant des effectifs égaux

D'autre part, les contraintes à respecter étaient doubles :

1 Pour pouvoir comparer l'effet des variables, il fallait observer, pour toutes les variables, le même nombre de modalités.
2 Pour la précision des calculs, il fallait avoir des effectifs comparables dans chaque univers.

Ces deux points nous ont conduit à retenir un découpage en univers d'effectifs équivalents. Là encore, un choix devait être opéré car ces effectifs pouvaient être calculés:

a En termes d'actes d'achats; il fallait alors avoir autant d'actes d'achat dans chaque univers.
b En termes d'individus; il fallait alors avoir autant de panélistes dans chaque univers.

Comme certains indices sont calculés sur la base d'actes d'achat et d'autres sont calculés sur la base d'individus, ce choix était particulièrement difficile.

Finalement, pour que chaque indice soit comparé sur le même univers, nous avons effectué les découpages de manière à avoir un nombre comparable d'actes d'achat dans chaque univers. L'inconvénient lié à ce mode de découpage est que certains univers peuvent se retrouver avec peu d'individus si ceux-ci sont majoritairement des consommateurs "fréquents". C'est le cas dans la strate des gros consommateurs (variable quantité consommée). Cependant, ceci est à notre avis moins gênant que de se trouver en présence d'univers comportant des tailles d'effectifs en "actes d'achats" très différents (par exemple pour l'univers des petits consommateurs CONSP nous aurions alors très peu d'actes d'achats), puisque l'acte d'achat constitue, en fin de compte, l'unité de base du comportement observé.

SS3-2 CODIFICATION DES VARIABLES QUALITATIVES

Pour des variables comme la CSP ou le Magasin, le découpage étant, par nature, déjà effectué dans la base de données SCAN 5000, le problème se posait en terme de regroupement de modalités afin de n'en avoir que trois pour chaque facteur manipulé. Dans la base SCAN 5000, la variable "CSP" comprend huit modalités :

1 Agriculteurs
2 Artisans Commerçants
3 Cadre Supérieurs
4 Cadres Moyens
5 Employés
6 Ouvriers
7 Retraités
8 Sans Activité

Nous avons regroupé cadres supérieurs avec cadres moyens (seul regroupement plausible) et conservé deux autres CSP ,les plus importantes (en terme d'actes d'achats): les employés et les retraités. Les autres CSP (agriculteurs, ouvriers, commerçants artisans, inactifs) n'ont pas été pris en compte lors de la manipulation de ce facteur car elles sont difficiles à regrouper et elles ne comportent que peu d'observations. Nous garderons à l'esprit, lors de l'interprétation des résultats, que les univers définis à partir de la variable CSP, représentent des effectifs un peu moins importants que les autres (voir les effectifs par univers, à l'Annexe 1).

Pour la variable *"Magasin"*, nous disposions des modalités correspondant aux dix magasins observés dans SCAN 5000. Afin d'effectuer des regroupements en trois modalités, nous avons choisi le critère de la surface (le plus courant en matière de

classification de points de vente). Nous avons aussi fait en sorte que ces regroupements donnent lieu à des effectifs comparables pour chaque univers (voir Annexe 1). Les regroupements effectués sont les suivants :

ENSEIGNE	SURFACE EN M2	CODIFICATION
Inter-Marché	900	MAGP (petit)
Inter-Marché	1040	
Pakbo	1200	
Casino	900	
Genty Catiar	1500	
Champion	1800	MAGM (moyen)
Sedico	1800	
Leclerc	4000	MAGG (grand)
Leclerc	3950	

Notons qu'avec la variable *"Magasin"* , nous manipulons finalement au moins deux facteurs:

1 Le point de vente lui même, qui est caractérisé par sa position géographique, les caractéristiques de sa zone de chalandise, etc. L'hypothèse à la base de la manipulation de ce facteur est alors la suivante: "la mesure de la concurrence entre les marques varie-t-elle significativement d'un point de vente à un autre ?".

2 Le type de magasin : Hyper/Super/Supérette, qui caractérise un grand nombre de variables telles que l'assortiment présent, le linéaire, les promotions, le profil des consommateurs, etc.. L'hypothèse à la base de la manipulation de ce facteur est la suivante: "La mesure de la concurrence entre les marques varie-t-elle significativement d'un mode de distribution à un autre ?".

SECTION 3 LES INDICES ETUDIES

Face à la multitude des indices de mesure de la concurrence entre les marques qui ont été relevés dans la littérature (voir première partie, chapitre II), l'idéal était de prélever un échantillon représentatif de ces différentes *"connotations"* du concept de concurrence entre les marques. Pour prélever cet échantillon, nous avons tenu compte des caractéristiques les plus importantes qui distinguent les indices entre eux . Ces caractéristiques (exposées dans la première partie), sont à la base de la distinction entre :

-indices déterministes / indices stochastiques
-indices individuels / indices agrégés (nous utiliserons par la suite les termes *"individu"* et *"ménage"* de manière indifférente, tout en étant conscients qu'en réalité, l'unité d'observation dans les panels est le ménage et non l'individu).

Nous avons choisi les indices étudiés ici, de manière à couvrir *"au mieux"* les quatre cas issus du croisement des deux facteurs précédents Nous allons maintenant présenter un à un chacun de ces indices, sachant qu'ils l'ont déjà été au chapitre II, dans la première partie (section 2 et 3).

SS1 INDICE AGREGE DETERMINISTE

L'indice de base de cette famille est l'indice de Lehmann (1972) qui est, rappelons le, un des premiers auteurs à avoir proposé l'utilisation de données de panels afin de construire des indices de mesure de la concurrence entre les marques. L'indice s'exprime sous la forme :

$$S_{ij} = \frac{N_{ij}}{N_i + N_j}$$

Où :

N_{ij} est le nombre d'individus ayant effectué au moins un transfert de la marque i vers la marque j sur la période étudiée .
$N_{i.}$ est le nombre d'individus ayant acheté la marque i sur la période
$N.j$ est le nombre d'individus ayant acheté la marque j sur la période

Notons que Lehmann proposait aussi une autre version de cet indice avec non pas N_{ij} mais $P_{j/i}$, la probabilité d'acheter la marque j sachant que la marque i a été choisie lors du dernier achat. Cependant, la première version de l'indice (celle que nous retenons ici) lui est apparue comme étant plus fiable (voir Lehmann 1972).

Une caractéristique de cet indice est qu'il utilise l'individu (le ménage) comme unité de mesure et non l'acte d'achat: le fait que cet individu ait effectué un transfert une fois ou dix fois de i vers j n'entre pas en ligne de compte. Les ménages ne sont donc pas pondérés par les quantités (ou fréquences) de leurs achats.

Nous remarquerons de plus qu'au dénominateur, figure le nombre total d'individus ayant acheté les marques i ou j (N_{ij}.+$N.j$) et non le produit $N_{i.}$ x $N.j$ (comme c'est le cas par

exemple pour l'indice de Rao et Sabavala qui sera présenté plus loin). Cet indice devrait ainsi être moins sensible aux variations des parts de marché des marques étudiées puisque la forme de l'indice est additive et non multiplicative. Nous appellerons cet indice "LEM" par la suite.

SS2 INDICE INDIVIDUEL DETERMINISTE

Merunka et Bourgeat (1988) ont proposé un indice construit à partir de la comparaison, pour chaque consommateur, des parts de choix des marques qu'il consomme .

$$S_{ij} = \frac{1}{\sum Q^l} \sum_{l=1}^{l} s_{ij}^l \cdot Q^l$$

$$s_{ij}^l = \frac{\min^l(q_i,q_j)}{\max^l(q_i,q_j)}$$

Où :

q_i et q_j sont les quantités des marques i et j achetées par le ménage l, sur la période d'étude. L'indice s_{ij} (au plan individuel), ne dépend que des quantités achetées par le panéliste concernant les seules marques i et j. Ainsi, le calcul de cet indice ne nécessite pas d'avoir défini au préalable un univers de concurrence .

Q^l est la quantité totale achetée par l'individu l au cours de la période. C'est le poids de l'individu dans l'agrégation des indices individuels. C'est à ce niveau seulement qu'intervient la définition de l'univers de concurrence.

Les indices individuels étant agrégés en pondérant chaque panéliste par sa quantité totale achetée, la structure du marché va être influencée par les gros consommateurs.

On notera que cet indice est une mesure de similarité"standardisée" qui possède les propriétés classiques d'un indice de similarités, notamment: $s_{ii} = 1$, pour tout i

Par rapport aux indices construits sur la base de transferts entre marques, il apporte un élément nouveau qui est l'évaluation de l'intensité de l'acte d'achat, par le biais de la quantité achetée.

Par contre, cet indice ne prend pas en compte la dynamique du comportement d'achat comme c'est le cas avec les transferts entre marques. Les exemples suivants illustrent ce point :

a) i i i j j j

b) i j i j i j

Ces deux historiques d'achats sont très différents d'un point de vue dynamique. Dans l'historique "a", le consommateur apparait comme étant fidèle à la marque j mais change

de marque, par exemple pour rechercher de la variété. Dans l'historique "b", soit les deux produits sont complémentaires avec une période inter-achat équivalente, soit l'individu est caractérisé par une très forte tendance à la recherche de variété.
L'indice proposé par Merunka et Bourgeat permettra de prendre en compte la tendance à la recherche de variété dans le cas "a" puisque les deux marques i et j seront considérées comme parfaitement substituables ce qui ne serait pas le cas avec l'utilisation d'un indice de transfert entre ces deux marques.
Dans le cas b, l'indice calculé montrera encore une très forte substitution entre les deux marques. Cet indice ne prend donc pas en compte la dynamique de l'achat, ce qui est :

- un avantage lorsque cette dynamique est le résultat d'une tendance à la Fidélité / Recherche de Variété, comme dans le cas "a", (le comportement de fidélité / recherche de variété a été souvent caractérisé, voir dans la première partie, chapitre I, section 2).
- un inconvénient, notamment au plan agrégé, puisqu'il ne prend pas en compte les éléments de la dynamique du marché.

D'autre part, cet indice présente une caractéristique intéressante, lorsque l'on raisonne non plus au niveau des couples de marques mais au niveau du marché tout entier. Il ne rend pas nécessaire la définition préalable de l'univers de concurrence. L'exemple suivant illustre ce point.

c) i j i i k i i i j

d) i j i i l i i i j

Les deux historiques précédents sont équivalents, aux marques k et l près. L'indice de Merunka et Bourgeat permettra de calculer des indices individuels sij (entre i et j) qui sont égaux dans les cas c et d. L'indice n'est pas sensible à la variation de la définition du marché figurant dans les cas c et d. Or, nous avons vu que la définition de l'univers de concurrence est un obstacle considérable à l'utilisation des indices fondés sur le concept dynamique de transfert entre marques. Nous avons ici un indice dont le calcul ne dépend pas de la pertinence de cette définition. sij est indépendant du comportement d'achat concernant les autres marques, hormis lors du calcul de la pondération des individus. Celle ci pourrait d'ailleurs être calculée sur la seule base des quantités totales de i et j achetées sur la période par le panéliste l.
Nous appellerons cet indice "MERR".
Nous avons calculé l'indice proposé par Merunka et Bourgeat dans une autre version, en remplaçant qilt (quantité de la marque i achetée par le panéliste l en l'instant t) par une codification binaire :

$q_{ilt} = 1$ si i est acheté en t
$q_{ilt} = 0$ dans les autres cas

Ce faisant, nous comptabilisons uniquement les "actes d'achats de marques", qui ne sont pas pondérés par la quantité de ces achats. Avec un tel indice, ce qui est important pour la mesure de la concurrence est "*l'acte d'achat*" lui même et non le poids de cet acte, qui peut varier en fonction de causes erratiques, notamment, la situation d'usage du produit, une promotion en magasin, etc ..Cet indice sera appelé "MERA".

SS3 INDICE AGREGE STOCHASTIQUE

Rao et Sabavala (1981) ont proposé un indice reposant sur le concept de "matrice de transition inter-périodes". L'indice de concurrence entre deux marques s'exprime de la manière suivante :

$$S_{ij} = \frac{n_{ij}\ n}{n_{i.}\ n_{.j}}$$

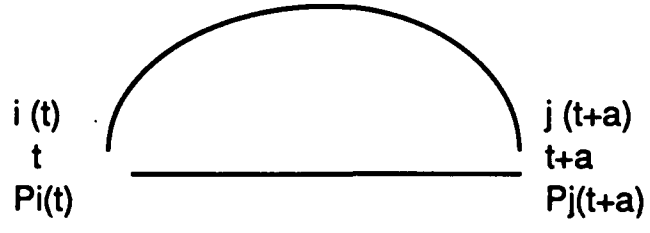
Pour une période d'observation définie entre t et t+a :

- n_{ij} est le nombre d'individus ayant acheté la marque i en t et la marque j en t+a
- $n_{i.}$ est le nombre total d'individus ayant acheté la marque i en t
- $n_{.j}$ est le nombre total d'individus ayant acheté la marque j en t+a
- n est le nombre total d'acheteurs dans le segment homogène étudié

Cet indice met finalement en rapport un effectif observé d'individus ayant effectué un transfert entre deux marques avec un effectif attendu sous hypothèse d'absence de lien concurrentiel entre les deux marques. Plus l'effectif observé n_{ij} est fort par rapport à l'effectif théorique attendu sous hypothèse d'indépendance et plus l'indice de concurrence sera élevé.

PROBLEMES DE MISE EN OEUVRE

Comme le notent les auteurs et nous l'avons vérifié, un calcul basé sur une seule matrice de transition à deux périodes ne présente aucun intérêt, même sur un gros échantillon, car il est très peu fiable et les n_{ij} pour chaque couple de marques sont presque toujours nuls. Ce faisant, pour obtenir des données robustes, les auteurs préconisent de calculer les indices à partir d'une matrice de transfert "moyenne" sur la période d'observation (ici une année). Pour chaque date d'arrivée (t+a) , nous avons calculé le nombre de transferts vers la marque j (noté j(t+a)) ainsi que les parts de marché correspondantes. Le graphique suivant illustre ce point.



$P_i(t)$ est la demande totale pour la marque i à la date t (dans le magasin)
$P_j(t+a)$ est la demande totale pour la marque j à la date

t+a (dans le même magasin)
t est la date "origine du transfert"
t+a est la date "arrivée du transfert"

Au plan agrégé (pour tous les individus et toutes les périodes) ceci revient à calculer pour chaque couple de marque (i,j) , le nombre d'individus ayant effectué un transfert de i vers j sur l'ensemble des périodes de transition. Le même panéliste peut ainsi participer plusieurs fois dans le calcul de n_{ij}. n_{ij} représente finalement le nombre d'actes d'achats de la marque i suivi d'un achat de la marque j. Nous obtenons finalement :

$$S_{ij} = \frac{N_{ij}}{N_{i.}\ N_{.j}}\ N$$

Où :

N_{ij} représente le nombre total de transferts de i vers j sur la période étudiée
N_i est la somme des achats par magasin, de la marque i, les jours de départ de ces transferts .
$N_{.j}$ est la somme des achats par magasin, de la marque j, les jours d'arrivée de ces transferts .
N est le nombre total d'actes d'achats effectué par l'échantillon, dans la catégorie de produits étudiée, sur la période d'observation.

Il est important de noter que lorsque les matrices de transitions sont agrégées, l'unité d'analyse n'est plus l'individu mais l'acte d'achat puisque chaque individu apparait autant de fois qu'il a effectué un transfert entre les deux marques étudiées.
Nous appellerons cet indice "SABA".

De manière à avoir une version de cet indice où l'unité d'analyse n'est pas l'acte mais l'individu, nous avons calculé l'indice suivant, appelé "RAO" par la suite. C'est une version de l'indice SABA dans laquelle les ménages ne sont pas pondérés par le nombre de transferts qu'ils effectuent.

$$S_{ij} = \frac{N_{ij}}{N_{i.}\ .N_{.j}}.N$$

où N_{ij} , $N_{i.}$ et $N_{.j}$ sont calculés, comme pour l'indice de Lehmann, en nombre de ménages ayant effectué au moins un transfert ou un achat sur la période d'observation.

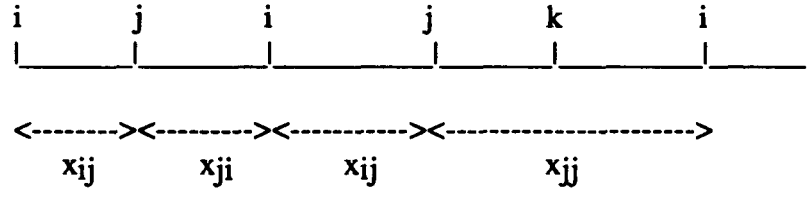
Le dénominateur représente le nombre attendu de panélistes effectuant un transfert entre les marques i et j sous hypothèse d'absence de lien de concurrence entre les marques. Par rapport à l'indice LEM, les indices SABA et RAO devraient être plus sensibles aux variations des parts de marché des marques puisque la forme du dénominateur est multiplicative. Avec l'indice SABA, ces parts de marché (les $N_{i.}$ et $N_{.j}$) sont calculées en actes d'achats, elles subissent donc des variations d'amplitude encore plus forte qu'avec l'indice RAO.

L'indice SABA est ainsi, à priori, un indice plus sensible aux variations du comportement de choix ce qui peut en pratique, constituer une faiblesse si l'indice perd de sa fiabilité parce qu'il est trop sensible.

SS4 INDICE INDIVIDUEL STOCHASTIQUE

Fraser et Bradford (1983, 1984) ont proposé un indice de mesure de la concurrence qui a été fortement critiqué dans la littérature (voir nos commentaires à ce sujet dans la première partie) mais peu étudié au plan empirique. Il est intéressant de confronter cet indice à d'autres indices de mesure de la concurrence. Cet indice fait de plus l'objet d'une notice technique qui figure à l'Annexe 2).

Rappelons que pour un couple de marques (i,j), Fraser et Bradford (F&B) proposent de retracer l'historique (dynamique et quantifié) de ces deux marques et de ces deux marques uniquement. L'indice, tel que le définissent les auteurs dans leur article, ne requiert pas une définition préalable du marché. La notion de transfert perd alors son sens, dans la mesure où, si l'on ne considère que deux marques lors de l'étude d'un historique d'achat. Il y a forcément transfert entre ces deux marques puisqu'elles sont les seules à être considérée! L'exemple suivant permet d'illustrer ce point très important.



Dans cet exemple, l'achat de k , même si la marque appartient réellement au marché étudié, n'a aucune influence sur les périodes inter-achats des deux marques i et j et sur le calcul de l'indice.

ESTIMATION DES PARAMETRES

Pour chaque individu l, nous pouvons estimer:

$$\overline{X_{ijl}} = 1/n \left[\sum_{t} (d_{jt} - d_{it'})/Q_{it} \right] \qquad (1)$$

Où :

$\overline{X_{ijl}}$ est la période inter-achat moyenne entre i et j pour le ménage l, cette valeur ne dépend pas des autres marques achetées par ce ménage.

d_{jt} est la date d'achat de la marque i

$d_{it'}$ est la date d'achat de la marque j, l'occasion d'achat t étant immédiatement antérieure à t'

Q_{it} est la quantité de i achetée en t

n représente le nombre d'achats de i et de j effectués par le ménage étudié, sur la période d'observation.

$\overline{X_{ijl}}$ est un intervalle moyen inter-achat entre les marques i et j, pondéré par l'inverse de la quantité achetée, ce qui permet d'en éliminer l'effet sur la période inter-achat. L'hypothèse qui est faite est que le ménage l utilise deux fois plus de temps pour consommer la quantité 2X que pour consommer la quantité X; il y a proportionnalité entre la quantité achetée et la durée entre deux achats, toutes choses égales par ailleurs (notamment les situations d'usage du produit).

De même, nous calculons V_{ijl}, la variance des intervalles inter-achats de ce même individu.

Pour estimer les périodes inter-achats agrégées sur l'échantillon, nous calculons:

$\overline{\overline{X_{ij}}}$, $\overline{\overline{X_{ii}}}$ et $\overline{\overline{X_{jj}}}$, moyennes des $\overline{X_{ijl}}$, $\overline{X_{iil}}$, $\overline{X_{jil}}$ et $\overline{X_{jjl}}$ sur l'ensemble des panélistes l.

Nous pouvons estimer ensuite ri, le paramètre de forme de la loi gamma qui gère le processus de choix de la marque i :

$$r_i = [1/N \sum_{l=1}^{N} \frac{V_{ijl}}{\overline{X_{ijl}}^2}]^{-1} \qquad (2)$$

l = 1 à N, est constitué par l'ensemble des individus ayant au moins deux intervalles inter-achats entre i et j sur la période (donc au moins 3 achats de i ou de j). Dans leur article, F&B ne retiennent que les individus ayant effectué au moins cinq achats des deux marques, ce qui limitait, à notre avis, trop sérieusement l'échantillon d'étude, notamment pour les marques à faible part de marché. Il est à noter que F&B ont retenu une période d'observation extrêmement longue (quatre années), ce qui leur permet d'estimer les paramètres de l'indice sur des chroniques d'achats très longues et donc d'obtenir des estimateurs précis de ces paramètres. Par contre, nous pouvons douter que sur une période aussi longue, l'hypothèse classique de la stabilité des préférences soit tenable, ne serait-ce que parce que l'offre varie considérablement sur des périodes aussi longues.

Au plan agrégé, l'indice s'écrit :

$$S_{ij} = \frac{(\overline{\overline{X_{ii}}} - \overline{\overline{X_{ij}}}) / \overline{\overline{\overline{X_{ii}}}}}{(r_i-1)/2.r_i} \qquad (3)$$

PROBLEMES PRATIQUES ET EXTENSIONS

1 Contrairement au postulat de base de F&B, l'hypothèse fondamentale sur laquelle repose l'indice n'est pas toujours vérifiée, à savoir:

$$H0 : \quad E(X_{ii}) > E(X_{ij})$$

L'Annexe 2 présente les pourcentages d'intervalles entre deux marques qui vérifient cette hypothèse H0. La formule précédente peut ainsi conduire à l'obtention d'indices négatifs.

D'autre part, nous avons souligné que l'indice proposé par F&B ne prend pas en compte l'univers de concurrence des deux marques, ce qui à notre avis dénature le concept de période inter-achat entre deux marques et supprime le caractère dynamique de la mesure. C'est pourquoi nous proposons le mode de calcul suivant. Les Xij sont estimés non pas sur la base des intervalles inter-achats calculés à partir des seules marques i et j, mais en prenant en compte toutes les autres marques substituts appartenant à l'univers de concurrence (comme c'est le cas pour le calcul des transferts entre marques). Il est ainsi nécessaire, pour calculer l'indice, d'avoir au préalable effectué une définition correcte du marché étudié.
Nous obtenons alors un indice qui présente des points communs avec les indices calculés sur la base de transferts entre marques puisqu'il fait intervenir un historique de type dynamique .

2 F&B standardisent leur indice par $((r_i-1)/2.r_i)$ qui est la valeur de $E(X_{ii}-X_{ij})/ E(X_{ii})$ sous hypothèse d'indépendance entre les deux marques (voir première partie, chapitre II, section 3). Nous regrettons que les auteurs n'aient pas apporté la démonstration de cette égalité qui, théoriquement et empiriquement, se révèle arbitraire. Notamment, lorsque r_i devient fort (valeurs supérieures à 2), les valeurs de l'indice s'éloignent franchement de leur intervalle de variation normal [0,1] . Les ramener de force à ces valeurs (0 ou 1) ne nous paraît pas une solution sérieuse, contrairement à ce que font les auteurs dans leur article. Or, voir Annexe 2, il est fréquent d'observer que la valeur de r_i s'éloigne des intervalles de variations courants de r_i. Notamment, quand la part de marché d'une marque i est faible, les valeurs de r_i deviennent fortes.

D'autre part, une erreur sur l'estimation de ri peut conduire à une forte variation de la valeur de l'indice. Les valeurs calculées risquent fort de dépendre de la qualité de cette estimation. Or, comme les simulations effectuées par les auteurs l'ont montré, cette estimation n'est fiable que lorsque le nombre d'achats de chaque ménage est important. C'est la raison pour laquelle les auteurs effectuent leurs mesures sur une période de quatre ans). F&B ont montré que la qualité de l'estimation est meilleure quand c'est le nombre d'actes d'achats des deux marques par le même panéliste qui augmente plutôt que le nombre de panélistes étudiés (c'est à dire la taille de l'échantillon). Le chercheur ou le responsable marketing ne peut pourtant pas, d'un point de vue pratique, limiter les analyses aux seuls ménages *"gros consommateurs"*.

Nous proposons de calculer un indice non standardisé qui ne dépendra pas de la qualité de l'estimation de ri, paramètre de forme de la loi gamma des intervalles inter-achats.

Soit :

$$Sij = \frac{\overline{X_{ii}} - \overline{X_{ij}}}{\overline{X_{ii}}} \quad (4)$$

Par croisement des points a et b, nous avons finalement calculé les 4 indices suivants.

		STANDARDISATION OUI	STANDARDISATION NON
DEFINITION DU MARCHE	OUI	FRASA	FRASB
	NON	BRAA	BRAB

BRAA correspond à l'indice tel qu'il a été proposé par Fraser et Bradford .
FRAA correspond au même indice que BRAA mais une fois que l'univers de concurrence auquel appartiennent les marques i et j est pris en compte.
BRAB et FRASB sont les versions de BRAA et FRASA estimées avec la formule (4) précédente

Notons au passage que les indices BRAA et BRAB nécessitent d'être calculés, pour chaque couple de marques, sur un ensemble d'historiques d'achats spécifiques à ces seules marques. En effet, nous avons vu que cet indice considère, pour calculer les Xij, que les historiques d'achats sont limités aux seules deux marques i et j. Il est donc nécessaire d'effectuer les calculs pour chaque couple de marques ce qui est très coûteux en temps d'accès et en temps de calcul.

SS5 INDICES CONSTRUITS A PARTIR DE "PSEUDO DISTANCES DU CHI2"

Les indices construits à partir de transferts entre marques (RAO, SABA, LEM ...) souffrent à notre idée des défauts majeurs suivants :

1 L'intérêt qu'il y a à calculer des matrices de concurrence inter-marques est de pouvoir ensuite utiliser des méthodes statistiques permettant une meilleure visualisation du champ concurrentiel ainsi caractérisé (MDS, Typologie ...). Or, pour utiliser ces méthodes, il est préférable de disposer d'indices vérifiant les propriétés suivantes :

A - Symétrie $s_{ij} = s_{ji}$.

En effet, les méthodes classiques sont en général mieux à même d'analyser des indices symétriques. Bien que certaines méthodes permettent de prendre en compte la non-symétrie, (voir première partie, chapitre III, section 2). Les auteurs surmontent souvent cette difficulté en appliquant les méthodes statistiques sur des matrices "rendues symétriques" en effectuant la moyenne de s_{ij} et s_{ji} ou en ne

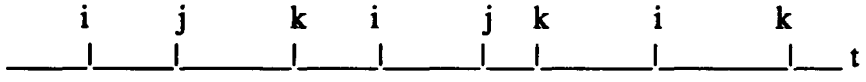
considérant qu'une moitié de la matrice. Ceci est à notre avis une forte réduction du phénomène concurrentiel, dans la mesure où un marché est par essence dynamique et que cette dynamique n'a souvent rien de symétrique.

D'autre part, nous avons souligné que les concepts de substituabilité et de concurrence ne sont pas symétriques. Dans le cas où une marque en attaque une autre sans que cette dernière ne réagisse d'autant, nous dirons que i concurrence plus j que j ne concurrence i. C'est cette dynamique est prise en compte lorsque les intensités concurrentielles sont mesurées à partir des transferts entre marques. Les transferts qui sont calculés sur un marché ont un sens, une direction, de i vers j ou de j vers i. Il est bon, pour mesurer la concurrence, de conserver cette information. Pour toutes ces raisons, il est important de pouvoir conserver l'assymétrie contenue dans les données de transferts.

- B - Une autre propriété souhaitable pour un indice de similarité est d'imposer que la similarité d'une marque à elle-même prenne la même valeur pour n'importe quelle marque et soit plus forte que la valeur de sa similarité avec n'importe quelle autre marque. Cette propriété n'a aucune raison d'être vérifiée pour un indice de concurrence. Le transfert de i vers j peut être proportionnellement plus fort que le transfert de i vers lui même. Il est alors hasardeux de considérer la diagonale d'une matrice d'indices lors de l'utilisation d'une méthode classique d'analyse des similarités.

- C - Les indices calculés peuvent manquer de fiabilité et de robustesse comme le soulignent Totten et Elrod (1988) puisque l'observation simple d'une chronique d'achats peut être trompeuse. La concurrence observée entre deux marques peut être en fait liée à l'introduction erronée de l'une des deux marques dans le marché tel qu'il aura été défini au préalable, comme l'illustre l'exemple suivant.

Exemple: Introduction erronée d'une marque dans l'univers de concurrence



Dans cet exemple, l'introduction erronée de la marque k va nous amener à établir un fort lien concurrentiel entre les deux marques j et k non concurrentes et à sous-estimer le lien entre les marques i et j, ceci à partir de mesures de transferts. Afin d'éliminer ce biais, il peut être intéressant comme le proposent Elrod et Totten de procéder à partir de la comparaison des relations qu'entretiennent ces deux marques avec toutes les autres.

En fait, si i et j présentent chacune le même pattern d'échange avec les autres marques, ceci est le signe que ces deux marques sont concurrentes. Il s'agit d'une mesure indirecte de la similarité entre deux objets: i est similaire à j parce que i et j ont toutes les deux les mêmes niveaux de ressemblance avec les autres objets. Cooper (1988) a présenté une démarche analogue (voir première partie, chapitre III, section 2) dans laquelle l'élasticité croisée entre deux marques i et j est exprimée comme étant le produit vectoriel de la *"Puissance de la marque i"* par la *"Réceptivité de la marque j"*.

- D - Les transferts entre deux marques, attendus sous hypothèse d'absence de concurrence entre ces deux marques sont en général exprimés comme étant proportionnels aux parts de marché des deux marques, ceci est le résultat de la "loi de proportionnalité des échanges" (cette loi est utilisée dans le calcul de l'indice de Rao et Sabavala, 1981). En fait, une telle expression fait implicitement l'hypothèse que ces deux marques bénéficient des mêmes proportions de consommateurs fidèles. Nous pensons que cette hypothèse n'est pas toujours compatible avec la réalité. Il est donc nécessaire de définir les niveaux de transferts attendus sous hypothèse d'absence de concurrence en tenant compte des différences de niveau de fidélité qui caractérisent les marques.

Les points A, B, C et D nous conduisent à proposer un indice prenant en compte les remarques précédentes. Cet indice est calculé en utilisant une *"pseudo-distance du chi2"* comme mesure du lien concurrentiel entre les marques. Soit S la matrice des transferts observés entre cinq marques (codifiées de 1 à 5) sur une période allant de l'instant t (occasion d'achat),en ligne, à l'instant t+n, en colonne, comme dans l'exemple ci-dessous.

EXEMPLE MATRICE DE TRANSFERTS ENTRE 5 MARQUES

1	2	3	4	5	
n11	n12	n13	n14	n15	SR1
n21	n22	n23	n24	n25	SR2
n31	n32	n33	n34	n35	SR3
n41	n42	n43	n44	n45	SR4
n51	n52	n53	n54	n55	SR5
SC1	SC2	SC3	SC4	SC5	

Dans cette matrice :

n_{ij} est le nombre de transferts (en individus ou en actes) de la marque i vers la marque j .
SC_i représente la somme des transferts vers la marque i en provenance de n'importe quelle autre marque (sauf elle même)
SR_i représente la somme des transferts de la marque i vers n'importe quelle autre marque (sauf elle même)

La comparaison des profils, et cela n'est pas sans nous rappeler le principe de l'analyse des correspondances, peut alors être conçue selon une double logique :

1 COMPARAISON DES PROFILS LIGNE

Chaque marque est caractérisée par son profil ligne qui lui même est constitué des transferts de cette marque vers les autres. Nous qualifierons cette approche de logique *"Défense"* puisque chaque marque est comparée sur la base de ce quelle perd par rapport aux autres (ce qui correspond au concept de réceptivité développé par Cooper dans le cadre des élasticité croisées). La concurrence entre les marque 1 et 3 (SD_{13}) s'écrit :

$$SD_{13} = SD_{31} = \sum_{\substack{l \neq 1 \\ l \neq 3}} \frac{1}{SC_l(3)} \left(\frac{n_{1l}}{SR_1} - \frac{n_{3l}}{SR_3} \right)^2$$

l indice de la colonne
n_{1l} nombre de transferts de 1 vers l
$SC_l(3)$ somme de la colonne l lorsque n_{3l} est retiré
$SC_l(3) = SC_l - n_{3l}$

$SC_l(3)$ permet de pondérer chaque terme de la distance par les volumes de transferts de toutes les marques, sauf les marques 1 et 3, vers la marque 1 et donc de prendre en compte l'échange moyen de cette marque avec les autres (et non sa part de marché). En effet, deux marques à parts de marché égales peuvent avoir des taux d'échanges moyens avec les autres marques très différents, en fonction du taux de fidélité qui les caractérisent. Les transferts attendus sous hypothèse d'absence de concurrence doivent donc prendre en compte les différences de niveaux de fidélité qui caractérisent les marques.

SR_l est la somme de la ligne l et représente les achats totaux de la marque l sur la période, moins le terme de la diagonale. SR_1 permet de pondérer chaque terme de différence (n_{1l}-n_{3l}) par l'échange moyen de la marque avec les autres marques.

2 COMPARAISON DES PROFILS COLONNE

Selon cette logique, chaque marque va être caractérisée par son profil colonne, c'est à dire des transferts des autres marques vers elle. Nous qualifierons cette logique de logique *"Attaque"* (ce qui correspond au concept de *Puissance* dans l'approche proposée par Cooper). Ainsi, les marques sont comparées sur la base de leurs forces, c'est à dire ce qu'elles prennent aux autres. Les calculs sont exactement les mêmes que précédemment, une fois que l'on a transposé la matrice S des transferts.

Une telle mesure comporte plusieurs avantages :

• L'indice obtenu a les propriétés d'une distance, il peut être employé sans problème avec des méthodes multivariées classiques. Notamment, il devrait permettre d'obtenir une meilleure qualité de l'ajustement (évaluée par l'indice de Stress) avec l'utilisation des méthodes d'analyse des similarités. En effet, il résout le problème de symétrie précédemment évoqué. L'assymétrie est prise en compte par la double

analyse ligne et colonne et chaque version de l'indice est individuellement symétrique. Toute marque est moins concurrente avec elle même qu'avec n'importe quelle autre: $S_{ij} = 0$ pour tout i.

• Cette mesure de la concurrence devrait être plus robuste que les autres à l'introduction de marques n'appartenant pas à l'univers de concurrence puisqu'elle repose sur des comparaisons de profils de transferts avec les autres marques et non sur les transferts directs. Par contre, l'indice risque d'être affecté par une forte sensibilité à la période d'observation, dans la mesure ou la similarité entre deux marques dépend des relations de ces deux marques avec toutes les autres.

• Cet indice prend en compte le niveau d'échange moyen d'une marque avec toutes les autres. En effet, le transfert observé entre deux marques non concurrentes (attribuable au simple hasard) est défini ici comme étant proportionnel au produit des transferts moyens de ces deux marques avec toutes les autres marques, sauf elle même. Ceci est différent des approches classiques (comme celle de Rao et Sabavala, ou le modèle Hendry ainsi que toutes les modélisations qui s'en sont inspirées) dans lesquelles le nombre d'échanges attendus entre deux marques sous hypothèse du simple hasard est proportionnel aux parts de marché de ces deux marques, c'est à dire aux transferts de chacune de ces deux marques avec toutes les autres, y compris elle même .

Deux types d'input peuvent être utilisés pour calculer cet indice :

- les transferts en "*nombre de ménages ayant effectué au moins un transfert*" et à partir desquels nous pouvons calculer les deux versions de l'indice: "*Défense*" (sur les profils colonnes) et "*Attaque*" (sur les profils ligne) .
- les transferts en "*nombre d'actes d'achats*", à partir desquels nous pouvons encore calculer les deux versions de l'indice (versions pondérées de l'indice précédent). Le tableau suivant résume les cas possibles et donne le nom de code des indices correspondants.

Type de Transfert	Défense	Attaque
Ménage	TOTT	TRTO
Acte d'achat	ELRO	TREL

SS6 INDICES DE TRANSFERTS BRUTS ENTRE MARQUES

Un indice très *"naturel"* de concurrence entre deux marques peut être établi à partir du nombre brut d'échanges observé sur la période, entre ces deux marques. Comme précédemment, nous calculerons ces échanges soit en termes de *"transferts individus"*, l'indice correspondant sera codifié "IND", soit en termes de *"transferts actes d'achats"*, l'indice correspondant sera appelé "ACT".

Ces indices sont particulièrement intéressants parce qu'ils constituent une base de comparaison pour tous les autres indices calculés ici, qui eux sont assez élaborés dans leur calcul et peuvent reposer sur des hypothèses plus contraignantes.

Le tableau suivante présente l'ensemble des indices étudiés ainsi que leur codification.

TABLEAU 2 INDICES CALCULES ET LEUR CODIFICATION

INDICE	CODE	N°
BRAA	BRAA	1
BRAB	BRAB	2
FRASA	FRASA	3
FRASB	FRASB	4
LEHMAN	LEM	5
MERUNKA V2	MERA	6
MERUNKA V1	MERR	7
SABAVALA	SABA	8
TOTT	TOTT	9
TRTO	TRTO	10
RAO	RAO	11
ELRO	ELRO	13
TREL	TREL	14
TRANSFERTS 1	IND	15
TRANSFERTS 2	ACT	16

Il était intéressant de calculer et de comparer des indices prenant en compte la recherche de variété, en particulier sur des marchés caractérisés par un fort niveau de recherche de variété, comme le marché des BRSA. Pour cela, nous avons calculé deux indices proposés dans la littérature et qui ont été exposés avec précision à la première partie, dans une section consacrée aux indices prenant en compte la recherche de variété. Les deux indices choisis étaient :

- l'indice proposé par Lattin et McAlister dans leur article de 1985
- l'indice proposé par Givon, dans la version donnée par son article de 1984, la version 1985 étant trop spécifique pour rentrer dans le cadre de cette étude .

Nous rappelons que ces indices doivent être estimés, pour chaque ménage, en utilisant soit des méthodes de maximisation sous contraintes (pour l'indice de Lattin et McAlister) soit une estimation par la méthode du maximum de vraisemblance (pour

l'indice de Givon). Nous avons calculé ces indices, en utilisant la bibliothèque de programmes de calcul numérique NAG, disponible au CNUSC à Montpellier, qui nous a permis d'effectuer les estimations pour un petit nombre de ménages. Malheureusement, ces calculs se sont avérés extrêmement lourds. Pour chaque ménage, il fallait mener des estimations pour chaque couple de marques, ce qui nous faisait supporter des coûts de traitement prohibitifs pour notre recherche. De plus, il ne semblait pas y avoir de convergence d'un ménage à l'autre et l'agrégation nécessaire à l'interprétation se révélait hasardeuse. Malgré tout l'intérêt que présentent selon nous ces approches, leur mise en oeuvre opérationnelle dans le but de mesurer la concurrence entre les marques, sur des données de panel à large échelle, nous parait impraticable et surtout très difficile à interpréter.

SECTION 4 PRESENTATION DES CALCULS ET DE LEUR MISE EN FORME

Dans cette section, nous allons présenter comment nous avons mis les données en forme afin de pouvoir les traiter dans le cadre des chapitres V et VI.

SS1 CALCULS ET MISE EN FORME DES RESULTATS BRUTS

Les calculs ont été menés en deux grandes phases. Dans la première phase, nous avons construit et extrait les bases de données et construit les programmes nécessaires pour effectuer nos calculs. Dans la deuxième phase, nous avons mis les données dans une forme nous permettant d'effectuer les traitements statistiques utilisés dans les chapitres V et VI.

SS1-1 EXTRACTION DES BASES ET CALCULS

Les bases *'Actes d'achats'* du panel SCAN 5000 sont disponibles dans une forme générale, c'est à dire retraçant les comportements relatifs à toutes les références d'un marché donné (sur le marché des Shampooings, il y a environ 450 références). Figurent dans cette base toutes les unités qui ont été achetées (marques, références). La même marque peut figurer en même temps à divers niveaux d'agrégation. Ainsi on trouvera les actes d'achats correspondants à Cadum 1*30cl et Cadum 2*30 cl mais aussi Cadum qui est l'agrégat des deux références précédentes.

Il fallait donc, dans un premier temps, extraire les unités d'achats correspondant aux univers de concurrence tels qu'ils ont été définis dans la section 2 précédente. D'autre part, certaines références ont fait l'objet de regroupements, notamment les petites marques qui ont été agrégées en une seule sous l'appellation "autres". Ces extractions ont été effectuées en utilisant le gestionnaire de base de données "NOMAD". Tous les indices n'avaient pas besoin d'être estimés directement sur la base de données SCAN 5000 , ce qui est coûteux. Nous avons donc effectué des estimations à partir des bases SCAN 5000 (qui auparavant avaient été mise en forme pour les traitements), pour les seuls intrants nécessaires aux calculs de tous les indices. Une fois stockés ces intrants, nous avons effectué le calcul définitif des indices étudiés ici.

Notons que, bien que tous les intrants nécessaires aient été calculés sur chacun des 18 univers pour chacun des quatres marchés, des pertes de fichiers (inexpliquées et inexplicables) nous ont contraint à ne pouvoir effectuer notre étude que sur une partie des fichiers issus du croisement indices / univers. Notamment, le marché des Shampooings s'est vu dépossédé d'une partie de ses univers de calcul correspondant à la manipulation des facteurs REVENU et CSP.

Nous avons programmé le calcul des indices en utilisant le langage de programmation de la bibliothèque statistique SAS. Tous ces programmes figurent à l'Annexe 3. Nous avons fait ces calculs sur le site informatique de la société Nielsen ainsi qu'au centre de calcul de Montpellier (CNUSC).

Chaque indice calculé sur un univers particulier et un marché particulier était stocké dans un fichier résultat sous la forme d'un vecteur à (n x n) composantes (respectivement 100,

196, 196 et 576 composantes pour les marchés Huile, Boisson, Parfum et Shamp qui comportent 10, 14, 14 et 24 marques).

Le détail des indices calculés, figure à l'annexe 4, pour chaque marché, sous forme de tableaux.

SS1-2 MISE EN FORME DES RESULTATS BRUTS

Pour chaque marché, les résultats se présentent sous la forme de fichiers. Chaque fichier correspond à un indice calculé sur un univers particulier, par exemple l'indice LEM calculé sur l'univers des Gros Revenus du marché Boisson. Nous dirons que le calcul d'un indice pour un marché particulier, sur un univers particulier, constitue une "mesure" particulière. Chaque mesure se présente, nous l'avons dit, sous la forme d'un vecteur à (n x n) composantes pour un marché comprenant n marques.

Puisque les analyses qui vont suivre sont essentiellement comparatives (comparaison des univers, comparaison des indices, comparaison de l'effet des facteurs manipulés sur les indices, etc ..), il nous fallait, pour aller plus loin, définir "le mode de comparaison de deux mesures" ainsi que "la similarité entre deux mesures".

SS1-2-1 DEFINITION DU MODE DE COMPARAISON DE DEUX MESURES

Nos comparaisons pouvaient se faire à deux niveaux :

- au niveau du couple de marques, en effectuant des comparaisons pour chacun des couples de marques.
- au niveau du marché total, en effectuant les comparaisons sur l'ensemble des couples de marques, c'est à dire sur la structure des indices de concurrence entre les marques appartenant au marché.

Nous avons retenu cette deuxième solution car la première est trop spécifique pour garantir une fiabilité et une facilité d'interprétation suffisante. Comme ce qui nous intéresse dans cette recherche est la structure du marché, conçue comme un ensemble de relations entre des marques plutôt que comme une succession de relations des marques considérées deux à deux.

D'autre part, les n x n composantes d'un vecteur de *"mesure"* appartiennent en fait à deux types différents :

- (n x n - n) éléments sont des indices de concurrence entre deux marques différentes i et j
- n éléments sont, par contre, des indices de concurrence d'une marque avec elle même (indices de fidélité à la marque).

Ce deuxième type d'éléments peut prendre une signification théorique et des valeurs numériques bien différentes d'un indice à l'autre. Les valeurs sont systématiquement égales à 0 ou 1 pour certains indices (FRASA, MERA, MERR, TOTT par exemple) ou bien peuvent prendre des valeurs quelconques pour d'autres indices comme par exemple LEM et SABA. Avec LEM, RAO et SABA ces éléments sont une mesure du niveau de fidélité à la marque et peuvent varier considérablement d'une marque à l'autre. Afin de ne pas introduire de biais dans les comparaisons, nous avons supprimé, pour chaque

mesure, les n composantes représentant la relation d'une marque avec elle-même. En effet, si nous comparons les indices FRASA et MERA sur le marché de l'huile à dix marques, ces deux indices ont 10% de leurs composantes égales, quelque soit l'univers de la mesure. Ceci n'est plus vrai si nous comparons les indices LEM et MERA. Les comparaisons ont ainsi été effectuées sur la base des ((n x n)-n) composantes restantes.

SS1-2-2 DEFINITION DE LA SIMILARITE ENTRE DEUX MESURES

Pour effectuer des comparaisons entre les mesures telles que nous venons de les définir, il fallait disposer d'une mesure de similarité. Le coefficient de corrélation entre deux vecteurs de mesure constitue un indice de similarité qui se prête bien à notre problème. Le plus utilisé de ces coefficients est le coefficient de corrélation linéaire de Pearson, que nous avons utilisé. Cependant, le lien entre deux mesures peut exister tout en étant non linéaire ce qui nous amènera à utiliser, parallèlement, une mesure de lien *"monotone"*:: le coefficient de rang de Spearman qui ne considère pour comparer deux vecteurs, que les relations d'ordre entre leurs composantes.

SS2 CREATION ET DESCRIPTION DU CONTENU DES FICHIERS DE CORRELATIONS

Afin de prendre en compte toutes les comparaisons possibles entre les différentes mesures effectuées, pour chacun des marchés, nous avons calculé toutes les corrélations possibles entre les mesures effectuées. Nous avons ainsi constitué quatre fichiers (un par marché) dans lesquels chaque ligne (enregistrement) correspond à une corrélation entre deux mesures. Par exemple, pour le marché de l'huile où nous avons effectué 198 mesures, il a été constitué un fichier comprenant 19503 observations (198 x 197/2). Chaque observation dans un fichier correspond à la corrélation entre deux mesures (couple de corrélation). Chaque enregistrement comporte ainsi trois zones :

- une première zone qui rassemble tous les éléments caractérisant le premier élément du couple de corrélation.
- une deuxième zone, parfaitement symétrique à la première, qui comprend tous les éléments caractérisant le deuxième élément du couple de corrélation.
- enfin, dans la troisième zone, figure la valeur du coefficient de corrélation qui a été calculé, selon deux versions: coefficient de Pearson (noté CORRA) et coefficient de Spearman (noté CORRB).

Schématisation d'un enregistrement :

{ ZONE 1 } { ZONE 2 } { ZONE 3 }

Nous allons maintenant décrire précisément le contenu de ces zones. Les zones 1 et 2 comprennent tous les éléments permettant de décrire une mesure particulière c'est à dire l'univers du calcul avec la codification des facteurs manipulés (ou non manipulés) et l'indice utilisé pour effectuer la mesure. Rappelons que par "univers" nous entendons un ensemble d'actes d'achats sur lequel les indices sont calculés. Si le facteur n'est pas

manipulé, l'univers est appelé GLOBAL. Il comporte alors la totalité des actes d'achats du marché étudié, sur une période de un an. Si un facteur particulier est manipulé, l'univers porte alors le nom de la modalité du facteur manipulé. Par exemple, si on a manipulé le facteur Revenu, l'univers REVG (gros revenu) comprend tous les actes d'achats qui ont été effectués par les panélistes caractérisés par cette modalité de la variable Revenu (les panélistes dont le revenu est supérieur à 9000 francs).
Les variables suivantes figurent successivement dans chacune des deux zones 1 et 2.

U (Univers): Variable alphanumérique donnant le nom de l'univers ayant servi de base au calcul de l'indice. Nous avons repris la codification des univers en dix huit modalités qui figure à la table 1 de la section 2 précédente, par exemple REVG pour l'univers des gros revenus, GLOBAL pour l'univers complet (aucun facteur n'a été manipulé).

F (Facteur) : Variable alphanumérique donnant le nom du facteur manipulé.

REV1 : Variable nominale à 4 catégories: les trois modalités (GROS, MOYEN, PETIT) et une quatrième: GLOBAL, caractérisant la non-manipulation du facteur REVENU.Nous avons donc :

1 REVG
2 REVM
3 REVP
4 GLOBAL

CSP : Variable nominale à 4 catégories : les trois modalités CADRE, EMPLOYE, RETRAITE et une quatrième, GLOBAL, caractérisant la non-manipulation du facteur CSP. Nous avons donc :

1 CSPS (CADRES)
2 CSP6 (OUVRIERS)
3 CSP7 (RETRAITES)
4 GLOBAL

MAG : Variable nominale à 4 catégories : les trois modalités GRANDE, MOYENNE, PETITE SURFACE et une quatrième, GLOBAL, caractérisant la non-manipulation du facteur MAGASIN. Nous avons donc:

1 MAGG
2 MAGM
3 MAGP
4 GLOBAL

TAYF : Variable nominale à 4 catégories : les trois modalités GRAND, MOYEN, PETIT du facteur TAILLE DU FOYER ainsi qu'une quatrième, GLOBAL, caractérisant la non-manipulation du facteur. Nous avons donc:

1 TAYG
2 TAYM
3 TAYP
4 GLOBAL

CONS : Variable nominale à 4 catégories : les trois modalités GROS, MOYEN, PETIT du facteur QUANTITE CONSOMMEE ainsi qu'une quatrième, GLOBAL, caractérisant la non-manipulation du facteur. Nous avons donc:

1 CONSG
2 CONSM
3 CONSP
4 GLOBAL

DATE : Variable nominale à 3 catégories : les deux modalités DATE1 et DATE2 ainsi qu'une troisième, GLOBAL, caractérisant la non-manipulation du facteur Temps. Nous avons donc:

1 DATE1
2 DATE2
3 GLOBAL

I (INDICE) : Cette variable caractérise l'indice qui a été utilisé pour mesurer la concurrence entre les marques, sur l'univers défini par les variables précédentes (on retiendra qu'un seul facteur est manipulé à la fois pour chaque univers). Nous avons repris la codification des indices présentée à la section 3 précédente (table 2).

Selon qu'elles caractérisent le premier élément du couple de corrélation (zone 1) ou le deuxième (zone 2), toutes les variables précédentes ont été affectées du suffixe 1 ou 2 .

Enfin, dans la zone 3, figurent les valeurs numériques des corrélations, CORRA et CORRB (la précision est de deux chiffres après la virgule) qui sont respectivement les coefficients de corrélation de Pearson et de Spearman entre deux mesures.

Un enregistrement se présente finalement de la manière suivante (un exemple figure à l'annexe 4) :

U1 F1 REV1 CSP1 MAG1 TAYF1 CONS1 DATE1 I1 ... } Zone 1

U2 F2 REV2 CSP2 MAG2 TAYF2 CONS2 DATE2 I2 ... } Zone 2

.....CORRA CORRB } Zone 3

SS3 CLASSIFICATION DES CORRELATIONS ANALYSEES

Chaque corrélation peut être considérée à la fois, soit comme une mesure de la similarité entre deux indices (chaque indice étant calculé sur un univers particulier), soit comme une mesure de la similarité entre deux univers sur lesquels ont été calculés les indices ---- Dans cette optique, nous allons qualifier les types de corrélations qui sont contenues dans les fichiers que nous venons de décrire .

1 CORRELATIONS ENTRE INDICES

Pour chaque couple de corrélation, peuvent se présenter deux cas :

- soit les deux indices I1 et I2 du couple sont identiques, nous parlerons alors de corrélation intra-indice
- soit les deux indices I1 et I2 sont différents, nous parlerons alors de corrélation inter-indices .

Toutes choses égales par ailleurs, en particulier pour les univers sur lesquels sont calculés les indices, cette corrélation exprime le degré de ressemblance existant entre les deux indices.

2 CORRELATIONS ENTRE UNIVERS

Lorsque deux mesures effectuées sur deux univers U1 et U2 sont comparées, il est possible de rencontrer trois types de corrélations :

- corrélations intra-univers: c'est le cas lorsque les deux univers de calcul sont identiques (U1=U2). La même modalité de facteur a été manipulée de part et d'autre du couple.
- corrélations intra-facteur: c'est le cas lorsque le même facteur a été manipulé de part et d'autre mais selon des modalités différentes (par exemple GROS et MOYENS Revenus). Ces deux univers sont alors complètement différents, leur intersection est nulle en terme d'actes d'achats. Nous les qualifierons d'univers complémentaires.
- corrélations inter-facteurs : c'est le cas lorsque deux facteurs différents ont été manipulés de part et d'autre du couple de corrélation (par exemple U1 est constitué des Gros Revenus et U2 est constitué des Gros Consommateurs). Ces deux univers, contrairement au cas précédent, ne sont pas complètement différents, ils comportent une partie de leurs actes d'achats en commun (les actes d'achats des panélistes à la fois Gros Consommateurs et Gros Revenus). L'importance de cette partie commune dépend de la similitude entre les deux modalités des facteurs. Si par exemple U1=REVG et U2=CSPS, nous pouvons nous attendre à ce que ces deux univers se recoupent fortement.

Aux deux types de corrélations entre indices peuvent correspondre les trois types de corrélations entre univers. L'ensemble des cas possibles est représenté dans le tableau suivant à six cases.

TABLEAU 3 TYPES DE CORRELATIONS ETUDIES

		MEME FACTEUR MANIPULE		FACTEURS DIFFERENTS
		F1 = F2	F1 = F2	F1 ≠ F2
		U1 = U2	U1 ∩ U2 = Ø	U1 ∩ U2 ≠ Ø
INDICES	I1=I2	CASE1	CASE2	CASE3
	I1≠ I2	CASE4	CASE5	CASE6

Nous allons décrire une à une les six cases de ce tableau, ces cases jouant, en effet, un rôle fondamental dans les analyses qui vont suivre (notamment dans le chapitre VI consacré à l'étude de l'effet des facteurs).

CASE1 Elle est constituée de corrélations mesurant la similarité entre deux univers identiques, à l'aide de deux indices identiques: toutes les corrélations y sont égales à 1. Cette case ne présente donc aucun intérêt.

CASE2 Elle est constituée de corrélations exprimant le degré de similarité entre deux mesures du même indice (I1=I2) effectuées sur deux univers différents mais qui sont issus de la manipulation du même facteur. Ces corrélations sont du type "*intra-facteur et intra-indice*". Elles permettent de mesurer la similarité entre les deux univers, puisque l'indice utilisé est le même. C'est l'effet du facteur manipulé sur la structure du marché qui est mesuré avec de telles corrélations:"*Est-ce que la forme des relations entre les marques reste identique lorsque l'on effectue la mesure sur les actes d'achats de deux catégories de panélistes (par exemple Gros et Moyens Revenus) ? "*. C'est à partir de ce type de corrélations que nous étudierons l'effet des facteurs manipulés sur la structure des marchés. Ces corrélations sont particulièrement intéressantes parce qu'elles permettent d'analyser l'effet dû à l'univers sur lequel est calculé l'indice, une fois que l'effet de l'instrument de mesure est contrôlé (I1=I2).
Nous appellerons par la suite cet effet : "*effet de structure*". C'est la part dans la mesure qui est due à l'objet mesuré lui même. Nous nommerons les corrélations de la case 2 "*corrélations intra-indice intra-facteur*".

CASE3 Elle est constituée de corrélations exprimant le degré de similarité de deux mesures effectuées à l'aide du même indice (I1=I2) sur deux univers U1 et U2 partiellement différents. Ces corrélations permettent de comparer la structure d'un marché pour deux univers issus de la manipulation de deux facteurs différents (par exemple Gros Revenus et CSP Supérieure). Nous les appellerons "*corrélations intra-indice et inter-univers*". Etant donné que les deux univers correspondant à cette mesure ne sont ni identiques ni complètement différents, ces corrélations correspondent à une situation hybride et seront peu utilisées dans les chapitres suivants.

CASE4 Elle est constituée de corrélations exprimant le degré de similarité de deux mesures effectuées à l'aide de deux indices I1 et I2 différents, sur un même univers. Nous mesurons ici la similarité de deux indices (deux instruments de mesure) lorsque les

calculs sont effectués sur le même univers. Tout comme la case2, cette case est particulièrement intéressante car elle permet de mesurer l'effet dû à l'instrument de mesure, une fois que l'effet dû à l'objet mesuré (l'univers) est contrôlé. Nous appellerons l'effet dû à l'instrument *"effet de mesure"*. Il correspond à la partie de la mesure qui n'est pas due à l'objet mesuré mais aux instruments utilisés pour effectuer la mesure. L'analyse de ces corrélations permet de répondre à la question "obtient-on la même structure de relations entre les marques en utilisant différents indices?". Ces corrélations seront appelées "corrélations inter-indices et intra-univers".

CASE5 Elle est constituée de corrélations exprimant le degré de similarité entre deux mesures effectuées sur deux univers U1 et U2 différents (issus de la manipulation de deux modalités différentes du même facteur) à l'aide de deux indices différents. Une telle similarité prend en compte deux effets : l'effet de l'univers et l'effet de l'instrument de mesure. Nous appellerons ces corrélations *"corrélations inter-indices et intra-facteur"*.

CASE6 Elle est constituée de corrélations mesurant la similarité entre deux indices différents calculés sur deux univers partiellement différents (qui sont issus de la manipulation de deux facteurs différents). Nous les appellerons:*"corrélations inter-indices et inter-facteurs"*. Comme les corrélations de la case trois, leur caractère hybride fait qu'elles seront peu étudiées.

Ce sont les cases deux, quatre et cinq qui sont les plus intéressantes et qui recevront le plus d'attention dans les analyses ultérieures, notamment au chapitre VI lors de l'analyse de l'effet des facteurs.

SS4 ANALYSE DESCRIPTIVE DU CONTENU DES FICHIERS DE CORRELATIONS

Nous allons présenter ici les résultats des traitements de base que nous avons effectués sur les quatre marchés. Le contenu de ces fichiers qui serviront de base à toute nos analyses peut être contrôlé en consultant les tableaux de tri à plat qui ont été effectués sur les variables de ces fichiers. Ces tableaux figurent à l'annexe 4. Y figure le nombre de manipulations qui ont été effectuées pour chaque facteur et chaque indice. Nous retiendrons que la taille importante de ces fichiers et l'importance du nombre de manipulations nous permettent de justifier l'utilisation qui sera faite de méthodes statistiques contraignantes en matière de taille d'échantillon (en particulier l'analyse de la variance).

SS4-1 ETUDE DES TENDANCES CENTRALES DES CORRELATIONS

Certains indices sont des mesures de distance, comme par exemple les indices TOTT, TRTO, ELRO et TREL. D'autres sont des mesures de similarité, comme par exemple LEM. Les corrélations entre ces deux types d'indices sont naturellement négatives. Afin d'éliminer l'effet de signe dû à ce phénomène, lors des calculs de moyennes, nous avons utilisé une nouvelle variable : ACORR1 qui est la valeur absolue de la variable CORR1 (valeur du coefficient de corrélation linéaire de Pearson). C'est cette variable qui sera la plus souvent utilisée car elle permet de mesurer la *"force"* du lien entre deux mesures quel qu'en soit le sens.

Sur chacun des marchés, nous avons étudié le niveau moyen des corrélations dans les différentes cases du tableau 4 suivant.

TABLEAU 4 MOYENNES PAR TYPES DE CORRELATIONS

		UNIVERS		
		QQ	F1=F2	U1=U2
INDICES	QQ *	1	2	3
	I1=I2	4	5	6

* QQ signifie quelconque

Case1 : corrélations quelconques : toutes les corrélations
Case2 : corrélations intra-facteur
Case3 : corrélations intra-univers
Case4 : corrélations intra-indice
Case5 : corrélations intra-indice et intra-facteur

Les résultats figurent à l'annexe 5 mais sont résumés dans les tableaux 4 suivants. Nous pouvons constater que la structure des moyennes de ACORR1 est identique quel que soit le marché. Il en est de même pour les écarts-types correspondants (hétérogénéité de la distribution des corrélations). Les distributions des corrélations sont donc similaires d'un marché à l'autre.

La comparaison des valeurs des moyennes figurant dans les différentes colonnes, pour une même ligne, permet d'analyser l'effet dû aux différents types d'univers de calcul (corrélations quelconques, intra-facteur et intra-univers).

La comparaison des valeurs des moyennes figurant dans les deux lignes, pour une même colonne, permet d'analyser l'effet dû à l'indice de mesure utilisé. Sur chaque marché, nous constatons que le niveau moyen des corrélations augmente sensiblement lorsque les deux indices sont égaux, quelque soit par ailleurs le type d'univers. Par contre, il n'y a pas de différences aussi fortes quand on compare, pour une même ligne, les valeurs figurant dans les différentes colonnes. Lorsque les deux univers sont identiques, le niveau moyen des corrélations augmente mais dans une proportion assez faible (par exemple .30 par rapport à .28 sur le marché des shampooings).

Nous retiendrons de ces analyses que le niveau de similarité des mesures dépend plus des instruments de mesure utilisés (les indices) que des objets mesurés (les univers).

Nous reviendrons plus en détails sur ce point dans la section 4 du chapitre VI qui est consacrée à l'étude comparative entre l'effet de mesure et l'effet de structure.

Nous pouvons enfin souligner que l'effet dû à l'indice et l'effet dû à l'univers sont remarquablement stables d'un marché à l'autre. Les tableaux suivants illustrent toutes ces observations pour les quatre marchés.

TABLEAU 4α HUILE

INDICES	UNIVERS		
	QQ	F1=F2	U1=U2
QQ *	.192	.181	.223
I1=I2	.340	.312	-

TABLEAU 4ß BOISSON

INDICES	UNIVERS		
	QQ	F1=F2	U1=U2
QQ *	.168	.150	.220
I1=I2	.310	.260	-

TABLEAU 4γ PARFUM

INDICES	UNIVERS		
	QQ	F1=F2	U1=U2
QQ *	.20	.18	.26
I1=I2	.360	.30	-

TABLEAU 4π SHAMPOOINGS

UNIVERS

INDICES	QQ	F1=F2	U1=U2
QQ *	.15	.12	.22
I1=I2	.30	.21	-

SS4-2 ETUDE DE LA DISTRIBUTION DES CORRELATIONS

Après avoir étudié les tendances centrales dans les cas correspondant aux cinq cases des tableaux précédents, nous allons étudier la distribution des valeurs de ACORR1 selon ces mêmes cinq cas. Notons tout d'abord (et nous garderons cette remarque à l'esprit), que la distribution des corrélations, quel que soit le cas ou le marché étudié, a fortement l'allure d'une loi Normale. Moyennes et médianes sont égales et la distribution est bien symétrique autour de sa tendance centrale. Les tests montrent une normalité fortement significative.

Nous pouvons étudier, à partir de ces distributions, les pourcentages de corrélations non significativement différentes de zéro. De telles corrélations indiquent que deux mesures effectuées sur le même marché, dans des conditions différentes (avec des indices et sur des univers qui peuvent varier), aboutissent à des patterns de concurrence entre les marques complètement différents. Les valeurs retenues (au seuil alpha 5%) en dessous desquelles les corrélations sont jugées non significatives sont les suivantes :*nj*

MARCHE	N	Cthéo
HUILE	90	.20
BOISSON	182	.15
PARFUM	182	.15
SHAMP	552	.085

N : nombre de composantes dans les vecteurs corrélés.
Ctheo : valeurs seuil des corrélations (au seuil alpha égale 5 %)

Les tableaux 5 (α à γ) qui suivent présentent les résultats obtenus sur les différents marchés. Les pourcentages des corrélations non significativement différentes de zéro sont importants, sur les quatre marchés. On peut observer que c'est sur le marché BOISSON que l'on trouve le plus faible pourcentage de corrélations non significativement nulles. Viennent ensuite les marchés SHAMP et PARFUM. Le marché HUILE est celui où l'on trouve le plus faible pourcentage de corrélations significatives (53.1 contre 82.9 pour les BOISSONS). Ceci est en partie explicable par le fait que le nombre de marques n'étant que de dix sur ce marché, la valeur seuil Cthéo est nettement plus élevée que sur les autres

marchés. Le pourcentage des corrélations significativement non nulles (%ACORRSS) augmente sensiblement lorsque les deux indices sont égaux (par exemple, il passe de 53.1 à 77.7 sur le marché de l'huile). Ce pourcentage augmente aussi, mais de façon moins sensible, lorsque les deux univers sont identiques (il passe de 53.1 à 66.6) sur le marché de l'huile.

Ceci vient confirmer les résultats qui figurent dans les tableaux 4 relatifs aux tendances centrales, à savoir que les niveaux des corrélations sont plus le fait des indices comparés que des univers comparés. Les effets respectifs de l'indice et de l'Univers sur la mesure sont identiques d'un marché à l'autre. Cette constatation est encourageante car les problèmes liés à la mesure semblent se répéter à l'identique d'un marché à l'autre.

TABLEAUX 5 ANALYSE DES DISTRIBUTIONS DES CORRELATIONS

- **CORREL** POURCENTAGE DES CORRELATIONS NEGATIVES
- **CORRN** POURCENTAGE DES CORRELATIONS NEGATIVES QUI SONT SIGNIFICATIVEMENT NON-NULLES AU SEUIL DE 5%
- **CORRP** POURCENTAGE DES CORRELATIONS POSITIVES QUI SONT SIGNIFICATIVEMENT NON-NULLES AU SEUIL DE 5%
- **ACORR** POURCENTAGE DES CORRELATIONS SIGNIFICATIVEMENT NON-NULLES, A 5%

TABLEAU 5α HUILE

VARIABLE	∀ I ∀ F	I1=I2 ∀ F	I1=I2 F1=F2	∀ I F1=F2	∀I U1=U2
CORREL	40.5	23.6	17.3	35.9	37.2
CORRN	30	33.2	10	27.1	27.8
CORRP	39.4	71.2	65.7	36.8	46.5
ACORR	33	60.9	66.1	33.3	37.8

TABLEAU 5ß BOISSON

VARIABLE	∀ I ∀ F	I1=I2 ∀ F	I1=I2 F1=F2	∀ I F1=F2	∀I U1=U2
CORREL	42.5	12	14.3	41.4	42.3
CORRN	42.3	33.3	NS	40	49.2
CORRP	50	78.3	NS	45.4	57
ACORR	46.7	72.9	72.5	43.2	53.7

TABLEAU 5γ PARFUM

VARIABLE	∀ I ∀ F	I1=I2 ∀ F	I1=I2 F1=F2	∀ I F1=F2	∀I U1=U2
CORREL	51.6	18.7	16.2	49.2	47.3
CORRN	49.7	53.1	NS	42.9	49.1
CORRP	51.7	79.9	NS	46.2	59.3
ACORR	50.4	74.9	65.6	45.7	54.5

TABLEAU 5π SHAMPOOINGS

VARIABLE	∀ I ∀ F	I1=I2 ∀ F	I1=I2 F1=F2	∀ I F1=F2	∀ I U1=U2
CORREL	39.4	9.4	10	40	35.7
CORRN	53.8	NS	NS	48	69.7
CORRP	60	NS	NS	52.4	69.4
ACORR	57.5	81.3	71.4	50.6	69.5

CHAPITRE V ETUDE DE VALIDITE DE LA MESURE DE LA CONCURRENCE ENTRE LES MARQUES

Depuis fort longtemps, les chercheurs en Science Sociale ont pris conscience qu'ils effectuent, à titres divers, des *"mesures"* et donc que pour cela ils utilisent des *"instruments de mesure"*. Utilisant des instruments de mesure, ces chercheurs ont développé des méthodes permettant d'apprécier les qualités de fiabilité et de justesse des instruments qu'ils utilisent Nous avons montré, dans la partie théorique de cette recherche, que l'analyse de la structure des marchés peut être découpée, d'un point de vue méthodologique, en trois phases à savoir :

a- La définition des frontières du marché
b- Le choix d'une mesure de la concurrence entre les marques
c- L'analyse de la structure sous-jacente aux relations de concurrence mesurées, cette phase donnant généralement lieu à des représentations cartographiées du marché .

A partir d'un tel découpage, l'aspect *"mesure"* apparait comme étant fondamental dans la phase b de cette méthodologie. De plus, il est bien évident que de la qualité de cette mesure dépend la qualité des structures qui pourront être caractérisées dans la phase c.

D'un point de vue théorique et pratique, il est donc difficile de contourner les problèmes de mesure lors de l'analyse de la structure d'un marché, du moins quand on adopte une démarche méthodologique proche de celle que nous avons proposée ici. La discipline qui étudie les problèmes de mesure spécifiques aux phénomènes psychologiques (la Psychométrie) a développé tout un corps de théories et de techniques que nous nous proposons d'appliquer au domaine de l'analyse de la structure des marchés et notamment à la mesure de la concurrence entre les marques.

Mais avant d'aller plus loin, il est bon de définir ce que la psychométrie entend par "mesure".

Bohrnstedt et Borgatta (1980, p24) définissent la mesure comme étant *"l'association de nombres à des objets suivant certaines règles"*. Cette définition, qui est couramment retenue, pèche pourtant par son excès de généralité. En effet, entre un objet et sa mesure, il n'existe pas une relation bi-univoque, la mesure ne définit pas l'objet mesuré.

Nunnally (1978 p3) nous propose une définition très pertinente de la mesure comme consistant en *"règles pour assigner des nombres aux objets de manière à rendre compte des quantités d'attributs possédées par ces objets"*. Une telle définition souligne que c'est bien une *"caractéristique particulière de l'objet"* (un attribut) qui est mesurée et non l'objet lui même. Ainsi définie, la mesure a pour but de quantifier la quantité d'un attribut possédée par un objet. De la même manière qu'un objet n'est pas défini par sa mesure, la relation entre deux marques n'est pas définie une fois qu'une mesure est effectuée. La

relation entre deux marques peut en effet être étudiée selon divers angles et si le chercheur souhaite effectuer une telle mesure, il est nécessaire d'avoir préalablement défini l'attribut objet de la mesure. Le concept de "concurrence entre les marques" est largement utilisé dans la littérature, mais il apparaît que peu de tentatives ont été faites pour essayer de mettre en place en face de ce concept une définition claire et unique, permettant d'opérationnaliser des mesures elles aussi uniques. Il semble au contraire que de nombreux auteurs ont apporté leur propre mesure de la concurrence sans se préoccuper d'intégrer les concepts sous-jacents à leur mesure au sein d'un corps théorique solide et construit.

Les auteurs préfèrent en général préciser le type de données qu'ils utilisent en se replaçant dans l'une des positions de la dichotomie qui oppose *"données de comportement ou données de "jugement"*. Cette identification entre un concept et sa mesure est le point de vue général de l'école dite "d'Opérationnalisme Strict", pour qui il y a identité entre le concept et sa mesure (voir Bagozzi 1980 p 122 pour une discussion générale à ce sujet). De cette confusion entre le concept de concurrence et sa mesure découle une multitude d'approches plus ou moins différentes qui se sont développées sans vraiment s'intégrer au sein d'un corps théorique structuré, comme c'est le cas pour le concept de *"Satisfaction du Consommateur"*.

Pour mesurer la concurrence entre les marques, le praticien se trouve confronté à une multitude d'approches différentes, sans disposer réellement de critères (théoriques ou pratiques) lui permettant de faire un choix parmi les approches possibles et encore moins pour comprendre et justifier son choix.

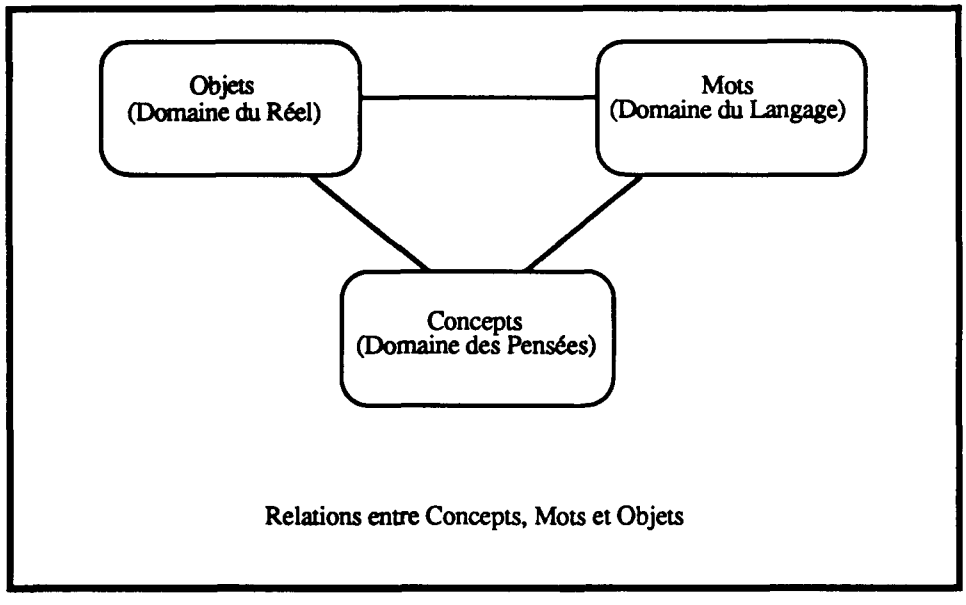
SECTION1 MESURE ET VALIDITE : ELEMENTS FONDAMENTAUX

SS1 LE PROBLEME DE LA MESURE DES CONCEPTS

Puisque qu'il est important de considérer la concurrence en tant que *"concept théorique"* avant de vouloir en effectuer la mesure il est important de commencer par décrire ce que signifie le terme de "concept".

Bagozzi (1980 p 114) définit un concept comme étant *"une unité de pensée"*. Un concept, toujours selon Bagozzi, prend sa signification à partir de sa relation avec le monde des mots et celui des objets (voir la figure 3 suivante).

FIGURE 3 LIEN ENTRE CONCEPTS MOTS ET OBJETS



"Adapté de Bagozzi, 1980 p 115"

Reprenant Carnap (1966), Bagozzi (1980 p 116) distingue trois types de concepts :

- concept de type *"classificatif"* comme par exemple le concept de "classification des produits" en Marketing (voir Murphy & Enis 1986).
- concept de type *"comparatif"*, comme le concept de *"préférence"*
- concept de type *"quantitatif"*, comme le concept de *"Satisfaction"*

Nous insistons sur le parallèle qu'il est possible de faire entre un tel découpage et les types de variables rencontrées en Sciences Sociales, à savoir: variables qualitatives, ordinales ou quantitatives (d'intervalle ou de ratio) .

Le concept de concurrence est, de par sa nature, essentiellement de type quantitatif. De la même manière que l'attitude est un concept théorique non observable dont le chercheur n'étudie que les reflets sur des échelles de mesure de l'attitude, la concurrence entre deux marques est un concept théorique non observable qu'il s'agit de mesurer. En Science Sociale, une grande partie des concepts étudiés se situent à un niveau théorique et non observable. Par exemple, depuis longtemps les psychologues sont conscients que le

concept d'intelligence qu'ils essaient de mesurer à l'aide de tests n'existe que par les définitions qu'ils en donnent et n'est donc pas directement observable. Ceci explique peut-être pourquoi les psychomètres ont, dès les années trente, étudié les moyens de s'assurer que les mesures qu'ils effectuent sont correctes.

Ce n'est que beaucoup plus tard (à la fin des années soixante dix) que les chercheurs en Marketing ont réalisé qu'ils se trouvent dans la même situation, lorsqu'ils étudient l'influence de la publicité sur les préférences, que les psychomètres lorsqu'ils étudient le lien entre intelligence et succès scolaire.

La difficulté qu'il y a à s'assurer que les mesures pratiquées sont correctes est certainement à l'origine de l'importance des travaux consacrés à la validité des mesures. L'utilisation qui en a été faite en Marketing n'est pourtant pas en rapport avec un tel effort. Tout se passe souvent comme si les concepts *s'imposaient d'eux mêmes* ainsi que leur mesure. Peter et Churchill ont mené deux recherches (voir Churchill et Peter 1984 et Peter et Churchill 1986) qui caractérisent et quantifient cette faiblesse. Dans notre domaine de recherche en particulier, quasiment aucune démarche n'a été faite (à notre connaissance) dans ce sens là.

Nous pouvons cependant citer Lehmann (1972) qui a comparé des cartes représentant des marchés obtenus selon l'approche traditionnelle des similarités perceptuelles à des cartes obtenues à partir de données de panels (transferts entre marques).

Plus récemment, Elrod et Winer (1988) ont effectué une comparaison entre cinq méthodes d'analyse de la structure des marchés. La comparaison a été faite sur la base de trois critères :

- prévision des ventes (c'est à dire la validité prédictive)
- validité faciale des cartes du marché obtenues à à partir de ces méthodes
- facilité (et coût) de mise en oeuvre de ces méthodes

Sans rentrer ici dans le détail, nous pouvons observer les points suivants:

- aucune de ces approches ne s'est inspirée des techniques et méthodes qui ont été développées dans le domaine de l'étude de la validité des mesures.
- ces deux recherches n'avaient pas pour objet d'étudier la qualité de la mesure. Notamment, dans leur étude comparative, Elrod et Winer n'isolent pas le problème de la mesure mais considèrent les méthodes comparées de manière globale. Leurs résultats portent ainsi sur l'ensemble: "mesure et méthode statistique de traitement de la mesure", qui correspond aux phases b et c de notre méthodologie d'Analyse de la Structure des Marchés.

Cette méthodologie ne permet pas de savoir si la valeur d'une approche est liée à la qualité de sa mesure où à celle du modèle analysant ces mesures (et vice-versa).

D'autre part, la comparaison que font les auteurs repose sur le critère de la validité prédictive (prédiction des parts de marché des marques) alors que pour certaines des méthodes (en particulier l'analyse factorielle des correspondances), l'objectif n'est pas de prévoir mais de représenter un ensemble d'indices de concurrence entre les marques.

SS2 THEORIE DE LA MESURE EN SCIENCE SOCIALE

Les théories de la mesure (et les modèles de la mesure correspondant) se sont construites autour des notions de "Fiabilité de la mesure" et, de manière plus récente et plus générale, autour de la notion de "Validité des mesures".

A partir du moment où ce qui est mesuré n'est pas directement observable, la difficulté ne réside pas tant dans la mesure mais dans l'évaluation de la qualité de cette mesure ! L'appellation "validité d'une mesure" recouvre un ensemble varié voire hétérogène d'aspects de la mesure, à savoir :

- fiabilité, consistance interne, stabilité
- validité concurrente
- validité de trait
- validité convergente et discriminante
- validité nomologique
- validité faciale
- validité prédictive
- etc ...

Ces différents types de validité peuvent être regroupés en deux grandes classes assez distinctes dans leur objet :

1 - Ce que nous appelons "Validité Interne" est une évaluation de la qualité de la mesure selon des critères internes à cette mesure. La validité interne permet d'assurer que ce qui est mesuré est bien mesuré et qu'on ne mesure que cela. Les critères les plus courants de validité sont la fiabilité, la validité convergente et la validité discriminante.

2 - Ce que nous appelons "Validité Externe" (à ne pas confondre avec la validité externe dans une expérimentation) a pour but de s'assurer que ce qui est mesuré correspond bien à ce qui devait être mesuré. Les critères les plus courants sont, dans cette optique là, la validité nomologique, la validité prédictive et la validité faciale.

Nous pouvons faire le parallèle entre ces deux types de validité et les problèmes posés par la mise en oeuvre d'une procédure d'échantillonnage. La validité interne correspond, dans ce cas, au fait de sélectionner un échantillon nous permettant d'effectuer une estimation précise de la variable étudiée. De nombreuses techniques, très sophistiquées, permettent de s'en assurer. Quand la validité interne d'un échantillon est bonne, cela signifie que cet échantillon est bien représentatif de la population dont il est extrait.

La validité externe pose un problème délicat qui est de savoir si la population, telle qu'elle a été définie (et si la validité interne est bonne, dont l'échantillon est représentatif) est pertinente eu égard au problème posé. En un mot, est-ce que la population, dont l'échantillon est représentatif, est bien celle de laquelle il fallait extraire l'échantillon ?
En fait, dans le cas de la mesure d'un concept comme dans celui de la représentation d'une population, le point le plus délicat est souvent le second, c'est à dire celui qui est lié à la validité externe. Le chercheur n'a jamais la certitude (faute de technique) que ce qu'il a mesuré (ou représenté) correspond bien à ce qu'il souhaitait mesurer (représenter) à l'origine.

SS3 FIABILITE DES MESURES

Guilford (1954 p 350) définit la fiabilité comme *"la proportion de la variance vraie du phénomène étudié qui est restituée par la mesure"*. Nunnally (1978 p 191) définit la fiabilité comme étant *"la propension avec laquelle une mesure est répétable"*.

La fiabilité d'une mesure est donc sa propension a fournir les mêmes valeurs dans les mêmes conditions de mesure.

SS3-1 LES DIFFERENTES APPROCHES DE LA FIABILITE

SS3-1-1 APPROCHE CLASSIQUE :

Les travaux de Spearman, au début du siècle, ont ouvert la voie des recherches sur la validité des mesures effectuées en Science Sociale. Le modèle classique de mesure s'exprime à partir de deux niveaux de relations :

Relation sur les scores :

$$X_O = X_t + X_e \quad (1)$$

Où :

X_O est la valeur observée de la mesure (du score)
X_t est la valeur vraie (non observable) de la mesure
X_e est l'erreur aléatoire de la mesure

Relation au niveau des variances des scores :

$$V_O = V_t + V_e \quad (2)$$

Où:

V_O est la variance observée de la mesure
V_t est la variance vraie qui inclut toute source de variance systématique
V_e est la variance due à l'erreur aléatoire, donc *"non systématique"*

Comme, par hypothèse, les termes aléatoires ont une espérance nulle (les erreurs se compensent), les échelles multiples, toutes choses égales par ailleurs, sont plus fiables. En effet, selon le modèle classique, plus les mesures du même concept sont nombreuses, plus la somme de leurs scores sera exempte d'erreur aléatoire .

A partir de ce modèle, la fiabilité, F, se définit comme étant la partie de la variance dans la mesure qui n'est pas due à une cause d'erreur aléatoire .

$$F = \frac{V_t}{V_O} = 1 - \frac{V_e}{V_O} = \frac{V_O - V_e}{V_O} \quad (3)$$

Plus l'erreur de type aléatoire est importante et moins la fiabilité est bonne .

SS3-1-2 LA THEORIE DU DOMAINE D'ECHANTILLONNAGE

Cette théorie présente un instrument de mesure (multiple) comme étant un échantillon aléatoire de tous les items possibles représentant le domaine du concept hypothétique étudié. Par exemple, pour étudier un trait particulier de la personnalité d'un individu, il est possible d'utiliser une infinité d'items différents. Tous ces items représentent le domaine du concept étudié. Le score vrai est alors défini comme étant le score qui serait obtenu en sommant l'ensemble des scores obtenus avec l'infinité des items appartenant au domaine. On parle ainsi de *"score du domaine"*.

$$X_t = E(X_i) \quad i = 1 \text{ à } -\infty \quad (4)$$

$E(X_i)$ est l'espérance des valeurs du score X_i .

Le coefficient de fiabilité d'un item peut alors être exprimé comme étant la corrélation moyenne de cet item avec le score vrai c'est à dire la totalité des items du domaine (voir Nunnally 1978 p197).

$$r_1(1,- - - k) = r_{1t} = (\overline{r_{1j}})^{1/2} \qquad (5)$$

$\overline{r_{1j}}$ est la corrélation moyenne de l'item j avec tous les autres

Cette relation s'exprime de la manière suivante : *"la corrélation de l'item 1 avec le score vrai du domaine est égale à la racine carrée de la corrélation moyenne de cet item avec tous les autres items du domaine"*. Comme en pratique il est impossible d'étudier tous les items, r_{1j} est estimé par re_{1j} qui est la corrélation moyenne de l'item 1 avec un échantillon (au sens de la théorie de l'échantillonnage) des items représentant le domaine.

Cette théorie, très séduisante d'un point de vue conceptuel, a en plus l'avantage d'être compatible avec de nombreux outils statistiques que nous utiliserons par la suite (alpha de Cronbach, analyse factorielle ..).

SS3-1-3 LE MODELE DES TESTS PARALLELES

C'est une théorie alternative à celle du domaine mais qui aboutit au même estimateur de la fiabilité que (5). Deux tests sont dits parallèles si :

- leurs variances sont égales
- leurs corrélations avec X_t, le score vrai, sont identiques
- la variance qui n'est pas due à X_t est purement aléatoire

L'expression de la corrélation entre deux tests parallèles 1 et 2 (notée r12) permet de montrer qu'on obtient exactement les mêmes résultats que ceux de la théorie du domaine, à savoir (Nunnally 1978 p 202):

$$r_{1t} = (r_{12})^{1/2} \quad (6)$$

r_{1t} est la corrélation entre le score 1 et le score vrai

SS3-1-4 LA THEORIE DES FACTEURS

Les deux théories précédentes supposent que tous les items d'une même échelle appartiennent à un domaine *"unidimensionnel"*. Elles restent encore vraies lorsque les items peuvent appartenir à différentes dimensions d'un même domaine. C'est le cas si nous étudions un concept théorique multidimensionnel. L'apport majeur de cette théorie est de considérer (selon Nunnally 1978 p 205) :

- que les items peuvent mesurer plusieurs facteurs et non un seul
- que l'intention du chercheur doit être d'étudier un domaine d'items qui ne concerne principalement qu'un facteur. En effet, même si la théorie du domaine vaut toujours lorsque ces items sont partagés en plusieurs facteurs, l'estimation de la corrélation avec le score vrai est plus précise lorsque les items ne sont dominés que par un seul facteur. La théorie des facteurs (voir une description détaillée dans Guilford 1954 p 355) est directement calquée sur la théorie de l'analyse factorielle qui considère que la variance observée peut être décomposée en :
  - variance imputable aux facteurs communs, Vfc
  - variance spécifique à l'item, Vfs
  - variance due de l'erreur aléatoire Ve

$$V_o = V_{fc} + V_{fs} + V_e \quad (7)$$

De la même manière:

$$X_o = X_{fc} + X_{fs} + X_e \quad (8)$$

Où:

- X_o est le score observé
- X_{fc} est le score imputable aux facteurs communs
- X_{fs} est le score spécifique à l'item
- X_e est le score dû à l'erreur aléatoire

SS3-2 MESURE DE LA FIABILITE

Les mesures de la fiabilité qui ont été proposées dans la littérature consacrée à la mesure sont nombreuses (voir Guilford 1954 pp 373 à 399 pour une revue complète, Parameswaran et al 1979, Peter 1979, Nunnally 1978 chap 6). Cependant, trois grandes approches sont en général retenues.

SS3-2-1 MESURE DE LA CONSISTANCE INTERNE (Split-Half) :

Cette technique découle directement de l'application de la théorie du domaine. Elle est fondée sur la constatation que la source la plus importante de l'erreur de mesure provient de l'erreur d'échantillonnage des items de mesure (Nunnally 1978 p 224).

Sous l'appellation "Split Half", sont regroupées un ensemble de techniques consistant à partager l'échantillon des items en deux (ou plus) parties afin de calculer la corrélation entre les scores des deux sous-ensembles d'items ainsi constitués, pour en estimer le degré de cohérence interne. Les méthodes permettant de partager l'ensemble des items en deux sont multiples, la plus simple consiste à effectuer le partage sur la base du numéro de l'item, pair ou impair (voir Guilford 1954 chap 14 pour une revue de ces méthodes).

La formule la plus connue et la plus satisfaisante au plan théorique est celle qu'a proposé Cronbach avec son "coefficient alpha" (voir Cronbach 1951). Ce coefficient alpha représente la corrélation moyenne entre tous les partages possibles, en deux parties, de l'échantillon des items du domaine. Du point de vue de la théorie du domaine d'échantillonnage, le coefficient alpha représente une limite haute de la fiabilité (la cohérence interne est au mieux égale à la valeur du score).

SS3-2-2 METHODE DES FORMES ALTERNATIVES :

Issue de la théorie des tests parallèles, cette approche permet d'estimer la fiabilité lorsque le chercheur ne se trouve plus en présence d'un échantillon d'items sélectionné dans l'univers du domaine, mais qu'il est en présence de deux formes de tests construites de manière indépendante (séparée). La méthode consiste alors à administrer ces deux formes alternatives de test en deux périodes de temps (en général espacées de deux semaines). Par rapport au test de consistance interne, cette approche a pour objectif de cerner parmi les trois sources d'erreur aléatoire suivantes, la ou les plus importantes (voir Nunnally 1978 p 231) :

- erreur due à des différences systématiques dans le contenu des deux tests
- erreur due à la subjectivité des répondants
- erreur due à la variabilité des réponses dans le temps

SS3-2-3 TEST-RETEST

Ce troisième type d'approche de la fiabilité est en fait un cas particulier du test des formes alternatives puisque le "même" test est présenté aux mêmes personnes à deux périodes de temps différentes.

Facile et économique a mettre en oeuvre, le retest à été souvent utilisé, malgré les critiques nombreuses que lui portent les psychométriciens .

L'élément essentiel qui lui est reproché est d'être largement influencé par l'effet de mémoire des répondants. De ce fait, les mêmes erreurs de réponse seraient commises sur les mêmes items aux deux époques de la mesure. Ce coefficient serait ainsi plus un indice de mémoire qu'un indice de la fiabilité de la mesure. D'autre part, ce test s'intègre mal dans la théorie du domaine d'échantillonnage. Comme le note Nunnally, (1978 p 234), l'intérêt du test-retest réside surtout dans son interprétation de manière négative : si le

niveau du retest est trop faible, celui de la forme alternative le serait encore plus. Le test-retest est alors considéré comme une condition nécessaire mais non suffisante de fiabilité. Son usage est par contre recommandé lorsque le chercheur n'anticipe pas d'effet de mémoire de la part du répondant. Le retest peut être utilisé selon une autre optique qui est de vérifier la stabilité des scores sur le long-terme (environ six mois) afin de prendre en compte d'éventuelles variations dans l'objet mesuré. Cette utilisation du test-retest est intéressante étudier et mesurer des caractéristiques qui a priori doivent comporter un degré de permanence.

SS3-3 GENERALISATION DU CONCEPT DE FIABILITE

Les différentes approches de la mesure de la fiabilité ont des bases théoriques qui peuvent être différentes et conduire à l'obtention de niveaux de fiabilité très différents d'une formule à l'autre. Parameswaran et al (1979) ont caractérisé ce problème en comparant les principales approches de la fiabilité. Arguant du fait que le niveau de la fiabilité dépend de ce que signifie le terme *"erreur de mesure"* puisque celle-ci peut avoir différentes origines, Cronbach a développé (Cronbach, Glezer et al 1972) une théorie appelée *"Théorie de la Généralisation"* ("generalizability theory"). Cette théorie repose sur la notion de *"conditions spécifiques d'échantillonnage"*. L'époque, l'instrument utilisé, l'observateur, de la mesure, sont autant de *"facettes"* différentes de cette mesure. Cette théorie est donc la généralisation de la Théorie du Domaine lorsque le domaine ne se limite pas à l'infinité des instruments de mesure possibles mais est généralisé à l'ensemble des époques, des observateurs, des situations possibles de la mesure. Le problème de la mesure consiste alors à choisir, parmi l'infinité des mesures possibles du concept étudié, un échantillon représentatif de ces mesures. Cette représentativité peut être formalisée par la corrélation entre le score de l'instrument utilisé et le score de l'univers (non observable) dont est représentatif l'instrument de mesure utilisé.

Avec un test-retest, en deux époques différentes avec les mêmes individus et le même instrument, seule la représentativité de la mesure selon la dimension "période d'observation" est contrôlée. Le chercheur ne sait pas, par contre, si la mesure est fiable selon les autres dimensions possibles de la mesure. De la même manière, la mesure de la fiabilité par le coefficient alpha revient à étudier la fiabilité de la mesure selon la seule facette "instrument".

La théorie de la généralisation montre que la fiabilité doit être envisagée selon toutes les composantes possibles de la variance d'une mesure. Comme le note Peter (1979), cette théorie a pour principal avantage de faire prendre conscience au chercheur que les coefficients de fiabilité qu'il calcule ne sont qu'une des nombreuses facettes de la fiabilité.

SS4 VALIDITE DES MESURES

Dans le modèle de mesure présenté par G Churchill (1979), le score observé d'une mesure s'exprime sous la forme :

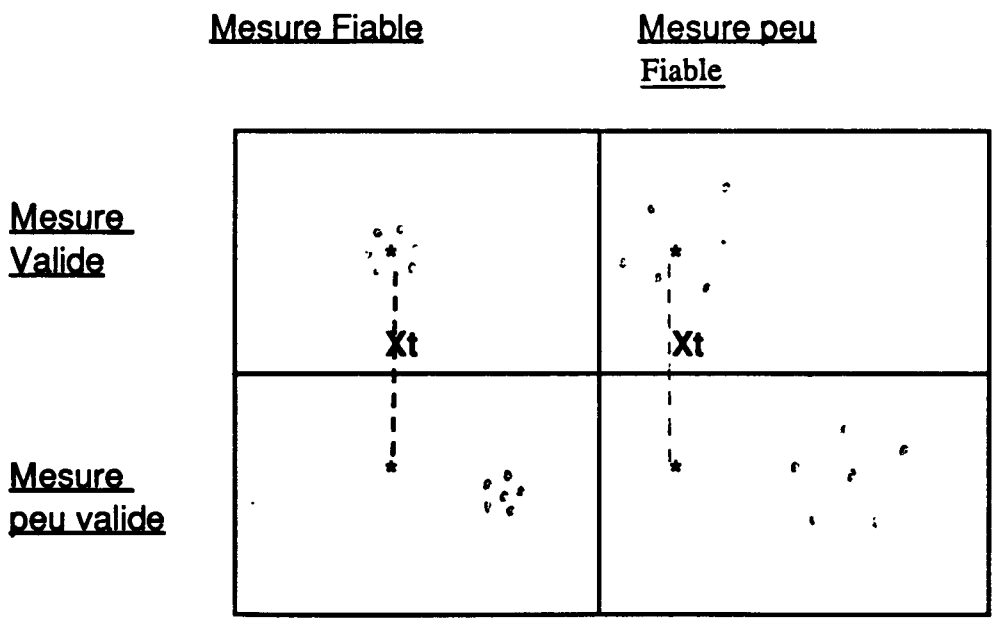
$$X_O = X_t + X_S + X_e$$

X_t : score vrai
X_S : score dû à une source d'erreur systématique, stable (par exemple une caractéristique permanente du répondant)
X_e : score dû à une erreur d'origine aléatoire

Ce modèle permet de faire la différence entre fiabilité et validité d'une mesure. En effet, de la faiblesse de X_e, va dépendre le niveau de fiabilité de la mesure. Or, nous avons vu que la meilleure manière de minimiser cette erreur est d'utiliser une échelle multiple permettant d'avoir un échantillon mieux représentatif des items du domaine étudié (ou de compenser les erreurs aléatoires selon le modèle classique).

X_S correspond par contre à une erreur de type systématique, qui est commise à chaque mesure. Cette erreur n'a rien d'aléatoire. Ainsi, même si X_e tend vers 0 (fiabilité maximale), la mesure n'en est pas pour autant valide. Comme le note Evrard (1985), la fiabilité est une *"condition nécessaire mais non suffisante de validité"*. La multiplicité des items ne garantit pas qu'une mesure soit valide. La figure suivante (extrait de cours, Evrard) permet de visualiser le problème.

FIGURE 4 REPRESENTATION SCHEMATIQUE DE LA VALIDITE ET DE LA FIABILITE



X_t: score vrai

SS4-1 ETUDE DE LA VALIDITE INTERNE

Guilford (1954) définit la validité d'un score par sa *"capacité à prévoir tout ce avec quoi il est corrélé, mais pas lui même, ce qui serait alors de la fiabilité"*. A propos de la théorie des facteurs, le même auteur (p 399) définit la validité comme étant *"la part dans la variance du score observée qui est expliquée par les facteurs communs"* (ce qui correspond, dans l'analyse factorielle, à la valeur des communalités).

Pour Campbell et Fiske (1959), il s'agit de la validité convergente et de la validité discriminante. Selon Churchill (1979), une mesure est valide quand *"des différences entre les scores observés reflètent de vraies différences dans les caractéristiques du concept mesuré"*. Peter (1981) parle de *"correspondance verticale entre un concept (non observable) et sa mesure "* ou bien *"du degré avec lequel la mesure correspond au concept qu'elle est supposée mesurer"*. Zaltman (1973 p 44) définit la validité comme étant "le degré avec lequel une opérationnalisation mesure le concept qu'elle est censée mesurer".

Constatant que les définitions données à la notion de validité ne recouvrent pas toute son étendue, Bagozzi (1980 p 117), représente la validité d'un construit à partir des six composantes suivantes :

1 Valeur Théorique du concept
2 Capacité à être Observable
3 Consistance Interne
4 Validité Convergente
5 Validité Discriminante
6 Validité Nomologique

Le premier point a trait à la valeur théorique d'un concept. Celle-ci dépend de la nature et de la consistance interne des mots qui sont utilisés pour le représenter. Cette valeur est essentiellement de type sémantique. Nous avons dit, dans l'introduction générale, que les concepts naissent au niveau théorique. Un concept doit donc être défini de manière claire et pertinente pour prétendre à être mesurable.

L'étude des attitudes est un exemple de domaine où les chercheurs ne sont pas vraiment arrivé à formuler une définition claire et unique du concept, nous avons vu dans la première partie qu'il en est de même pour le concept de concurrence.

Le deuxième point est lié à la capacité qu'a le concept à être observable. Un concept tire sa validité du degré avec lequel il est possible de le relier à d'autres concepts, qui eux sont observables.

Le troisième point correspond à la définition classique de la fiabilité (consistance interne, test-retest).

Les points quatre et cinq correspondent aux notions de validités convergente et discriminante. Ces deux points sont fondamentaux du point de vue de l'étude de la validité interne d'une mesure, ils constituent ce qui est appelé la *"validité de traits"*. Nous allons les décrire en détail.

SS4-2 VALIDITE CONVERGENTE ET VALIDITE DISCRIMINANTE

Ces critères de validité ont été formalisés de façon très simple par Campbell et Fiske (1959). La logique de leur démarche consiste à apprécier le degré avec lequel la mesure capte le concept et ne capte que lui.

Une fois définis, d'un point de vue théorique, les contours du concept à mesurer (par exemple l'attitude, la satisfaction, la concurrence entre deux marques), il s'agit d'étudier si les mesures qui sont faites de ce concept convergent bien et ne sont pas contaminées par d'autres concepts que le chercheur ne souhaitait pas mesurer.

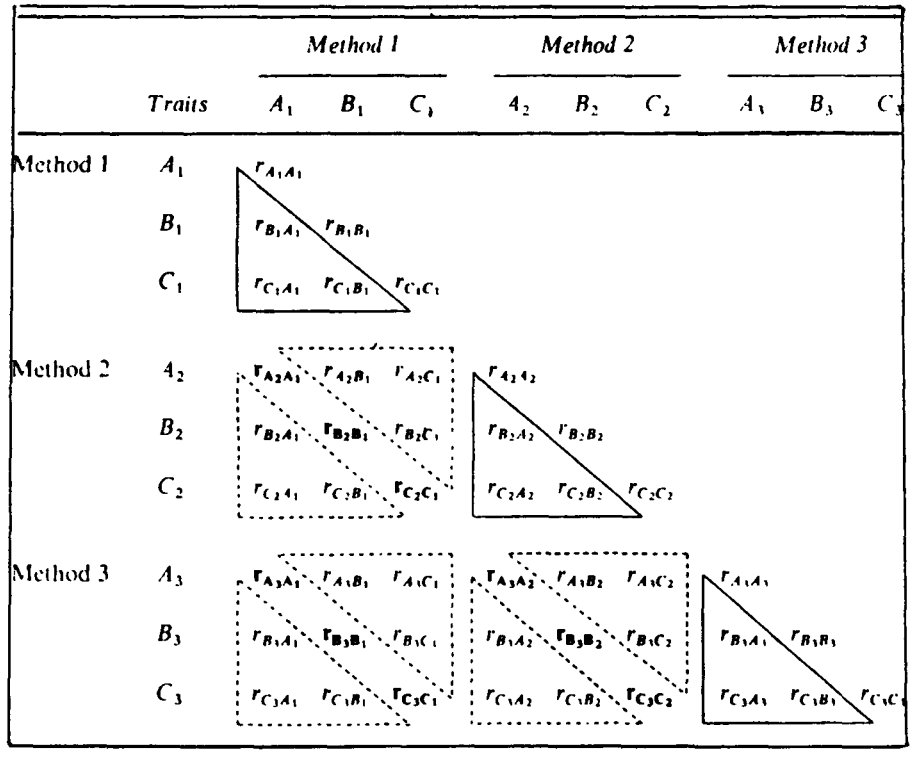
La validité de traits comprend deux dimensions :

- la validité convergente
- la validité discriminante

Campbell et Fiske ont proposé un outil permettant d'apprécier ces deux types de validité. Cet outil se présente sous la forme d'une matrice de corrélations entre différents traits (concepts théoriques mesurés) et différents instruments de mesure. Cette matrice est appelée Matrice Multi-Traits Multi-Méthodes (MTMM). Il est souhaitable que les traits ne soient pas trop différents afin de fournir un bon test de discrimination de même qu'il faut utiliser des mesures aussi indépendantes que possibles, afin de ne pas sur-estimer le niveau de la convergence. L'exemple suivant (voir Figure 5), adapté de Campbell et Fiske par Bagozzi (1980 p 130) présente une matrice MTMM ainsi que les différents éléments qui la composent.

FIGURE 5 EXEMPLE DE MATRICE MTMM

A Multitrait-Multimethod Matrix for Three Traits and Three Methods



"Extrait de Bagozzi 1980 p 130"

Trois traits A,B et C (les trois composantes cognitive, affective et conative de l'attitude) sont mesurés à l'aide de trois méthodes (instruments) de mesure des attitudes (échelles de Thurstone, de Likert et de Guttman).

Dans la diagonale principale, figurent les valeurs des coefficients de fiabilité de chacune des mesures (il y a neuf mesures).

En caractères gras, figurent les diagonales de validité, qui sont constituées par les corrélations entre des mesures du même trait effectuées à l'aide de méthodes différentes.

Dans les triangles en trait plein, appelés triangles hétéro-traits mono-méthode, figurent les corrélations entre différents traits mesurés par la même méthode. Puisqu'il s'agit de traits différents, ces corrélations représentent la variance imputable à l'instrument de mesure.

Enfin, dans les triangles en traits pointillés, appelés triangles hétéro-traits hétéro-méthodes, figurent les corrélations entre différents traits mesurés à l'aide de différentes méthodes.

Campbell et Fiske (C&F) formalisent les critères de validité convergente et discriminante dans les termes suivants (traduit de Campbell et Fiske 1959 p 82) : *"A partir du diagramme précédent, la validité repose sur la vérification de quatre contraintes. Les valeurs dans la diagonale de fiabilité doivent être significativement différentes de zéro et suffisantes pour encourager des analyses ultérieures de la validité. Cette contrainte est une preuve de la validité convergente. Deuxièmement, les valeurs de la diagonale de validité doivent être plus grandes que les valeurs correspondantes dans les colonnes et les lignes hétéro-traits et hétéro-méthodes. Ceci signifie qu'une valeur de validité pour une variable doit être plus importante que les corrélations de cette variable avec n'importe quelle autre n'ayant ni de trait ni de méthode en commun avec elle ..Troisièmement, une variable doit être mieux liée avec un effort indépendant pour mesurer le même trait qu'avec des efforts pour mesurer d'autres traits mais avec la même méthode. Pour une variable, il faut ainsi comparer ses valeurs dans la diagonale de validité avec ses valeurs dans les triangles hétéro-traits mono-méthode. Une quatrième condition est que le même pattern de relations entre les traits se retrouve dans tous les triangles hétéro-traits mono ou hétéro-méthodes ..."*

La matrice MTMM constitue, avec la fiabilité, un des outils essentiels pour l'étude de la validité des mesures. Elle a été cependant critiquée par plusieurs auteurs qui lui reprochent en général les points suivants (voir Schmitt, Coyle et al 1977, Jackson 1969):

- les critères de C&F n'offrent pas une mesure quantitative, objective, du degré avec lequel la matrice remplit les critères de convergence et de discrimination.
- puisque la mesure n'est pas complètement fiable, les valeurs des corrélations dans la matrice sont sujettes à fluctuations.
- l'analyse de la matrice ne permet pas d'appréhender globalement l'ensemble des corrélations, puisqu'il faut se cantonner à effectuer des analyses par couples de corrélations. Ceci est très réducteur, surtout si le concept mesuré inclut plus d'un trait (concept multi-facettes).

Nous pouvons constater que l'essentiel de la critique porte sur l'aspect informel des critères proposés par C&F. De nombreuses alternatives ont été proposées pour analyser la matrice MTMM, nous les séparerons en deux grandes approches:

-a- Une approche empirique, où sont utilisées des méthodes permettant simplement d'étudier la structure des corrélations contenues dans la matrice MTMM. Plusieurs méthodes ont été proposée: l'analyse de la variance (Stanley 1961, cité dans Schmitt et al 1977), diverses formes plus ou moins compliquées d'analyses factorielles ou en composantes principales (voir Schmitt et al 1977 pour une comparaison de ces approches). D'autre part, Heeler et Ray (1972) ont proposé l'utilisation jointe de l'analyse typologique hiérarchique et de l'analyse des similarités.

-b- Une approche confirmatoire où est testée l'adéquation de la matrice des corrélations observées avec un pattern de corrélations hypothétique, fixé à l'avance en fonction des caractéristiques que devrait présenter la mesure pour être valide. Dans cette optique là, Joreskog (1969) a proposé d'utiliser le *"path analysis"*, *"l'analyse factorielle confirmatoire"* avec estimation par le maximum de vraisemblance, cette méthode permettant d'effectuer des tests d'hypothèses sur la structure sous-jacente aux données analysées. Plus récemment, Joreskog (voir Joreskog dans Fornell 1982 p 200), Bagozzi (1980) ont proposé l'utilisation des modèles d'équations simultanées avec variables latentes et erreur sur la mesure (voir Bagozzi 1977, Fornell et Larcker 1981), et en particulier le modèle Lisrel (voir Joreskog et Sorbom 1981, 1982) pour étudier la matrice MTMM. Toutes ces méthodes permettent d'apporter leur puissance et leur finesse d'analyse mais ne changent en rien les principes de base de la méthode qui réside dans la matrice MTMM originelle.

Nous conclurons sur la validité interne en notant l'opposition qui existe entre fiabilité et validité convergente. La théorie du domaine nous a montré que la fiabilité d'une mesure peut être exprimée comme étant la corrélation d'une mesure avec le score vrai. Si, à l'extrême, le domaine du construit était réduit à un seul item, celui qui est utilisé pour effectuer la mesure, alors la fiabilité serait maximale. Ce faisant, le champ du domaine couvert par le concept serait considérablement réduit, la validité serait donc mauvaise. Guilford (1954) note que l'utilisation d'une décomposition factorielle du concept permet d'améliorer la validité puisqu'elle conduit à mieux couvrir le domaine.

Tous les points proposés par Bagozzi et qui viennent d'être passés en revue concernent la qualité de la mesure étudiée uniquement sous son angle interne. C'est ce que nous appelerons les *"critères internes au laboratoire de mesure"*.

Une fois que la mesure a satisfait à ces critères internes de validité, il est nécessaire de la sortir de son contexte de laboratoire pour la confronter à la réalité des théories et des faits. Notamment, le sixième critère de validité proposé par Bagozzi concerne la validité Nomologique, c'est à dire la capacité du concept étudié à s'intégrer dans un réseau de relations déjà existantes. C'est un critère que nous qualifierons *"d'externe au laboratoire de mesure"*

SS5 VALIDITE EXTERNE

Lorsque la validité interne d'une mesure est établie, le chercheur s'est en fait garanti:

- que la mesure est fiable selon la ou les dimensions à partir de laquelle cette fiabilité a été établie
- que les différentes mesures captent bien le même concept et n'en captent pas en même temps d'autres, en particulier ceux sur la base desquels est établie la validité discriminante.

Ce dont le chercheur est par contre moins sûr est que *"ce qui est mesuré correspond bien à ce qu'il souhaitait mesurer à l'origine"*. Il s'agit là d'un problème de validité externe.

Il n'existe pas d'approche théorique aussi formalisée pour étudier la validité externe que pour étudier la validité interne. Une telle démarche reste sujette aux incertitudes de toute démarche déductive confrontée au falsificationnisme tel que Popper (1963) l'a défini, à savoir que *"Une théorie est tenue pour vraie (ici une mesure valide d'un point de vue externe) tant que le contraire n'a pas été prouvé"*. Par contre, il est possible par approches successives, d'augmenter ses chances d'avoir obtenu une mesure présentant une bonne validité externe. Trois grandes approches sont possibles .

SS5-1 VALIDITE NOMOLOGIQUE

L'opération consiste à apprécier la validité de la mesure en la sortant de son contexte interne, afin de la confronter à un corps théorique déjà présent. Est ce que la mesure est compatible avec la théorie admise par la communauté scientifique ? Pour répondre à cette question, la mesure est confrontée à un réseau de relations qui ont été déjà établies et le chercheur va évaluer avec quel degré la mesure est compatible avec ce réseau. Les méthodes basées sur l'analyse des structures de covariance (Path-Analysis, Lisrel) sont particulièrement bien adaptées à ce type de recherche. Nous trouverons des exemples d'application de ces méthodes dans Bagozzi (1980 chapitre V), Bagozzi dans Fornell (1982) pp 145 à 183 pour une application à la mesure des attitudes, Churchill et Pecotich (1982) pour une application à la mesure de la satisfaction du consommateur.

SS5-2 VALIDITE PREDICTIVE

Bagozzi (dans Fornell 1982 pp 145 à 183) définit comme étant valide du point de vue prédictif, *"une mesure qui est reliée empiriquement, en tant qu'antécédant ou en tant que conséquence, à la mesure d'une autre variable"*. L'objectif est d'essayer d'effectuer des prévisions de la mesure à valider à partir d'autres mesures ou inversement de prévoir d'autres mesures à partir de la mesure à valider.

SS5-3 VALIDITE FACIALE

Il s'agit là de vérifier si la mesure est compatible avec le bon sens, ou avec l'avis d'experts utilisateurs de ces mesures. Ce type de validité est très important, comme le notent par exemple Shocker et al (1987).

Tout ce que nous venons d'exposer d'un point de vue théorique et qui à été développé dans le contexte de la mesure des phénomènes psychologiques à l'aide de batteries de tests, a été appliqué en Marketing pour l'étude du comportement du consommateur. Par contre, comme nous l'avons constaté dans l'introduction de ce chapitre, aucune étude (à notre connaissance) n'a tenté d'étudier la validité de la mesure de la concurrence entre les marques sous l'angle de la théorie de la mesure.

Pourtant, toutes les conditions pour une telle approche nous paraissent réunies:

- nous avons souligné, dans la partie théorique de cette recherche, que toute analyse de la structure d'un marché passe (de manière explicite ou implicite) par une phase de mesure: la mesure de la concurrence entre les marques.
- le concept mesuré est par essence "théorique", mais la revue de littérature que nous avons effectuée ne caractérise pas un consensus à ce sujet. Nous avons vu qu'à ce niveau deux grands courants s'affrontent:

  le courant de l'approche "jugement"
  le courant de l'approche "comportement"

Au sein même de la deuxième approche sur laquelle nous nous sommes centrés dans cette recherche, le concept de concurrence n'est pas clairement défini. Mieux il ne semble pas nécessaire, pour un grand nombre d'auteurs, de présenter une définition de ce concept. De plus, il est opérationnalisé par des mesures dont les fondements théoriques sont forts différents :

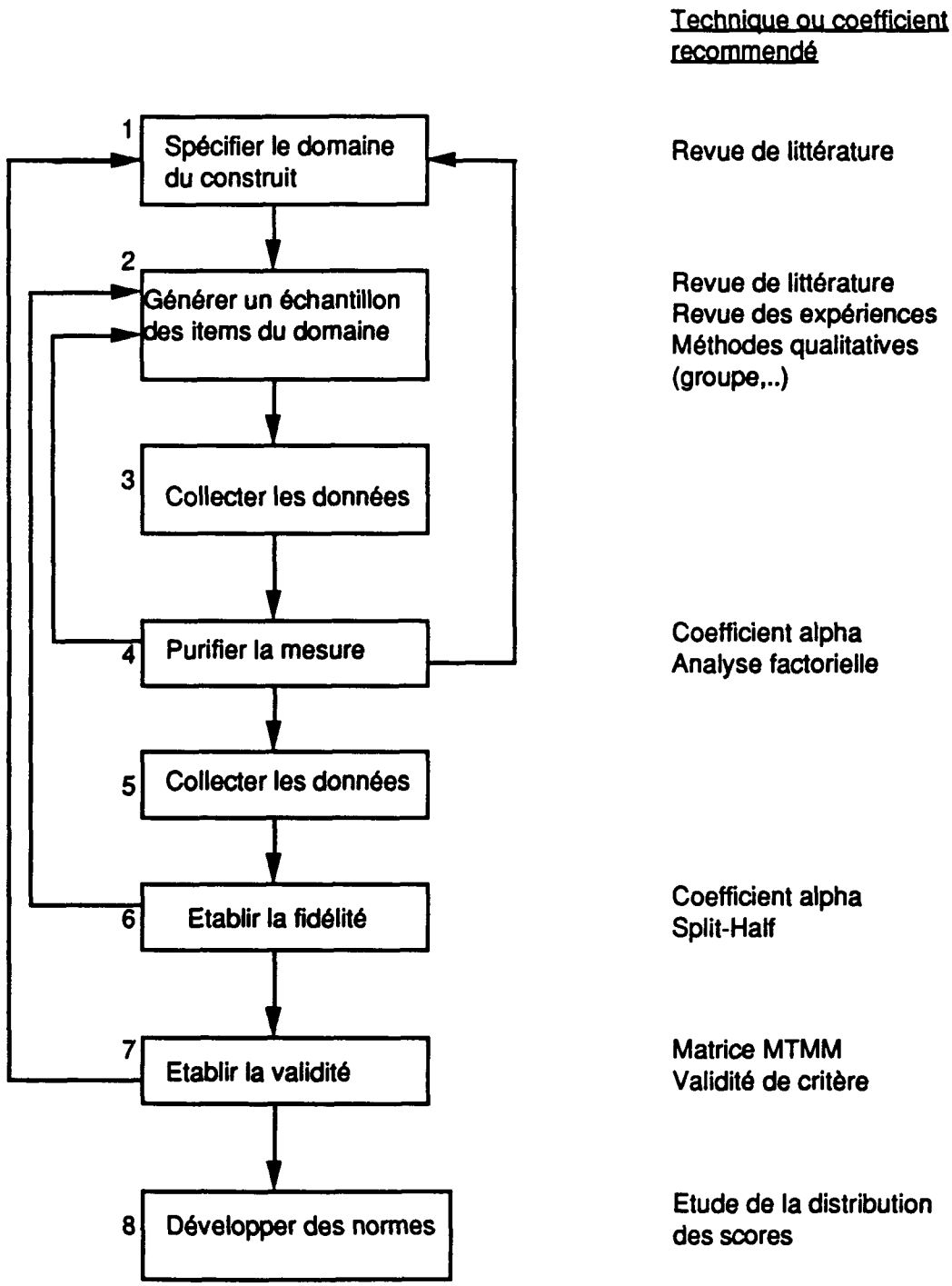
coefficients d'élasticité
intervalles inter-achats
transferts entre marques (brand switching)
parts de choix
etc ...

Nous nous proposons d'effectuer, dans cette recherche, une étude de la validité du concept de concurrence entre les marques tel qu'il est opérationnalisé à travers les mesures qui en ont été proposées dans la littérature Marketing. Nous pourrons alors conclure sur la validité des mesures et des analyses de la structure des marchés qui découlent de ces mesures. Ces études seront menées à partir des corrélations entre les indices calculés sur l'univers "Global" qui est constitué de l'échantillon total des actes d'achat .

Churchill (1979) a proposé un paradigme *"afin de développer de meilleures mesures"* dont nous pouvons, *"ex-post"*, nous inspirer pour étudier la validité du concept de concurrence (voir figure 6 suivante).

Largement inspiré de la théorie du domaine d'échantillonnage, ce paradigme est cohérent avec les éléments théoriques que nous avons présentés dans la section précédente. Nous reprendrons ici la logique de ce paradigme tout en l'adaptant à notre vision de l'étude de la validité des mesures en deux grandes étapes: validité interne et validité externe.

FIGURE 6 PARADIGME POUR L'OBTENTION DE MEILLEURES MESURES



"Adapté de Churchill 1979 p 66"

• 1 Fixer les contours du construits

La première étape dans ce paradigme, consiste à fixer les limites du domaine du construit (concept) étudié, c'est à dire à en définir les frontières. Cette définition passe naturellement par la définition *"théorique"* du concept de concurrence. Notons que le domaine que nous étudions est déjà considérablement restreint puisque nous avons opté pour une approche *"marché"* (par opposition à une approche *"consommateur"*) en étudiant la concurrence entre les marques à partir de données de comportement. Peu de définitions du concept de concurrence sont proposées dans la littérature (voir partie I, chapitre I, section 4). Les articles présentant l'utilisation des indices que nous allons étudier ici n'ont pas abordé de manière claire le problème de la définition de la concurrence sous-jacente à leur indice. Nous pouvons constater que les frontières du domaine du concept de *"concurrence entre les marques"* sont mal définies. Par concéquent, nous pouvons poser l'hypothèse que les différentes mesures de la concurrence qui sont proposées dans la littérature, ne convergent pas vers la mesure du même concept.

Nous retiendrons finalement la définition suivante du concept de concurrence: *"la relation entre deux unités au sein d'un univers de choix, telle qu'elle se manifeste sur le marché à travers les comportements d'achat des consommateurs"* (c'est une définition suffisamment large pour englober tous les indices que nous allons étudier).

• 2 Le choix d'un échantillon des items du domaine

La deuxième phase du paradigme consiste à choisir un échantillon des items appartenant au domaine du construit. A partir de la définition qui est donnée au concept (les frontières fixées au domaine du concept) un échantillon est prélevé parmi l'infinité des connotations possibles de ce concept. Lorsque nous avons présenté les différents indices présents dans la littérature, nous avons utilisé les critères de différenciation suivants :

- indices individuels / agrégés
- indices déterministes / stochastiques

Les indices que nous avons retenu dans le cadre de cette recherche essaient de couvrir les quatre cas résultant du croisement de ces deux critères .

INDICE	INDIVIDUEL	AGREGE
DETERMINISTE-	LEHMANN	MERUNKA
STOCHASTIQUE	RAO ET ET SABAVALA	FRASER ET BRADFORD

Nous considérons que ces indices constituent un échantillon des *"connotations"* du concept de concurrence .

La troisième phase du paradigme concerne la *"purification de la mesure"*, c'est à dire l'obtention de mesures fiables du concept étudié. Cette phase très importante fait l'objet de la section suivante.

SECTION2 FIABILILITE DES MESURES ET ANALYSE DE LA STRUCTURE DES MARCHES

Nous étudierons ici les deux aspects classiques de la fiabilité, à savoir la consistance interne et la stabilité dans le temps.

SS1 COHERENCE INTERNE DE LA MESURE DE LA CONCURRENCE

Dans cette partie, nous allons étudier la cohérence interne des indices étudiés en utilisant plusieurs approches notamment :

- la corrélation moyenne de chaque indice avec tous les autres
- le coefficient alpha de Cronbach
- l'analyse de la structure interne des indices (en utilisant l'analyse factorielle)

Ces trois approches mises en oeuvre sur les quatre marchés étudiés devraient nous permettre d'obtenir des résultats possédant un degré de généralité satisfaisant.

SS1-1 ETUDE DES CORRELATIONS MOYENNES

Selon la théorie du domaine d'échantillonnage, un ensemble d'items appartiennent au même domaine si les corrélations moyennes de ces items entre eux sont égales et suffisamment élevées (voir Churchill 1979).

$r_i(1..N) = \text{constante}$ pour tout $i = 1$ à n

$r_i(1..N)$ est la corrélation moyenne entre N items
$r_i(1..N)$ représente l'importance du "coeur commun" aux différents items.

Ce résultat reste vrai lorsque le concept mesuré possède plusieurs dimensions. Dans ce cas, la fiabilité de la mesure est calculée sur chacune des dimensions. Notons, que l'idée selon laquelle le concept de concurrence comprendrait plusieurs dimensions a été rarement évoquée dans la littérature marketing. Nous pouvons pourtant imaginer que la concurrence comporte plusieurs dimensions. Le concept d'élasticité (si nous admettons qu'il constitue une mesure de la concurrence), en est un bon exemple, puisqu'il peut être calculé relativement à différentes variables d'action marketing.

Nous avons donc calculé les corrélations moyennes entre chaque indice et les autres, sur les quatre marchés. Les résultats figurent dans les tableaux suivants. Nous allons successivement commenter les résultats figurant dans ces tableaux.

TABLEAU 6α VALEURS DES CORRELATIONS MOYENNES MARCHE HUILE

Indice	MOY1	MOY2	MOY3	MOY4
BRAA	.086	-	.130	-
BRAB	-	-	-	-
FRASA	.182	-	.216	-
FRASB	.254	-	.270	-
LEM	.358	.584	.485	.754
MERA	.208	-	.213	-
MERR	230	-	.235	-
SABA	180	-	.232	-
TOTT	269	.454	.389	.676
TRTO	268	.415	.432	.699
RAO	.212	-	.356	-
ELRO	303	.493	.400	.723
TREL	245	.409	.422	.755
IND	331	.536	.473	.739
ACT	315	.516	.474	.739

MOY1: Moyenne de la corrélation de l'indice avec tous les autres
MOY2 : Moyenne de la corrélation de l'indice avec tous les autres après suppression de BRAA FRASA FRASB MERA MERR SABA RAO
MOY3 : Même calcul que MOY1 mais en utilisant le coefficient de Spearman
MOY4 : Même calcul que MOY2 mais en utilisant le coefficient de Spearman

L'examen des corrélations du tableau 6α montre que le niveau de cohérence interne de l'ensemble des indices est faible. De plus, nous pouvons observer que ces corrélations présentent des niveaux différents, ce qui n'est pas un gage d'appartenance de l'ensemble des items au même domaine. Notamment, BRAA semble ne rien partager en commun avec les autres. De même, FRASA, MERA, MERR et SABA ont un niveau de lien moyen avec les autres qui est faible.

A la vue de ces résultats, il paraît peu probable que les indices appartiennent tous au même domaine: ils ne sont pas tous la mesure du même concept de concurrence. Après élimination des indices les moins liés en moyenne aux autres (BRAA, FRASA, FRASB, SABA, MERA, MERR et RAO), nous constatons que le niveau moyen des corrélations restantes s'est nettement amélioré et qu'il s'est homogénéisé. Les indices restant apparaissent comme appartenant au même domaine.
LEM apparait comme l'indice le plus central de la mesure avec une corrélation de .58 .
IND et ACT sont bien liés en moyenne aux autres indices.

Il est intéressant de noter que RAO et SABA, tous deux calculés à partir de données de changements entre marques, sont mal corrélés aux autres indices de même nature (en particulier avec LEM). Nous reviendrons plus loin sur ce point.

Nous venons de mesurer le degré d'association entre les indices à partir de leur coefficient de corrélation linéaire (coefficient de Pearson). Si une relation existe entre deux indices mais qu'elle n'est pas linéaire, une telle mesure risque de sous-estimer la force du lien. Ceci est peut être le cas pour deux indices comme LEM et SABA dont le dénominateur est une forme additive des parts de choix pour le premier et multiplicative pour le second.

Au lieu de mesurer la force du lien linéaire entre deux indices, il est intéressant de mesurer la force du lien monotone qui les unit en utilisant le coefficient de rang de Spearman. Ce faisant, nous assouplissons les conditions que doivent remplir les deux indices pour être liés puisqu'ils n'ont plus qu'à vérifier une relation de type monotone. Nous pouvons ainsi espérer augmenter le lien apparent entre des indices comme RAO, SABA, LEM.

Nous avons donc effectué les mêmes calculs en utilisant cette fois le coefficient de corrélation de rang de Spearman. Les résultats figurent dans le tableau 6α (colonnes MOY3 et MOY4).

Nous pouvons constater que le niveau général des corrélations a augmenté, signe que ce coefficient facilite l'obtention de niveaux de liens élevés (ce qui est logique puisqu'il mesure un degré d'association de type plus général qu'un lien linéaire). Par contre, il n'y a pas d'augmentation relative du lien de SABA et de RAO avec les autres indices.

TABLEAU 6β VALEURS DES CORRELATIONS MOYENNES
MARCHE BOISSON

INDICES	MOY1	MOY2	MOY3	MOY4
BRAA	.116	-	.098	-
BRAB	.176	-	.099	-
FRASA	.318	.329	.313	.352
FRASB	.312	.342	.311	.365
LEM	.389	.525	.400	.499
MERA	.163	-	.159	-
MERR	.169	-	.172	-
SABA	.220	-	.366	.407
TOTT	.226	.269	.237	.292
TRTO	.253	.337	.252	.332
RAO	.342	.461	.333	.461
ELRO	.234	.251	.294	.359
TREL	.237	.324	.240	.334
IND	.341	.465	.371	.451
ACT	.357	.485	.386	.462

MOY1 : Moyenne de l'indice avec tous les autres
MOY2 : Moyenne de l'indice avec tous les autres après suppression de BRAA BRAB MERA MERR SABA
MOY3 : Même calcul que MOY1 mais en utilisant le coefficient de Spearman
MOY4 : Même calcul que MOY3 mais après suppression de BRAA BRAB MERA et MERR

Nous obtenons sensiblement les mêmes résultats que précédemment. BRAA et BRAB sont très faiblement liés aux autres, de même que MERA et MERR. Après élimination des cinq indices les moins reliés au domaine (BRAA, BRAB, MERA, MERR, SABA), nous obtenons des niveaux de corrélations plus élevés et plus homogènes. LEM apparait comme étant l'indice le plus central de la mesure.

Avec l'utilisation du coefficient de Spearman, nous constatons sur ce marché une nette amélioration du niveau moyen des corrélations. Plus particulièrement, les liens qui étaient déjà faibles avec l'utilisation du coefficient de Pearson restent faibles alors que les liens déjà forts le deviennent encore plus.

Le lien de SABA avec les autres s'est considérablement amélioré puisqu'il devient un des indices centraux de la mesure. Cet indice apparait sur ce marché comme ayant une relation non pas linéaire avec les autres mais de type monotone (notamment avec LEM).

TABLEAU 6γ VALEURS DES CORRELATIONS MOYENNES
MARCHE PARFUM

INDICES	MOY1	MOY2	MOY3	MOY4
BRAA	.147	-	.130	-
BRAB	.149	-	.170	-
FRASA	.327	.443	.442	.549
FRASB	.368	.458	.474	.560
LEM	.452	.606	.574	.708
MERR	.142	-	.151	-
SABA	.132	-	.167	-
TOTT	.448	.584	.530	.652
TRTO	.443	.577	.530	.658
RAO	.347	.462	.552	.669
ELRO	.431	.580	.533	.643
TREL	.403	.547	.539	.659
IND	.358	.463	.517	.662
ACT	.409	.525	.521	.672

MOY1 : Moyenne de l'indice avec tous les autres
MOY2 : Moyenne de l'indice avec tous les autres après suppression de BRAA BRAB MERR SABA
MOY3 : Même calcul que MOY1 mais en utilisant le coefficient de Spearman
MOY4 : Même calcul que MOY3 mais après suppression de BRAA BRAB MERR et SABA

Sur ce marché, le niveau moyen des corrélations est supérieur à celui enregistré sur les deux marchés précédents (ce qui caractérise a priori une mesure plus fidèle). Nous constatons encore que BRAA, BRAB, MERR et SABA sont faiblement liés aux autres indices (voir tableau 6γ). Après élimination de ces quatre indices, le niveau moyen des corrélations devient plus élevé et plus homogène.

LEM apparait encore une fois comme un indice central. L'utilisation du coefficient de Spearman accentue encore les résultats observés précédemment. Notons que sur ce marché, le lien moyen de SABA avec les autres ne s'améliore pas pour autant avec l'utilisation du coefficient de Spearman.

TABLEAU 6π VALEURS DES CORRELATIONS MOYENNES
MARCHE SHAMP

INDICE	MOY1	MOY2	MOY3	MOY4
FRASA	.402	.502	.46	.574
FRASB	.423	.534	.493	.616
LEM	.537	.677	.603	.756
MERA	.204	-	.251	-
MERR	.196	-	.198	-
SABA	.314	.392	.536	.782
RAO	.442	.545	.453	.553
IND	.504	.639	.597	.551
ACT	.510	.647	.623	.785

MOY1 : Moyenne de l'indice avec tous les autres
MOY2 : Moyenne de l'indice avec tous les autres après élimination de MERA et MERR
MOY3 : Même calcul que MOY1 mais en utilisant le coefficient de Spearman
MOY4 : Même que MOY2 mais en utilisant le coefficient de Spearman

Sur ce marché, le niveau des corrélations moyennes est satisfaisant. L'indice le plus central reste LEM, avec IND et ACT.

MERA et MERR restent les indices les moins liés.

L'indice SABA se situe à un niveau intermédiaire. L'utilisation du coefficient de Spearman améliore très sensiblement le lien moyen de cet indice avec les autres.

Après élimination de MERA et MERR, nous obtenons des corrélations reflétant un fort degré de cohérence interne.

SS1-2 MESURE DE LA COHERENCE INTERNE AVEC LE COEFFICIENT ALPHA DE CRONBACH

La mesure la plus courante de la cohérence interne d'un ensemble d'items est le coefficient alpha de Cronbach (voir Cronbach 1951). Il est intéressant d'utiliser une telle mesure pour analyser la cohérence interne de la batterie d'indices que nous étudions. Peter (1979) a présenté en détail les modalités du calcul de ce coefficient.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{k} \sigma_i^2 + 2\sum_{i\,j} \sigma_{ij}}\right)$$

k est le nombre d'items de l'échelle multiple

σ_i^2 est la variance de l'item i

σ_j^2 est la covariance entre les items i et j

Cependant, telle qu'elle a été conçue, la formule du coefficient alpha ne peut être directement appliquée ici parce que les indices ne sont pas construits sur des échelles de mesure homogènes. Il n'est pas possible d'effectuer des opérations (qui sont comparatives) entre les différents éléments de la matrice de variance covariance inter-indices puisque, ce faisant, nous donnerions, dans la formule du calcul, plus de poids aux indices qui présentent des niveaux de variance élevés en valeur absolue. Pour contourner le problème, nous avons effectué les calculs sur les indices standardisés (c'est à dire en valeur centrée-réduite). Ceci revient à substituer, dans la formule du alpha, les corrélations (qui sont rappelons le une mesure standardisée de covariance) aux variances et aux covariances.

Nous avons ainsi calculé, pour chaque marché, la valeur du coefficient pour l'ensemble des indices, ainsi que la valeur du coefficient lorsque l'on supprime tour à tour chacun des indices. Ceci permet d'étudier dans quel sens évolue la consistance interne lorsqu'un indice particulier.est supprimé. Si la valeur du coefficient chute fortement cela signifie que l'indice supprimé est bien au coeur de la mesure. Inversement, si la valeur du coefficient augmente quand un indice est supprimé, cela signifie que l'indice en question n'appartient pas au domaine du construit étudié.

Ces calculs ont été programmés avec le langage de la bibliothèque SAS. Les résultats figurent dans le tableau 7 suivant.

TABLEAU 7 VALEURS DU COEFFICIENT ALPHA

	HUILE	BOISSON	PARFUM	SHAMP
TOTAL	.816	.838	.871	.853

Valeur du coefficient après élimination de :

BRAA	-	.847	.877	-
BRAB	.822	.838	.877	-
FRASA	.812	.822	.862	.836
FRASB	.803	.822	.886	.832
LEM	.789	.813	.850	.812
MERA	.808	.839	-	.866
MERR	.804	.838	.878	.867
SABA	.811	.832	.879	.850
TOTT	.801	.832	.850	-
TRTO	.802	.829	.851	-
RAO	.808	.819	.861	.830
ELRO	.796	.831	.853	-
TREL	.805	.831	.855	-
IND	.793	.819	.860	.819
ACT .	.795	.817	.855	.818

Ces résultats confirment les résultats obtenus précédemment à partir des corrélations moyennes. BRAA et BRAB, sur les trois marchés où ils ont été calculés, nuisent à la cohérence interne de la mesure. Les valeurs du coefficient alpha augmentent lorsqu'ils sont supprimés de la mesure. Notons au passage que les variations observées du coefficient sont faibles. Ceci s'explique parce que beaucoup d'indices participent au calcul du coefficient alpha, l'impact de chacun d'entre eux est ainsi minimisé.

D'autre part, lorsqu'un indice est supprimé, il faut tenir compte de l'effet mécanique de diminution du coefficient liée à la diminution du nombre d'indices entrant dans le calcul (voir Cox 1980 à ce sujet). Cet effet mécanique peut compenser (voire dépasser) l'augmentation liée à l'amélioration de la consistance interne de l'échelle, si l'indice supprimé possède peu de variance commune avec les autres .Nous constatons encore que LEM est l'indice le plus central de la mesure puisque le coefficient chute fortement lorsque cet item est supprimé, ceci sur tous les marchés.

Les indices MERA et MERR apparaissent peu liés à tous les autres. SABA est aussi dans le même cas. Tous ces indices détériorent la cohérence interne de la mesure, nous pouvons donc considérer qu'ils n'appartiennent pas au même domaine que les autres.

Par contre, nous pouvons observer que FRASA et FRASB appartiennent assez bien au domaine.

A partir de ces constatations, nous avons recommencé les calculs après suppression des indices qui sont systématiquement les moins liés sur tous les marchés, à savoir BRAA, BRAB, MERA et MERR. Les résultats figurent dans le tableau 8 suivant.

TABLEAU 8 VALEURS DU COEFFICIENT ALPHA APRES SUPPRESSION DE BRAA BRAB MERA MERR

	HUILE	BOISSON	PARFUM	SHAMP
TOTAL	.809	.840	.882	.868

.868

Valeur du coefficient après élimination de :

FRASA	.806	.840	.873	.852
FRASB	.797	.827	.871	.848
LEM	.765	.824	.860	.824
SABA	.813	.802	.893	.864
TOTT	.887	.833	.863	-
TRTO	.792	.833	.863	-
RAO	.809	.828	.872	.846
ELRO	.789	.812	.863	-
TREL	.796	.833	.866	-
IND	.770	.829	.871	.830
ACT	.780	.809	.866	.830

Nous ferons, à partir du tableau 8, les commentaires suivants :

- la valeur du coefficient alpha augmente sur tous les marchés (sauf sur HUILE) montrant ainsi que l'amélioration de la cohérence interne fait plus que compenser la diminution mécanique du coefficient due à la diminution du nombre des indices.
- la valeur du coefficient alpha ne fait que diminuer lorsque, tour à tour, sont supprimés les indices restant (excepté pour l'indice SABA sur les marchés HUILE et PARFUM). Ceci est la preuve que la cohérence interne de l'échelle ne peut plus améliorée.

Comme dans la section précédente, nous avons effectué les mêmes calculs en utilisant le coefficient de Spearman. Les résultats figurent pour les marchés BOISSON et SHAMP dans le tableau 9. Sur ces marchés nous pouvons constater une nette amélioration du degré d'appartenance de l'indice SABA à la mesure et à un moindre niveau une amélioration de celui de l'indice RAO.

TABLEAU 9 VALEUR DU COEFFICIENT ALPHA APRES SUPPRESSION DE BRAA BRAB MERA MERR
(Calcul sur la base du coefficient de Spearman)

	BOISSON	SHAMP
TOTAL	.864	.902

Valeurs du coefficient après suppression de :

FRASA	.850	.893
FRASB	.848	.888
LEM	.831	.871
SABA	.839	.880
TOTT	.857	-
TRTO	.854	-
RAO	.840	.895
ELRO	.849	-
TREL	.854	-
IND	.835	.872
ACT	.834	.867

A partir des résultats précédents (corrélations moyennes et coefficient alpha), nous pouvons en conclure que certains indices sont mal reliés aux autres, ce qui s'explique de deux manières différentes:

- soit le domaine du construit étudié (la concurrence entre les marques) comprend plusieurs dimensions. Les indices caractérisés par un faible niveau d'appartenance au domaine appartiendraient en fait à une autre (ou d'autres) dimension du construit.
- soit les indices mal reliés n'appartiennent pas au domaine du construit

Afin de mieux étudier ce problème, nous avons effectué une série d'analyses factorielles. Ces analyses vont permettre de comprendre la structure interne des indices. Notons, que Guilford (1954 p 403) conseillait déjà l'utilisation de cette technique quand le chercheur pense que le domaine du construit comprend plusieurs dimensions.

SS1-3 ETUDE DE LA COHERENCE INTERNE PAR L'ANALYSE FACTORIELLE

Une telle analyse vient en application directe de la théorie des facteurs (voir section1) qui stipule que le domaine étudié peut être partagé en sous-groupes d'items, chaque sous-groupe correspondant à un facteur, c'est à dire à une facette particulière du construit. Nous avons ainsi mené une série d'analyses factorielles sur les matrices des corrélations inter-indices (calculés sur l'univers GLOBAL) qui ont été constituées pour chacun des marchés. La méthode d'estimation est la plus courante, à savoir l'analyse en composantes principales. Une fois les facteurs principaux extraits, des rotations sont effectuées dans l'espace des facteurs les plus explicatifs. Afin de faciliter l'interprétation, deux grands types de rotations peuvent être menées: des rotations orthogonales (qui contraignent les facteurs à rester orthogonaux) ou des rotations obliques, qui n'imposent plus aux facteurs de rester orthogonaux (voir Harman 1976 chap 13 et 14). Les rotations orthogonales ont été effectuées en utilisant le critère Varimax. Ce critère a pour objet de faire en sorte que chaque item pèse le plus fortement possible sur un ou plusieurs facteurs et le plus faiblement possible sur les autres. Cet effet de contraste est obtenu en maximisant, sur chaque facteur, la variance des carrés des saturations des items sur le facteur. Les rotations obliques consistent à simplifier le pattern des saturations sur les facteurs mais cette fois en permettant aux facteurs de n'être plus orthogonaux. L'intérêt de ce type de rotations, pour nous, est de vérifier si les familles qui se sont constituées sur chacun des facteurs sont indépendantes ou liées. Si elles sont significativement liées, cela indique qu'elles sont les facettes du même domaine et non des familles mesurant des concepts indépendants.

Nous allons commenter les résultats obtenus successivement sur les quatre marchés.

α MARCHE HUILE

L'examen des valeurs propres (voir Annexe 6α) montre que la structure des indices comporte quatre dimensions essentielles (qui correspondent à des valeurs propres supérieures à 1). Ces quatre dimensions expliquent 75% de la variance totale.

Le lien de chaque indice avec un facteur (loading ou encore saturation ou corrélation de l'indice avec le facteur) représente la part de la variance que partage l'indice avec le facteur. C'est une mesure de la fiabilité de l'indice, relativement au domaine correspondant à ce facteur.

Dans cet espace à quatre dimensions, une rotation orthogonale Varimax nous permet d'observer le pattern suivant des facteurs (voir annexe 6α: Factor Pattern , les valeurs sont arrondies et multipliées par cent, les saturations significatives sont repérées par un astérisque *).

La composition des quatre facteurs F1, F2, F3 et F4 est la suivante :

F1 : il est composé des indices LEM,IND,ACT
F2 : il est composé des indices TOTT, TRTO, ELRO et TREL
F3 : il regroupe les indices MERA et MERR ainsi que RAO et SABA (dans une moindre mesure)

F4 : Ce facteur est essentiellement constitué par les indices BRAA, FRASA et FRASB .

Afin de vérifier si la structure obtenue est réellement orthogonale, c'est à dire afin de contrôler si les familles liées à ces quatre facteurs sont indépendantes, nous avons effectué des rotations obliques.

Après rotation oblique (toujours dans le même espace à quatre facteurs), nous pouvons observer que la structure des facteurs n'a pas changé (Annexe 6α).
Par contre, les deux facteurs F1 et F2 sont corrélés (.40), signe que les indices correspondant à ces deux familles sont liés et appartiennent vraisemblablement au même domaine. Les autres couples de facteurs sont par contre nettement orthogonaux, les indices qui les composent partagent donc peu de variance commune.

L'étude des communalités (Annexe 6α), montre que tous les indices appartiennent globalement bien au domaine représenté par les quatre facteurs réunis, excepté les indices BRAA et SABA.

L'examen de ces communalités permet d'observer quels sont les indices dont la variance est bien ou mal reconstituée par l'espace dans lequel la rotation a été effectuée. La communalité d'un indice (voir Harman 1976 chap V) représente le coefficient de corrélation multiple de cet indice avec les facteurs de l'espace dans lequel la rotation est effectuée. C'est son coefficient de corrélation multiple avec la partie de la variance qui est commune à tous les autres indices.
Pour un espace à p dimensions, (ces p dimensions représentent ici p familles d'indices), la communalité représente l'appartenance moyenne de l'indice au domaine constitué par ces p familles d'indices. La communalité peut ainsi être considérée comme une mesure de cohésion de l'indice avec le domaine (au sens de la théorie du domaine) représenté par les p familles (ou facteurs).

Par rapport au modèle de mesure de Churchill (1979 p 65) présenté à la section 1, la communalité d'un indice représente la part de sa variance totale qui n'est ni spécifique à cet indice, ni aléatoire. Au sens de la théorie du domaine, la communalité peut être interprétée comme le pourcentage de variance vraie contenue dans l'item (vraie au sens où cette variance appartient au domaine).

Ceci n'est vrai, bien sûr, que si tous les facteurs caractérisés par l'analyse sont des facettes représentatives du "même" domaine étudié. Dans notre cas, si nous considérons que le facteur F4 qui porte les indices BRAA et BRAB n'appartient pas au même domaine que les autres, alors la communalité n'est plus une mesure de l'appartenance au domaine.

Nous avons déjà noté le faible lien que présente l'indice SABA (et à un moindre niveau RAO) avec les autres indices construits à partir de données de "transferts entre marques". Ce faible lien est dû, selon nous, à la forme multiplicative du dénominateur qui représente le nombre d'échanges attendus sous hypothèse d'indépendance, c'est à dire d'absence de concurrence entre les deux marques. Nous avons jusqu'à présent mesuré le degré d'association entre les indices à partir de leur coefficient de corrélation linéaire (coefficient de Pearson). Si une relation existe entre deux indices, mais qu'elle n'est pas linéaire, une telle mesure risque fort de sous-estimer la force du lien. Ceci est peut être le cas pour deux indices comme LEM et SABA puisque le premier contient à son dénominateur une forme additive et le second une forme multiplicative des parts de choix. Nous avons donc utilisé le coefficient de corrélation de Spearman afin d'étudier la

structure des liens "monotones" entre les indices. La même analyse factorielle a été effectuée en utilisant le coefficient de rang de Spearman dans la matrice des corrélations. Les résultats figurent à l'annexe 7α.

Les quatre premiers facteurs représentent 84% de la variance totale. Nous obtenons donc une structure plus forte qu'avec les coefficients de Pearson. Après rotation orthogonale dans cet espace à quatre facteurs, la structure suivante est caractérisée:

F1 LEM TOTT TRTO ELRO TREL IND ACT
F2 MERA MERR SABA RAO
F3 FRASA FRASB
F4 BRAA

Les deux facteurs F1 et F2 de l'analyse précédente n'en forment plus qu'un seul. Sur F4, BRAA est isolé, il ne partage même pas de variance avec FRASA et FRASB.
Le pattern obtenu après rotation oblique dans cet espace est sensiblement le même. Les quatre facteurs restent orthogonaux (voir Annexe 7α).
Le tableau des communalités montre un degré d'appartenance à l'espace commun qui est plus fort que dans le cas précédent (cf Annexe 7α). Cependant, malgré l'utilisation du coefficient de Spearman, SABA reste l'indice le plus "extérieur" à la mesure .

ß MARCHE BOISSON

L'examen des valeurs propres (voir Annexe 6ß) montre que la structure des indices comporte essentiellement cinq dimensions (qui correspondent à des valeurs propres supérieures à 1). Ces cinq dimensions expliquent 82% de la variance totale .
Dans cet espace à cinq dimensions, une rotation orthogonale nous permet d'observer le pattern suivant des facteurs (voir annexe 6ß: corrélations des indices avec les facteurs après rotation orthogonale).

F1: Il est formé par les indices FRASA, FRASB, LEM, RAO, IND, ACT et SABA.
F2: Il est formé par les indices TOT, TRTO, RAO et TREL
F3: Ce facteur isole les indices BRAA, BRAB, FRASA et FRASB.
F4: Ce facteur est essentiellement constitué par les indices MERA et MERR.
F5: Ce facteur est composé des indices mal représentés sur les quatre premiers facteurs : BRAA, SABA et ELRO

Afin de vérifier si la structure obtenue est réellement orthogonale, nous avons mené une rotation oblique (annexe 6ß). La structure des indices n'apparait pas modifiée, excepté pour SABA qui pèse maintenant significativement sur le facteur F1.
La structure reste assez orthogonale, les corrélations entre les facteurs sont toutes proches de zéro, excepté pour les facteurs 1 et 3 qui ont une corrélation de .28. Cette corrélation de .28 témoigne de l'existence d'un lien entre les indices construits sur la base de transferts entre marques (LEM RAO etc) et les indices construits à partir des intervalles inter-achats dans les versions FRASA et FRASB.

L'examen des communalités permet d'observer quels sont les indices dont la variance est bien ou mal reconstituée par l'espace dans lequel la rotation a été effectuée. Sur ce marché, nous observons (Annexe 6ß) que SABA partage peu de variance avec l'espace commun, sa communalité n'est que de .45 .

La même analyse effectuée à partir des coefficients de rang de Spearman présente les caractéristiques suivantes (voir Annexe 7ß).

Le pattern des facteurs n'est pas modifié dans sa structure. Nous pouvons cependant observer que l'indice SABA pèse maintenant très fortement sur F1. Une rotation oblique ne modifie pas ce pattern. Notons la corrélation de .35 entre F1 et F3, qui confirme le lien existant entre les deux familles correspondantes (l'une calculée sur la base de transferts entre marques, l'autre à partir d'intervalles inter-achats). Pour finir, l'examen des communalités montre que le niveau d'appartenance de chacun des indices au domaine s'est amélioré. L'utilisation d'une mesure de lien monotone améliore le niveau de fiabilité des indices. Notamment, l'indice SABA se révèle avoir maintenant une bonne communalité dans l'espace des facteurs.

γ MARCHE PARFUM

Trois facteurs apparaissent significatifs et expliquent 70% de la variance totale (voir Annexe 6γ). Après rotation varimax dans cet espace à trois facteurs, la structure des indices se présente de la façon suivante :

F1 : Il est composé des indices TOTT, TRTO, ELRO, TREL
F2 : Il est composé des indices FRASA, FRASB, LEM, RAO, IND et ACT
F3 : Ce facteur isole les indices BRAA, BRAB

Une rotation oblique ne modifie pas le pattern des facteurs. Nous pouvons observer que les deux premiers facteurs sont bien liés (corrélation de .40) ce qui caractérise encore une bonne consistance interne entre les indices des deux familles correspondantes. Le troisième facteur reste par contre orthogonal aux deux premiers. L'examen des communalités montre que MERR et SABA partargent très peu avec le domaine défini par les trois facteurs précédents.

La même analyse effectuée sur la matrice des coefficients de rang de Spearman caractérise encore trois facteurs qui expliquent 81% de la variance totale (Annexe 7γ). Le pattern des facteurs n'a pas été modifié, il présente la structure suivante :

F1: Il est composé des indices FRASA, FRASB, LEM, TOT, TRTO, RAO, ELRO, TREL, IND et ACT
F2: Il est composé des indices TOTT, TRTO, ELRO, TREL ainsi que de MERR et SABA
F3: Ce facteur isole les indices BRAA et BRAB

L'observation des communalités indique que l'utilisation de l'indice de Spearman n'améliore pas, sur ce marché, l'appartenance de SABA au domaine (communalité égale à .42). De même l'indice MERR, avec une communalité de .42 est mal reconstitué dans cet espace à trois facteurs.

Après rotation oblique, nous constatons qu'il n'y a pas de changement de la structure. Les facteurs F1 et F2 sont assez fortement liés, avec une corrélation de .43 (voir Annexe 7γ).

A partir des valeurs des communalités, nous observons encore le faible degré d'appartenance de MERR et de SABA au domaine.

π MARCHE SHAMP

Trois facteurs apparaissent significatifs après examen des valeurs propres (voir Annexe 6π) et représentent 86% de la variance totale. Le pattern des facteurs caractérise la répartition suivante des indices en trois familles .

F1: Il est composé par les indices LEM RAO IND et ACT (tous des indices construits à partir de transferts entre marques) .
F2: Il est composé des indices FRASA et FRASB
F3: Il est composé par les indices MERA et MERR

Après examen des communalités, SABA (et RAO à un moindre degré) apparait faiblement lié au domaine. En effet, il n'est relié significativement à aucun des facteurs.

La rotation oblique n'entraîne pas de modification du pattern des facteurs. On observe par contre que les facteurs F1 et F2 deviennent bien corrélés (.43), ces deux familles sont donc deux facettes du même domaine. F3 (MERA et MERR) reste par contre orthogonal aux deux autres, ce qui confirme ce que nous avons déjà constaté sur les marchés précédents, à savoir la non appartenance de ces indices au domaine représenté par les autres.

L'analyse menée sur les corrélations de rang de Spearman fournit les éléments suivants (voir Annexe 7π). Trois facteurs apparaissent significatifs et expliquent 90% de la variance totale. Le pattern des facteurs est le suivant :

F1 : il est constitué par les indices LEM SABA RAO IND et ACT
F2 : il est constitué par les indices FRASA et FRASB.
F3 : il est constitué par les indices MERA et MERR

Nous constatons que l'indice SABA est maintenant mieux rattaché au domaine, il pèse fortement sur F1.

L'examen des communalités confirme l'amélioration du degré d'appartenance de SABA au domaine, ce qui n'est par contre pas le cas pour RAO.

Sur l'analyse oblique, encore une fois, apparaît une forte corrélation (.53) entre les deux familles F1 et F2 ainsi que la nette indépendance de la troisième (F3) qui correspond aux indices MERA et MERR.

Nous pouvons maintenant procéder à deux types de conclusions concernant l'étude de la fiabilité des indices de mesure de la concurrence.

1 D'UN POINT DE VUE METHODOLOGIQUE

L'utilisation du coefficient de corrélation de rang, qui mesure la force d'un lien de type monotone, accentue le lien entre les indices. Plus exactement, il diminue le niveau des relations les plus faibles et inversement augmente le niveau des liens les plus forts, ce qui a pour effet de renforcer la structure interne des indices.

Sur les quatre marchés, nous avons observé, à nombre de facteurs égal, une augmentation du pourcentage de variance expliquée et une augmentation des valeurs moyennes des communalités qui accompagne l'utilisation du coefficient de corrélation de Spearman. L'utilisation du coefficient de rang améliore le degré de cohérence interne de la mesure, qui est la forme principale de la fiabilité. De plus, sur deux marchés, il améliore très sensiblement le degré de fiabilité de l'indice SABA.

2 DU POINT DE VUE DE LA MESURE

Concernant la fiabilité des indices de mesure de la concurrence, nous avons caractérisé trois grandes familles naturelles d'indices à savoir :

F1: indices calculés à partir des intervalles inter-achats, dans les versions FRASA et FRASB.
F2: indices calculés à partir de transferts entre marques (brand switching).
F3: indices construits sur les parts de choix des marques.

Les deux indices BRAA et BRAB (correspondant à la version originelle proposée par Fraser et Bradford) ne sont liés à aucune de ces trois familles. En particulier, ils s'associent mal avec les indices FRASA et FRASB qui pourtant sont eux aussi calculés à partir d'intervalles inter-achats (voir Chapitre IV section 3). Si trois grandes familles apparaissent, les analyses obliques ont montré que deux d'entre elles sont fortement liées: ce sont les familles F1 et F3.

Les indices FRASA, FRASB, LEM, RAO, TOTT, TRTO, ELRO, TREL, IND et ACT sont globalement bien liés. Même quand ces indices sont "saturés" sur des facteurs différents, nous avons constaté que ces facteurs sont corrélés dans les rotations obliques. LEM, de même que IND et ACT, apparaissent sur presque toutes ces analyses, comme étant les indices les plus au "coeur" de l'ensemble des indices. RAO et surtout SABA, appartiennent à ce même univers mais avec un faible degré de cohésion (notamment pour l'indice SABA). Nous dirons que l'indice SABA appartient au même domaine dont LEM est l'élément central. En effet, même si le lien de SABA (son loading avec le facteur concerné) est faible sur presque tous les marchés, nous avons pu noter la présence d'un lien de cet indice avec cette famille et uniquement cette famille. Nous en conclurons qu'il appartient au même domaine que LEM, mais que son degré de fiabilité est en général insuffisant.

Notons que cette fiabilité s'est considérablement améliorée sur deux marchés (BOISSON et SHAMP) avec l'utilisation du coefficient de corrélation de rang de Spearman. Ceci n'est pas le cas des indices MERA et MERR qui eux, de toute évidence, appartiennent à un autre domaine. La famille constituée par les indices MERA et MERR reste toujours indépendante des deux premières. Nous pouvons considérer que les indices

MERA et MERR appartiennent à un domaine différent de tous les autres. Ils mesurent un autre concept que les indices précédents .

Pour finir, nous soulignerons que les deux types d'analyses précédentes de la fiabilité (par les corrélations moyennes et par le coefficient alpha), n'avaient pas permis de caractériser aussi finement un tel phénomène. Nous ne pouvons donc, à l'issue de cette phase, que recommander l'utilisation de l'analyse factorielle pour l'étude de la cohérence interne d'une échelle de mesure multiple.

SS2 MESURE DE LA FIABILITE DANS LE TEMPS (TEST-RETEST)

Le coefficient le plus fréquemment utilisé pour étudier la fiabilité d'une mesure est le test-retest. Rappelons que ce test revient à comparer deux mesures pratiquées à l'aide du même instrument, en deux époques différentes (en général espacées de deux semaines).

Comme nous l'avons noté dans la section précédente, beaucoup de psychomètres (par exemple Nunnally 1978 p 233) déconseillent l'utilisation d'un tel test ou bien, réduisent la portée des conclusions qu'il permet de tirer. Il est en effet reproché à ce test de surestimer le niveau de la fiabilité. Nunnally propose de l'utiliser en tant que condition nécessaire de fiabilité, sa valeur devant être élevée pour que la mesure présente des signes de fiabilité.

Il est par contre des cas où les auteurs conseillent l'utilisation du test-retest, notamment quand le chercheur est intéressé par l'obtention de mesures stables dans le temps. Nunnally (p 235) parle même de test-retest sur le long terme (long run test) en évoquant des périodes espacées de six mois.

C'est dans ce cas que nous nous trouvons avec la mesure de la concurrence entre les marques. Il est en effet souhaitable, dans ce contexte, d'obtenir des mesures présentant un gage de stabilité dans le temps suffisant. Rappelons à cet effet, qu'une des hypothèses sur laquelle reposent de nombreux modèles stochastiques de choix est la *"stabilité des préférences qui sont à l'origine des comportements observés"* (voir Massy, Montgomery et Morrison 1970). Le modèle de Hendry que nous avons décrit dans la première partie, au chapitre II, section 2, repose sur une telle hypothèse (voir Kalwani et al 1977). Etudier la stabilité des indices de concurrence dans le temps apparait donc comme un élément très important, puisque la stabilité des préférences (dans le contexte des modèles stochastiques de choix il faut entendre par préférence la préférence révélée par les choix, c'est à dire la part de choix de la marque) devrait se retrouver en terme de stabilité des indices de concurrence entre marques.

Les coefficients de test-retest vont en fait nous renseigner à un double point de vue.

1- Sur la stabilité des mesures quand elles sont répétées dans le temps, ce qui est l'objectif classique du test.

2- Sur la robustesse des mesures à un partage en deux de l'échantillon des observations retenues pour effectuer la mesure. En effet, dans notre cadre particulier de recherche, il a fallu découper la période d'observation en deux sous-périodes de six mois (et donc l'échantillon des actes d'achats en deux sous-échantillons de taille égale) afin de pouvoir effectuer des mesures sur deux périodes du temps. Ce test mesure ainsi une certaine forme de robustesse des indices, à savoir la stabilité des estimateurs des indices de concurrence entre les marques lorsque la taille de l'échantillon est partagée en deux.

La faiblesse de cette approche réside dans le fait qu'il sera difficile de juger lequel des deux effets possibles (instabilité dans le temps ou sensibilité à la taille de l'échantillon) est à l'origine d'un faible coefficient de test-retest.
Les résultats figurent dans le tableau 10 suivant.

TABLEAU 10 INDICES DE FIABILITE DANS LE TEMPS TEST RETEST (coefficient de corrélation de Pearson)

	HUILE	BOISSON	PARFUM	SHAMP
FRASA	-	.38	.08 *	.31
FRASB	-	.43	-.02 *	.17
LEM	.84	.78	.77	.52
MERA	-	.03 *	-	-.01 *
MERR	.17 *	.19	.28	.01 *
SABA	.50	.32	.01 *	.18
TOTT	.70	.38	.31	-
TRTO	.15 *	.14 *	.87	-
RAO	.49	.40	.29	.17
ELRO	.43	.45	.45	-
TREL	-.12 *	.13 *	.85	-
IND	.93	.92	.93	.81
ACT	.90	.92	.92	.82

*: non significatif au seuil de risque alpha = 5 %. Le nombre de marques étant différent entre les marchés sur lesquels sont calculées les corrélations, les valeurs du seuil de signification du coefficient de Pearson sont différentes d'un marché à l'autre.

Les résultats, pour certains indices, sont assez surprenants, nous allons donc nous attarder à les commenter.

Tout d'abord, le premier résultat qui apparait est que les données de transferts bruts (indices IND et ACT) sont extrêmement stables sur le long terme (long terme étant entendu pour des périodes de six mois). Nous observons en effet des corrélations supérieures à .90 pour les deux indices IND et ACT sur les quatre marchés. De plus (voir table 6), les corrélations entre ces deux indices (IND et ACT) sur une même période sont encore plus fortes que la valeur du test-retest (corrélation de l'indice avec lui-même, quand il est calculé sur deux périodes différentes).

TABLEAU 11 LIEN ENTRE IND ET ACT

	HUILE	BOISSON	PARFUM	SHAMP
INDt1.ACTt1	.999	.957	.957	.993
INDt1.ACTt2	.988	.885	.914	.813
INDt2.ACTt1	.931	.891	.882	.808
INDt2.ACTt2	.977	.965	.966	.984

Sur les quatre marchés, ces corrélations sont extrêmement fortes, ce qui nous apporte deux résultats intéressants :

- il y a pratiquement identité entre transfert en *"nombre de ménages ayant changé entre deux marques"* et transferts en *"nombre d'actes de changement entre ces deux marques"*. Rappelons que l'indice ACT est une version de l'indice IND qui est pondérée par les transferts effectués par les ménages/. Ceci constitue un résultat très important. La structure des transferts entre marques est similaire lorsqu'elle est évaluée à l'aide de transferts pondérés ou non pondérés (indice IND) par la fréquence d'achat des ménages (indice ACT).
- la structure des transferts bruts entre marques est extrêmement stable sur le long terme ainsi qu'au partage en deux de l'échantillon, ce qui n'est pas le cas de tous les indices ceci constitue encore un résultat très important.

C'est la raison pour laquelle, de par son mode de calcul, l'indice LEM est parmi tous les autres indices, celui qui présente les meilleurs scores en test-retest. Par contre, les indices RAO et SABA, qui du point de vue de la nature des données sont similaires à LEM, ne sont pas aussi fiables sur le long-terme.

Nous avons déjà présenté la raison, à notre avis, d'une telle instabilité qui est due à l'utilisation du produit des probabilités marginales de choix des deux marques pour évaluer un flux de changements attendu sous hypothèse d'indépendance.

$$S_{ij} = \frac{N_{ij}}{N_{i.}\, N_{.j}}$$

S_{ij} : Indice de concurrence entre les marques i et j
N_{ij} : Transferts observés entre i et j
$N_{i.}\, N_{.j}$: Transferts entre les marques i et j attendus sous hypothèse d'indépendance des deux marques

De faibles variations dans l'estimation de $N_{i.}$ et $N_{.j}$ (estimation dans la mesure où nous ne disposons que d'un échantillon des actes d'achats sur le marché) peuvent entraîner de fortes variations sur S_{ij}. Ceci est d'autant plus vrai, comme c'est le cas avec SABA, si les transferts sont comptabilisés en *"actes"*, qui connaissent des variations en valeurs absolues encore plus fortes.

Pour les indices BRAA et BRAB, les valeurs en test-retest n'ont pas pu être calculées, les valeurs des indices sur les deux strates de temps n'ayant pas elles-mêmes été calculées.

Pour les autres indices, les coefficients de test-retest ont été calculés en utilisant alternativement les coefficients de Pearson et de Spearman .

Nous pouvons constater que les deux versions "Défense", ELRO et TOTT, sur les trois marchés où ils sont calculés, présentent des niveaux de fiabilité satisfaisants et comparables. Par contre, nous observons que sur HUILE et BOISSON, les deux indices en version "Attaque", à savoir TRTO et TREL, sont peu fiables.

Ces observations suggèrent le fait qu'un marché peut apparaître comme stable selon un point de vue (ici "Défense") et au contraire se révéler instable selon un autre point de vue (ici "Attaque"). Ces éléments tendent à montrer que la concurrence n'est pas symétrique.

Les indices MERA et MERR apparaissent, sur tous les marchés, comme très sensibles à l'effet du temps, donc peu fiables selon l'esprit du test-retest. Comment interpréter ce résultat alarmant qui tend à indiquer qu'un même marché présente des structures très différentes d'une période à l'autre ?

D'une part, puisque ces indices sont calculés, dans un premier temps, au plan individuel, la réduction de la période en deux tranches de six mois peut biaiser considérablement les estimations au plan individuel. En effet, pour un indice agrégé, même si les comportements de chaque individu sont modifiés, les comportements peuvent se compenser : certains changements de comportements rencontrés chez certains individus peuvent être compensés par les changements inverses du comportement d'autres individus. Lorsque les estimations sont faites au plan individuel, ceci n'est plus vrai et ne le devient pas lorsqu'on agrège les estimations individuelles. Bref, le comportement des individus n'est pas stable alors que celui de la population le serait.

D'autre part, comme les estimations sont faites sur la base des quantités achetées, l'indice peut subir l'effet de saisonnalité de la consommation (notamment sur les marchés Parfum et Boisson).

Les indices FRASA et FRASB présentent un niveau de fiabilité médiocre dans le temps, excepté sur le marché PARFUM où ils sont tout à fait instables.

L'observation des niveaux de fiabilité obtenus avec le coefficient de Spearman, ne change rien à ce qui a été dit précédemment (voir tableau 12).

TABLEAU 12 INDICES DE FIABILITE DANS LE TEMPS :TEST RETEST (Coefficient de Spearman)

	HUILE	BOISSON	PARFUM	SHAMP
FRASA	-	.28	.19	.39
FRASB	-	.41	-.02 *	.28
LEM	.76	.64	.53	.46
MERA	-	.06 *	-	-.04 *
MERR	.28	.18	.30	.02 *
SABA	.49	.41	.17	.20
TOTT	.33	.38	.53	-
TRTO	.12 *	.24	.87	-
RAO	.45	.38	.36	.22
ELRO	.23	.59	.59	-
TREL	-.17	.30	.83	-

* Non significatif au seuil de risque alpha = 10%

CONCLUSIONS DES ETUDES SUR LA FIABILITE DE LA MESURE DE LA CONCURRENCE

Nous venons d'étudier la fiabilité de plusieurs indices de mesure de la concurrence, en utilisant plusieurs types d'analyses :

- fiabilité interne
- test-retest

A partir des résultats obtenus, nous pouvons porter deux types de conclusions concernant la qualité de la mesure de la concurrence entre les marques et les aspects méthodologiques d'une étude de fiabilité de ce type.

CONCLUSIONS RELATIVES A LA FIABILITE DE LA MESURE

Au sein de l'échantillon des mesures de la concurrence que nous avons prélevé dans la littérature portant sur l'analyse de la structure des marchés à partir de données de comportements, il n'y a pas un bon niveau de cohérence des mesures obtenues, au sens de la théorie du domaine. Tous les indices n'appartiennent pas au domaine représentatif du même construit. Trois familles bien indépendantes se sont dessinées :

- Une première famille, la plus importante, qui comprend tous les indices construits à partir des données de transferts entre marques (LEM, SABA, RAO, TOTT, TRTO, ELRO, TREL, IND et ACT) mais aussi FRASA et FRASB qui sont construits à partir des intervalles inter-achats des marques. Même si, dans les analyses factorielles, nous avons pu voir peser ces indices sur différents facteurs, ces facteurs étaient (dans les rotations obliques) suffisamment corrélés pour démontrer le degré de lien entre ces indices .

SABA apparait en général peu lié avec cette famille, de même que RAO, dans une moindre mesure. De tels indices mesurent pourtant le même concept mais comportent manifestement un fort degré d'erreur de mesure, au sens de la théorie du domaine d'échantillonnage. Nous avons proposé une explication à cela, en considérant que la forme multiplicative du dénominateur de cet indice crée des perturbations dans le calcul.
L'indice LEM apparait comme l'élément central de cette famille, que ce soit en terme de consistance interne ou de stabilité dans le temps.

- Une deuxième famille comprend les deux versions de l'indice de Merunka et Bourgeat que nous avons calculées, indices construits sur la base des parts de choix des marques. Ces deux versions sont très fortement liées. Cette mesure de la concurrence ne partage apparemment rien de commun avec la famille précédente. Les deux indices sont presque toujours isolés sur une dimension particulière et indépendante de toutes les autres. Nous constatons ici l'orthogonalité (l'indépendance) des domaines dont sont issues les deux familles d'indices précédentes.
D'autre part, ces indices apparaissent comme ayant une stabilité dans le temps (fiabilité), ou en tout cas au partage en deux de l'échantillon, très faible .

- Une troisième famille est constituée par les indices BRAA et BRAB, qui sont issus du modèle proposé par Fraser et Bradford (1983). Ces deux indices sont fortement indépendants de tous les autres indices. Notamment, ils partagent très peu de variance avec les deux indices FRASA et FRASB qui sont eux aussi calculés à partir d'estimations des intervalles inter-achats, mais en prenant en compte la totalité des marques en concurrence pour la définition de ces périodes (voir chapitre IV section 3).

Enfin, nous avons montré que les mesures de transferts bruts entre marques (indices IND et ACT) sont significativement plus stables dans le temps que les autres indices.

CONCLUSIONS RELATIVES AUX PROBLEMES METHODOLOGIQUES LIES A L'ANALYSE DE LA FIABILITE

Nous avons montré que l'analyse factorielle, dont la base théorique est proche de celle de la théorie de la mesure, offre un outil très complet pour analyser la cohérence interne d'une échelle multiple. Les différents résultats offerts par cette méthode (pourcentages de variance expliquée, saturations, communalités, rotations, corrélations entre les facteurs) sont autant d'informations précieuses car elles sont cohérentes avec la théorie du domaine d'échantillonnage.

D'autre part, nous avons pu constater que l'utilisation d'un indice particulier d'association entre les mesures (linéaire ou non-linéaire) peut conditionner dans certains cas les résultats obtenus en matière de fiabilité. Notamment, l'utilisation d'un indicateur de lien non-linéaire augmente le niveau apparent de la fiabilité. Ceci peut se justifier de deux façons différentes:

- soit les liens entre deux mesures existent mais sont non-linéaires, dans ce cas il est préférable d'utiliser un indicateur de lien non-linéaire (comme le coefficient de Spearman) pour procéder à des analyses de fiabilité.

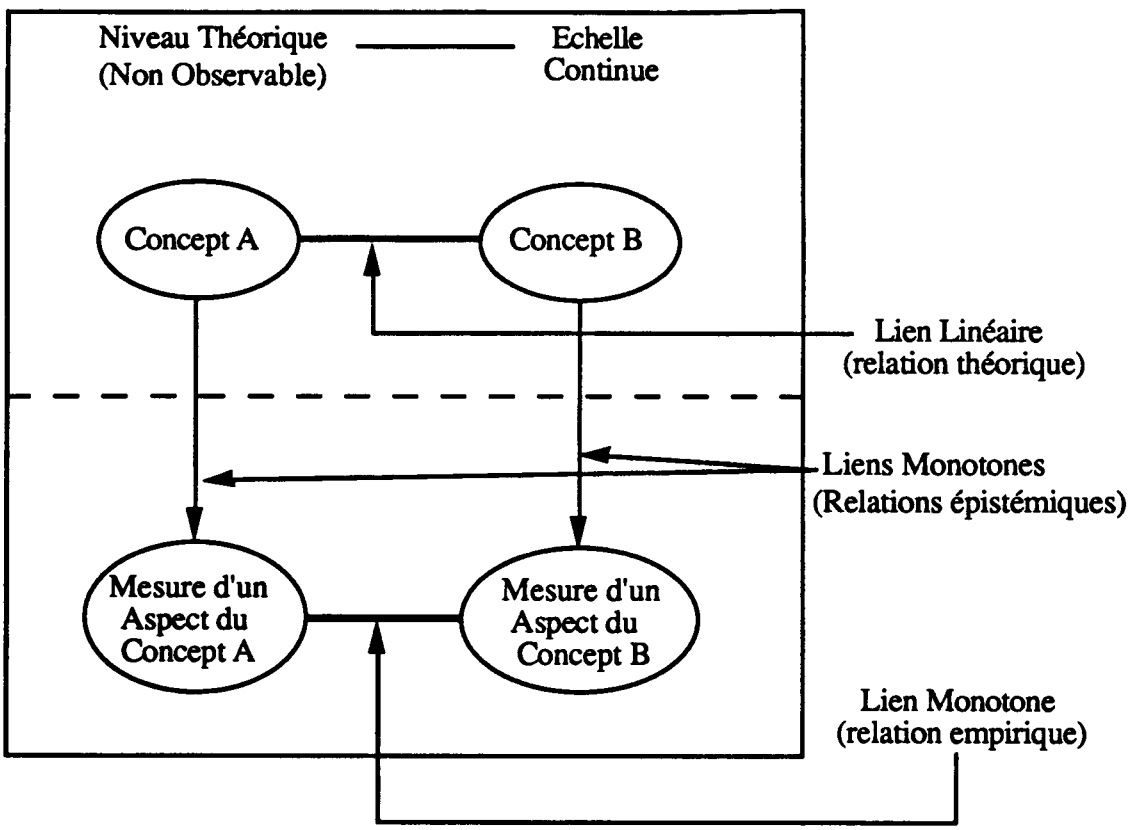
- soit les liens entre les deux mesures sont linéaires, mais c'est l'erreur de mesure qui les fait apparaître comme "non linéaires". Dans ce cas, il est préférable d'utiliser une mesure de lien linéaire (le coefficient de Pearson). Bohrnstedt et Borgatta (1980, chap 1) ont critiqué l'utilisation des coefficients de rang pour apprécier le lien entre des échelles de type ordinal parce que ces coefficients pour être calculables utilisent des propriétés métriques. Notamment, pour pouvoir additionner les rangs, ils faut faire l'hypothèse qu'il y a le même écart entre tous les rangs (ceci n'est pas vrai pour le coefficient de Kendall qui se calcule de manière purement ordinale).

Pour avoir plus d'éléments concernant cette cette deuxième explication nous renvoyons le lecteur au chapitre du livre de Bohrnstedt et Borgatta (chap 1, 1980). Le modèle suggéré par ces deux auteurs est le suivant:

- au niveau théorique (latent) la plupart des concepts que nous étudions (comme la concurrence entre deux marques) sont de type métrique et la répartition de leurs valeurs est distribuée selon la loi normale.
- au niveau empirique, les mesures que nous en faisons sont *'grossières'* et entachées d'erreur de mesure.
- c'est pourquoi il existe un lien monotone (non linéaire) entre le concept latent et sa mesure ou bien entre plusieurs mesures.

Ceci est une autre manière de considérer que nos mesures sont affectées par l'erreur. Ce que nous prenons pour "relation monotone" ne serait en fait que le résultat de *"l'erreur de mesure"*, selon cette théorie. Ce n'est donc pas en utilisant des méthodes "ordinales" que l'on analyserait correctement les données. La figure suivante illustre ce point de vue.

FIGURE 7 TYPE D'ECHELLE ET NIVEAU DE MESURE



SECTION 3 ETUDE DE LA VALIDITE DE TRAITS

Les phases 5 et 6 du paradigme de Churchill (voir Figure 6 à la section 1) ne présentent pas d'intérêt dans le cadre de cette recherche. Elles correspondent en effet à la validation croisée sur de nouvelles données des résultats de la phase précédente.

Nous aborderons directement la septième phase qui consiste à établir la validité dite "de traits" de la mesure.

Cette validité comporte deux grandes dimensions qui sont la validité convergente et la validité discriminante et que nous nous proposons d'étudier ici.

La matrice Multi-Traits et Multi-Méthodes (MTMM) présentée par Campbell et Fiske (1959), constitue certainement l'instrument le plus satisfaisant que nous ayons à notre disposition. C'est d'ailleurs le plus connu et le plus utilisé même s'il présente un certain nombre de faiblesses (voir section1).

SS1 APPROCHE CLASSIQUE DE LA VALIDITE DE TRAITS

Nous allons appliquer ici les principes de base de la matrice MTMM à l'étude de la validité de la mesure de la concurrence entre les marques.

Rappelons que le principe de cette matrice est de faire des comparaisons entre différentes mesures, effectuées sur différents traits (concepts hypothétiques) et d'en tirer des conclusions:

- quant à la capacité des méthodes à "converger" lorsqu'elles mesurent le même traits (concept).
- quant à la capacité des méthodes à "discriminer" les différents traits (concepts).

En théorie, suivant les recommandations de Campbell et Fiske (C&F), il est souhaitable que :

- les instruments de mesure (méthodes) soient aussi différents que possible afin que le test de convergence ne soit pas attribuable à la similarité des instruments
- les traits mesurés (concepts) ne soient pas trop différents, afin que le test de discrimination ne soit pas trop facile.

Dans notre cadre d'analyse, les traits sont au nombre de quatre et sont constitués par les marchés sur lesquels nous avons calculé les indices .

Trait1 HUILE
Trait2 BOISSON
Trait3 PARFUM
Trait4 SHAMP

Les méthodes sont représentées par les différents indices qui ont été calculés (quinze indices en tout).

I1 BRAA
I2 BRAB
I3 FRASA
I4 FRASB
I5 LEM
I6 MERA
I7 MERR
I8 SABA
I9 TOTT
I10 TRTO
I11 RAO
I12 ELRO
I13 TREL
I14 IND
I15 ACT

Les différentes méthodes ne sont pas indépendantes puisque certains indices sont calculés sur des bases très voisines (par exemple LEM, IND, ACT, RAO ...). Néanmoins, ce problème d'indépendance ne doit pas constituer un obstacle définitif aux analyses de validité. Comme plusieurs auteurs l'ont noté (en particulier C&F), il est rare de disposer, dans la réalité, d'instruments de mesure parfaitement indépendants, surtout quand comme ici, il faut mener des analyses avec des instruments de mesure déjà existants. L'essentiel par contre est de bien rester conscient du fait que les méthodes partagent une part plus ou moins grande de variance entre elles et d'évaluer l'importance de cette variance afin d'en tenir compte dans la comparaison des corrélations.

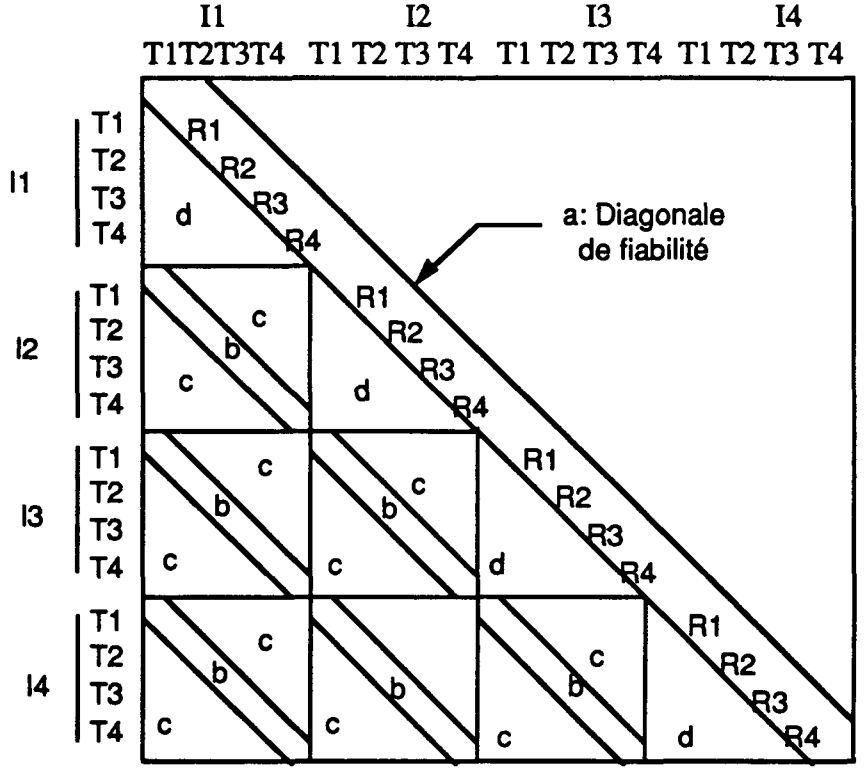
D'ailleurs, Campbell et Fiske soulignaient déjà qu'une telle indépendance est difficile à obtenir dans la réalité mais que l'on peut quand même apprécier la validité. Comme le notent Campbell et Fiske p 84: *" What if the overlap of method variance were higher ? ...The degree of elevation of the validity diagonal above the heterotrait heteromethod triangles remains comparable and relative validity can still be evaluated ..".*

Ici, nous pouvons considérer que les instruments les plus indépendants sont ceux qui sont calculés à partir de types de données différents. Coexistent ainsi trois grandes catégories:

- les indices construits sur la base des intervalles inter-achats
- les indices construits sur la base des parts de choix
- les indices construits sur la base des transferts entre marques

Nous allons préciser maintenant le contenu de la matrice MTMM construite dans le cadre de cette recherche. La matrice fictive suivante étudie la validité de quatre indices (I1, I2, I3, I4) calculés sur quatre marchés (T1 à T4).

FIGURE 8 MATRICE MTMM POUR QUATRE INDICES CALCULES SUR QUATRE MARCHES



Nous trouvons dans cette matrice, successivement, les éléments suivants :

a : diagonale de fiabilité
b : diagonales de validité
c : triangles hétéro-marchés hétéro-méthodes
d : triangles hétéro-marchés mono-méthode

- dans la diagonale principale (zone a), figurent les coefficients de fiabilité obtenus pour chaque indice en test-retest. Ces coefficients représentent, rappelons le, la stabilité de l'indice dans le temps ainsi que sa robustesse à un partage en deux de l'échantillon (voir section 2 précédente).
- les diagonales de convergence (zones b) contiennent les corrélations entre des mesures effectuées à l'aide d'indices différents, calculés sur le même marché. C'est à partir de ces valeurs que nous pourrons (ou non) établir la convergence des instruments de mesure. La condition "minimale" requise par C&F est que ces corrélations soient significativement supérieures à zéro et d'un niveau suffisant pour prétendre poursuivre l'investigation. Dans notre étude, obtenir des valeurs élevées dans les diagonales de validité signifie que la même structure de relations entre les marques est obtenue à partir de deux indices de mesure différents. Pour que le

concept de concurrence soit "identifiable", il est important que ces valeurs soient élevées.

- les triangles hétéro-traits et hétéro-méthodes (zones c) sont constitués par les corrélations de mesures effectuées en utilisant différents indices calculés sur différents marchés. Elles représentent la variance commune à deux mesures due à la fois aux objets mesurés mais aussi à une cause systématique comme le lien entre les indices. Il s'agit encore ici d'un effet de méthode. Ces corrélations ont beaucoup d'importance. Il faut en tenir compte pour interpréter le niveau des corrélations dans les diagonales de validité (zone a). En effet, le niveau d'une corrélation de validité ne peut être directement interprété si les indices correspondants présentent un niveau de lien important dans les triangles de type c. Un niveau de convergence (une corrélation de la zone a) doit finalement être établi une fois éliminé l'effet systématique qui n'est pas dû à l'objet mesuré (les valeurs correspondantes dans les zones c).
- les triangles hétéro-traits mono-méthode (zones d) sont constitués par les corrélations de mesures effectuées avec le même indice, calculé sur les différents marchés. Ces corrélations permettent d'évaluer le degré d'effet de mesure lié à un instrument (indice) particulier. Si ces valeurs sont faibles, cela signifie qu'il y a peu de variance dans la mesure qui est attribuable à l'instrument lui même. Il est ainsi souhaitable d'avoir de faibles valeurs (non significativement différentes de zéro) dans ces triangles.

Pour pouvoir comparer des indices qui sont calculés sur des marchés comportant un nombre de marques différents (respectivement 10, 14, 15, 24 marques) nous avons ramené ce nombre de marques à 10 sur tous les marchés. Pour cela nous avons éliminé les marques aux parts de marché les plus faibles en ne conservant à chaque fois que les dix marques les plus importantes.

SS1-1 ETUDE DU CRITERE DE VALIDITE CONVERGENTE

Campbell et Fiske proposent une condition nécessaire, mais non suffisante, de validité qui est que les corrélations dans les diagonales de validité (zone b) soient significativement non-nulles et d'un niveau suffisant pour prétendre poursuivre plus loin l'étude.

Bien que ce critère soit simple, toute la difficulté réside dans son opérationnalisation, notamment dans notre cas où la matrice comporte un nombre important d'indices.

Nous proposons dans un premier temps de compter, pour chaque indice, le nombre de corrélations significativement non-nulles qu'il compte dans les diagonales de validité, par rapport au nombre total de corrélations qui figurent dans ces diagonales (zones b). Les résultats figurent dans le tableau suivant.

TABLEAU 14 VALIDITE CONVERGENTE (*) (Coefficient de corrélation de Pearson)

	HUILE	BOISSON		SHAMP	SCORE TOTAL et % **
BRAA	1/13	2/14	1/13	-	4/40 10
BRAB	-	4/14	1/13	-	5/27 19
FRASA	4/13	9/14	9/13	6/8	28/48 58
FRASB	7/13	8/14	10/13	6/8	31/48 65
LEM	9/13	9/14	10/13	6/8	34/48 71
MERA	5/13	4/14	-	1/8	10/35 29
MERR	5/13	3/14	4/13	1/8	13/48 27
SABA	5/13	8/14	4/13	6/8	23/48 48
TOTT	6/13	4/14	10/13	-	20/40 50
TRTO	9/13	7/14	10/13	-	26/40 65
RAO	8/13	10/14	9/13	6/8	33/48 69
ELRO	9/13	7/14	10/13	-	26/40 65
TREL	8/13	6/14	10/13	-	24/40 60
IND	9/13	7/14	10/13	6/8	32/48 67
ACT	9/13	9/14	10/13	6/8	34/48 71

* Nombre d'éléments dans les diagonales de validité qui sont significativement non-nuls au seuil de risque alpha 5% et nombre total de ces éléments .
** Score total tous marchés confondus : c'est la somme des valeurs obtenues sur les quatre marchés. Figure à côté de ce score, le pourcentage d'éléments significatifs ramené au nombre total d'éléments.

Dans ce tableau, figurent pour chaque indice et chaque marché, le nombre de corrélations significativement non-nulles (au seuil alpha de 5 %) et le nombre total d'éléments des diagonales de validité correspondant à l'indice. La colonne "Score Total" indique le score obtenu par l'indice sur les quatre marchés. Ce score figure en valeur absolue et en pourcentage. Exemple de lecture: LEM vérifie 34 fois sur 48 le critère de convergence , pour les quatre marchés, ce qui représente un taux de 71 % de validité convergente.

Sur les trois marchés où ils sont calculés, BRAA et BRAB ne convergent avec aucun autre indice, excepté entre eux. A la vue de scores aussi faibles, nous en concluons que les deux indices mesurent autre chose que le domaine constitué par l'ensemble des autres indices. Par conséquent, nous n'irons pas plus loin dans l'étude de la validité de ces deux indices.

FRASA et FRASB convergent assez bien, respectivement 58% et 65% de leurs diagonales de convergence sont significativement non-nulles.

LEM, IND, ACT, RAO sont les indices bénéficiant des meilleures validités convergentes .

De la même manière, TOTT, TRTO, ELRO et TREL présentent de bons niveaux de convergence. L'examen de la matrice des corrélations (voir Annexe 8) nous montre que ces indices sont les seuls à présenter un certain niveau de convergence avec les deux indices MERA et MERR.
SABA n'est que moyennement convergent, sur les quatre marchés.
Enfin, nous constatons encore le faible lien de MERA et MERR avec les autres indices puisqu'ils n'ont respectivement que 29% et 27% de leurs éléments significativement non-nuls.

Nous avons effectué les mêmes opérations en utilisant le coefficient de corrélation de Spearman. Les résultats figurent dans le tableau suivant.

TABLEAU 15 VALIDITES CONVERGENTES (*) (Coefficient de corrélation de Spearman)

	HUILE	BOISSON	PARFUM	SHAMP	SCORE TOTAL et %(*)
BRAA	1/13	1/14	1/13	-	3/40 7.5
BRAB	-	1/14	1/13	-	2/27 7.5
FRASA	6/13	6/14	9/13	7/8	28/4 58
FRASB	6/13	7/14	10/13	6/8	29/48 60
LEM	10/53	8/14	9/13	6/8	33/48 69
MERA	4/13	2/14	-	2/8	8/35 23
MERR	5/13	4/14	6/13	1/8	16/48 33
SABA	4/13	9/14	6/13	6/8	25/48 52
TOTT	9/13	6/14	11/13	-	26/40 65
TRTO	10/13	5/14	11/13	-	26/40 65
RAO	9/13	10/14	10/13	6/8	35/48 73
ELRO	7/13	10/14	11/13	-	28/40 70
TREL	7/13	6/14	11/13	-	24/40 60
IND	9/13	7/14	9/13	6/8	31/48 78
ACT	9/13	8/14	9/13	6/8	32/48 80

* Nombre d'éléments dans les diagonales de validité qui sont significativement non-nuls au seuil de risque alpha 5% et nombre total de ces éléments .
** Score total tous marchés confondus : c'est la somme des valeurs obtenues sur les quatre marchés. Figure à côté de ce score, le pourcentage d'éléments significatifs ramené au nombre total d'éléments.

L'examen de ce tableau montre, par rapport au cas précédent, que BRAA, BRAB, MERA et MERR restent à des niveaux de convergence très faibles. Pour les autres indices nous observons une homogénéisation de leur degré de convergence.

SS1-2 CRITERES DE VALIDITE DISCRIMINANTE

Toutes les conclusions que nous venons de porter ne sont que relatives à des conditions nécessaires de la validité de traits. Il nous faut maintenant examiner la validité discriminante. Rappelons que dans notre contexte de recherche, les conditions de validité discriminante formulées par C&F peuvent s'énoncer de la façon suivante:

Critère 1 : Toute corrélation entre deux indices sur le même marché (diagonale de validité : zone b) doit être plus forte que la corrélation entre ces deux mêmes indices sur deux marchés différents (triangles hétéro-traits et hétéro-méthodes, zone c).
Ceci équivaut à comparer chaque élément des diagonales de validité avec les valeurs, dans la même ligne et de la même colonne, des deux triangles hétéro-traits hétéro-méthodes adjacents à la diagonale.

Critère 2 : Toute corrélation entre deux mesures effectuées avec deux indices calculés sur le même marché T_i (diagonale de validité) doit être supérieure à toute corrélation des mesures de chacun de ces deux indices avec lui même, lorsqu'ils sont calculés sur deux marchés T_i et T_j différents.
Chaque corrélation dans une diagonale de validité (zone b) doit donc être supérieure aux corrélations correspondantes dans les deux triangles hétéro-traits mono-méthode correspondant aux deux indices (zone d).

Campbell et Fiske ont suggéré de compter le nombre de relations de ce type qui sont vérifiées et de le comparer au nombre de relations vérifiées attendues sous hypothèse du simple hasard. Cependant, les deux critères proposés par C&F, aussi intéressants soient-ils au plan conceptuel, sont difficiles à opérationnaliser dans la réalité. Notamment, deux problèmes se posent :

- comment les mettre en oeuvre sur une matrice de taille importante comme la nôtre ?
- si les contraintes de discrimination sont vérifiées mais que les niveaux des corrélations dans la diagonale de convergence ne sont pas significatifs, que doit on en conclure ?

Nous avons calculé, pour chaque indice, le nombre total de relations qu'il vérifie selon chacun des deux critères de validité convergente, rapporté au nombre total de relations pour chacun de ces deux critères. Ce comptage a été effectué uniquement dans le cas où l'élément correspondant dans la diagonale de validité était significativement non-nul. Les autres cas n'ont pas été pris en compte. Les résultats pour les deux critères (appelés VD1 et VD2) figurent dans le tableau suivant.

TABLEAU 16 SYNTHESE DES RESULTATS DE VALIDITE CONVERGENTE ET DISCRIMINANTE

	%VALID	%VD1	%VD2	%TVD1	%TVD2
FRASA	58	99	94	57	54
FRASB	65	99	91	64	59
LEM	71	100	99	71	70
MERA	29	80	94	23	27
MERR	27	79	90	21	24
SABA	48	99	90	47	43
TOTT	50	98	94	49	47
TRTO	96	92	62	60	65
RAO	69	100	99	69	68
ELRO	65	96	87	62	57
TREL	60	94	93	56	56
IND	67	98	91	66	61
ACT	71	97	93	69	66

Dans la colonne %VALID figurent les pourcentages d'éléments des diagonales de validité qui sont significativement non-nuls (obtenus à partir du coefficient de corrélation de Pearson). C'est la mesure du degré de convergence de l'indice que nous avons étudiée dans le paragraphe SS1-1.
Dans les colonnes %VD1 et %VD2, figure le pourcentage de relations vérifiées pour les critères VD1 et VD2 ceci, en ne prenant en compte que les cas où la diagonale de validité est significative.

Nous constatons que lorsque les éléments de la diagonale de validité convergente sont significativement non-nuls il n'y a pas de problème de validité discriminante.

Par contre, quand nous prenons en compte la totalité des éléments des diagonales de convergence, (même quand les valeurs correspondantes ne sont pas significatives), nous observons alors que les indices MERA et MERR tombent à des niveaux très bas, ceci étant dû à leur faible niveau de convergence avec les autres indices. Ces valeurs figurent dans les colonnes %TVD1 et %TVD2 .

Les résultats précédents suggèrent qu'une fois franchi le cap de la validité convergente, les indices ne rencontrent pas de problèmes de validité discriminante. Le problème de la convergence découle directement des problèmes de consistance interne qui ont été caractérisés dans la section précédente consacrée à l'étude de la fiabilité des indices.

Une fois écartés les indices BRAA, BRAB, MERA et MERR à cause de leur faible degré de convergence, nous obtenons un ensemble d'indices vérifiant assez bien les conditions de validité de traits .

C&F proposaient dans leur méthodologie originelle un troisième critère de validité discriminante suivant lequel la structure des corrélations dans les triangles hétéro-traits hétéro-méthodes (zones c) doit être identique d'un triangle à l'autre.

Pour étudier ce critère, il serait possible de calculer tous les coefficients de corrélation de rang (Spearman) entre les éléments appartenant à tous les triangles de type c. Afin de

ne pas trop alourdir les calculs et devant la difficulté d'interprétation de tous ces coefficients, nous n'avons pas étudié ce troisième critère.

Comme nous l'avons noté dans la section 1 consacrée à une approche théorique des problèmes de mesure, plusieurs chercheurs ont proposé des alternatives à l'analyse classique de la matrice MTMM. C'est ce que nous avons effectué dans le paragraphe suivant.

SS2 AUTRES APPROCHES DE LA VALIDITE DE TRAITS

De nombreuses techniques ont été proposées pour obtenir une analyse plus efficace de la matrice MTMM. Cet ensemble de techniques peut être partagé en deux courants:

- une approche essentiellement *"descriptive"* (exploratoire) qui utilise des méthodes statistiques à caractère *"exploratoire"*.
- une approche *"confirmatoire"* qui utilise des méthodes elles mêmes confirmatoires (ce que Fornell 1982 appelle les méthodes de la deuxième génération multivariée).

Nous utiliserons ici des méthodes issues de ces deux courants pour analyser la matrice MTMM. Notons cependant qu'une même méthode peut appartenir aux deux courants, l'appartenance d'une méthode à un courant étant surtout liée à l'état d'esprit avec lequel la méthode est utilisée.

SS2-1 APPROCHE EXPLORATOIRE

Parmi l'ensemble des techniques statistiques utilisées pour "ré-analyser la matrice MTMM", nous pouvons citer l'analyse de la variance (Stanley 1961), l'analyse des similarités et l'analyse typologique (Heeler et Ray 1972) ainsi que les méthodes factorielles qui ont été les plus utilisées (voir Schmitt 1977). Ceci s'explique pour plusieurs raisons :

- nous avons déjà souligné que d'un point de vue conceptuel, la théorie de la mesure (en particulier la théorie du Domaine d'Echantillonnage et la Théorie des Facteurs) est parfaitement compatible avec les fondements de l'analyse factorielle, l'une et l'autre ayant évolué en parallèle .
- le pouvoir de synthèse de l'analyse factorielle, la richesse de ses résultats et sa facilité d'interprétation font qu'elle constitue un outil privilégié pour les chercheurs en Sciences Sociales.

SS2-1-1 UTILISATION DE L'ANALYSE FACTORIELLE

Les critères de validité convergente peuvent être formulés de la façon suivante, avec l'analyse factorielle .

A Validité Convergente

Le carré du loading (saturation) d'un indice avec un facteur représente le pourcentage de variance de cet indice qui est expliqué par le facteur, c'est à dire la variance que cet indice partage avec le facteur. La saturation d'un indice avec un facteur peut donc être interprétée comme étant le degré de convergence de l'indice avec le concept correspondant au facteur. Guilford (1954 p 54) notait déjà ce point. Pour qu'il y ait une bonne validité convergente, il faut que les indices calculés sur un marché particulier pèsent fortement sur le même facteur.

B Validité Discriminante

Pour établir la validité discriminante, il faut observer que chaque indice ne pèse "que" sur le marché sur lequel il est calculé. Ceci peut être exprimé sous la forme de deux conditions :

C1 - Tout indice doit peser plus fort sur le facteur correspondant au marché sur lequel il a été calculé que sur n'importe quel autre facteur.

C2 - Tout indice mesuré sur un marché particulier doit peser plus fort sur le facteur correspondant à ce marché que tout autre indice qui n'a pas été mesuré sur ce marché.

Les deux conditions précédentes peuvent être étudiées dans deux types d'espaces des facteurs :

- soit un espace dans lequel les facteurs sont orthogonaux et c'est dans ce cas que les conditions de validité de traits sont les plus difficiles à remplir.
- soit un espace dans lequel les facteurs sont obliques. Les conditions de validité discriminante sont alors plus faciles à vérifier puisqu'il peut y avoir transfert de la variance entre deux indices qui ne sont pas calculés sur le même marché (signe d'une mauvaise validité discriminante) par l'intermédiaire de la variance entre les facteurs puisque ceux-ci peuvent être corrélés.

Nous avons appliqué la méthode sur la matrice des corrélations entre les 15 indices calculés sur les quatre marchés (ce qui représente au total une matrice à 52 lignes et colonnes). Les résultats figurent à l'Annexe 9.
71% de la variance est expliquée par les dix premiers facteurs. L'examen des valeurs propres indique que l'essentiel de la structure se retrouve dans les cinq premiers facteurs qui représentent 50% de la variance totale.

Après rotation orthogonale (critère Varimax) dans l'espace des dix premiers facteurs, nous obtenons les résultats suivants:

Sur le premier facteur, figurent des indices appartenant à deux marchés: HUILE et PARFUM .

F1 MERA MERR RAO ELRO calculés sur le marché HUILE, FRASA, LEM, MERA, TRTTO, TOTT, RAO, ELRO, TREL calculés sur le marché PARFUM

Il est assez difficile d'interpréter la nature exacte de ce facteur. Il capte à la fois un effet de méthode (les indices) mais aussi un effet de marché. Un tel résultat ne caractérise pas un fort degré de validité de traits concernant l'ensemble des indices.

Les quatre facteurs suivants sont par contre nettement caractéristiques des quatres marchés, montrant que si la validité de traits n'est pas parfaite, elle existe tout de même.

Par ordre d'importance décroissante sur chaque facteur, nous trouvons:

F2 LEM, IND, ACT, FRASA, FRASB, RAO, SABA, ELRO calculés sur le marché BOISSON

F3 LEM, IND, ACT, RAO, FRASB, FRASA, SABA calculés sur le marché SHAMP

F4 LEM, ACT, IND, FRASB, RAO, FRASA, TOTT, TRTO, ELRO, TREL calculés sur le marché PARFUM

F5 TOT TRTO ELRO TREL LEM calculés sur le marché HUILE mais aussi MERA MERR calculés sur BOISSON qui pèsent fortement sur ce facteur

.

Nous pouvons constater que BRAA, BRAB, MERR et MERA ne pèsent sur aucun des facteurs correspondant aux marchés sur lesquels ils ont été calculés. Par contre, LEM apparait sur chacun des quatre marchés comme étant central à la mesure. Les indices FRASA et FRASB apparaissent eux aussi comme centraux dans la mesure, excepté sur le marché HUILE .

Après ces cinq premiers facteurs, les facteurs suivant apparaissent plutôt comme étant des facteurs qui caractérisent un "effet de méthode". Sur ces facteurs pèsent en effet les mêmes types d'indices calculés sur des marchés différents. Par exemple, nous citerons:

F6 LEM ACT IND calculés sur HUILE et IND ACT calculés sur SHAMP
F8 FRASA FRAB calculés sur les quatre marchés ainsi que BRAA et BRAB calculés sur BOISSON .

La variance de tels facteurs peut être assimilée à de l'effet de méthode.

Comme dans la section précédente, nous avons effectué des rotations obliques. Comme prévu, les conditions de validité de traits sont alors mieux remplies (voir Annexe 9). La structure des facteurs après rotation est la suivante :

F1 MERA, MERR, RAO, ELRO calculés sur HUILE. TRTTO, TOTT, ELRO,TREL calculés sur PARFUM
F2 LEM, IND, ACT, FRASA, FRASB, RAO, ELRO, SABA calculés sur BOISSON
F3 LEM, IND, ACT, RAO, FRASB, FRASA, SABA calculés sur SHAMP
F4 LEM, ACT, IND, FRASB, RAO, FRASA calculés sur PARFUM
F5 TOTT, TRTO, ELRO, TREL calculés sur HUILE mais aussi MERA MERR calculés sur BOISSON pèsent fortement sur ce facteur.
F6 FRASA, FRASB, IND, ACT, LEM calculés sur HUILE.

L'examen des corrélations entre les facteurs montre que les facteurs restent bien orthogonaux (indépendants) excepté pour les facteurs F1 et F4 qui présentent une corrélation de .31. Cette corrélation ne nuit pas à la validité discriminante puisque sur ces deux facteurs pèsent des indices calculés sur le même marché.

Nous avons effectué la même analyse en utilisant le coefficient de corrélation de Spearman. Les résultats figurent à l'Annexe 10. Comme nous l'avions noté dans l'analyse précédente, les résultats ne sont pas modifiés dans leur forme générale. Nous retiendrons de cette analyse les éléments suivants :

- les dix premiers facteurs expliquent maintenant 78% de la variance (52% pour les quatre premiers). La structure commune est donc légèrement plus forte qu'avec le coefficient de Pearson.
- les indices BRAA, BRAB, MERA, MERR, présentent toujours un mauvais degré de convergence avec les autres.
- pour les autres indices, ,nous constatons une amélioration du respect des conditions de validité de traits définies plus haut.
- l'indice SABA voit sa convergence évoluer très favorablement sur deux marchés, BOISSON et SHAMP, grâce à l'utilisation du coefficient de Spearman.

SS2-1-2 UTILISATION JOINTE DE L'ANALYSE TYPOLOGIQUE ET DE L'ANALYSE DES SIMILARITES

L'utilisation de ces méthodes a déjà été proposée dans la littérature Marketing afin d'étudier la structure d'une matrice MTMM. Heeler et Ray (1972) ont proposé l'utilisation jointe de l'analyse typologique hiérarchique et de l'analyse non-métrique des similarités (MDS). Plus récemment, Fornell et Denison (dans Fornell 1982) ont proposé une approche assez formalisée de l'étude de la validité des concepts utilisant le MDS. Les auteurs justifient l'utilisation d'une telle méthode par les raisons suivantes :

- elle nécessite très peu d'hypothèses sur les formes fonctionnelles qui relient les items aux facteurs (cette forme est par hypothèse linéaire dans l'analyse en composantes principales).

- elle nécessite peu d'hypothèses sur la nature des données utilisées.

Avec une telle analyse, la relation entre deux indices est interprétée en termes de distances. Fornell et Denison ont présenté les critères formels suivants pour apprécier la validité de traits :

- a chaque concept i, correspond une région Ri dans l'espace qui lui est spécifique.
- chaque variable censée mesurer le concept i doit être plus proche des autres variables censées mesurer i que n'importe quelle autre variable ne mesurant pas i.

Nous avons utilisé ici l'analyse typologique hiérarchique (Johnson 1967) avec, comme mesure de similarité entre deux groupes, la distance moyenne.

Pour effectuer l'analyse des similarités, nous ne disposions pas comme Fornell et Denison (F&D) d'un modèle de MDS *"contraint"*. Nous nous limiterons donc à une analyse exploratoire (par opposition à confirmatoire). Nous avons utilisé le modèle ALSCAL, disponible dans la bibliothèque SAS, dans sa forme non métrique. Les analyses ont été mises en oeuvre dans les conditions suivantes :

- en utilisant le coefficient de corrélation de Pearson ou de Spearman.
- à partir de la matrice complète des corrélations (nous avons alors 52 objets: les indices calculés sur les quatre marchés) mais aussi sur la matrice réduite, après suppression des indices BRAA BRAB MERA MERR qui se sont révélés nuire à la validité de traits dans les analyses précédentes (nous avons alors 40 objets).

Quatre analyses ont été menées. Les éléments permettant d'apprécier la qualité des analyses de similarité figurent dans le tableau 17 suivant qui fournit, dans chaque case, successivement , les valeurs du stress et le carré du coefficient de corrélation multiple (R2), pour chacune des analyses .

Le stress (S) est celui qui est obtenu avec la formule 1 de Kruskall (voir Kruskall dans Davies et Coxon 1981, p 59). Il s'interprète comme un pourcentage de relations monotones non satisfaites entre les distances sur la représentation graphique et les disparités (les disparités sont une fonction monotone des dissimilarités figurant dans la matrice analysée). Ici, les corrélations (indices de similarité) ont été transformées en dissimilarité par la transformation :

- $DISS_{ij} = 1 - ACORR_{ij}$

- $DISS_{ij}$ est la dissimilarité entre les indices i et j
- $ACORR_{ij}$ est le coefficient de corrélation entre les indices i et j
- R2 est la proportion de la variance des disparités qui est expliquée par le modèle (voir Schiffman et al 1981 p 175).

TABLEAU 17 STRESS DES REPRESENTATIONS POUR LES 4 CAS ETUDIES

Nombre de dim-ensions	52 Objets Pearson S R2	52 Objets Spearman S R2	40 Objets Pearson S R2	40 Objets Spearman S R2
6	.12 .76	.11 .85	.09 .88	.06 .95
5	.15 .70	.13 .80	.12 .83	.09 .93
4	.18 .62	.17 .72	.15 .78	.12 .87
3	.24 .51	.23 .61	.19 .70	.16 .80

Nous observons que:

- la suppression des indices peu liés aux autres améliore nettement la qualité de l'ajustement (ceci est bien sûr attribuable en partie à la diminution du nombre d'objets, mais cet effet ne doit pas être relativement trop fort vu le nombre important d'indices (objets) qui sont analysés).
- l'utilisation de l'indice de Spearman améliore systématiquement la qualité de l'ajustement.

Afin de ne pas trop alourdir nos résultats, seul le dernier cas (40 variables, coefficient de Spearman) sera présenté ici. Les autres analyses figurent dans l'annexe 11.

L'analyse jointe des résultats de la typologie et des cartes obtenues avec le MDS permet de constater que la validité est globalement bien atteinte. Les quatre marchés se détachent nettement les uns des autres, notamment dans l'analyse sur la matrice réduite.

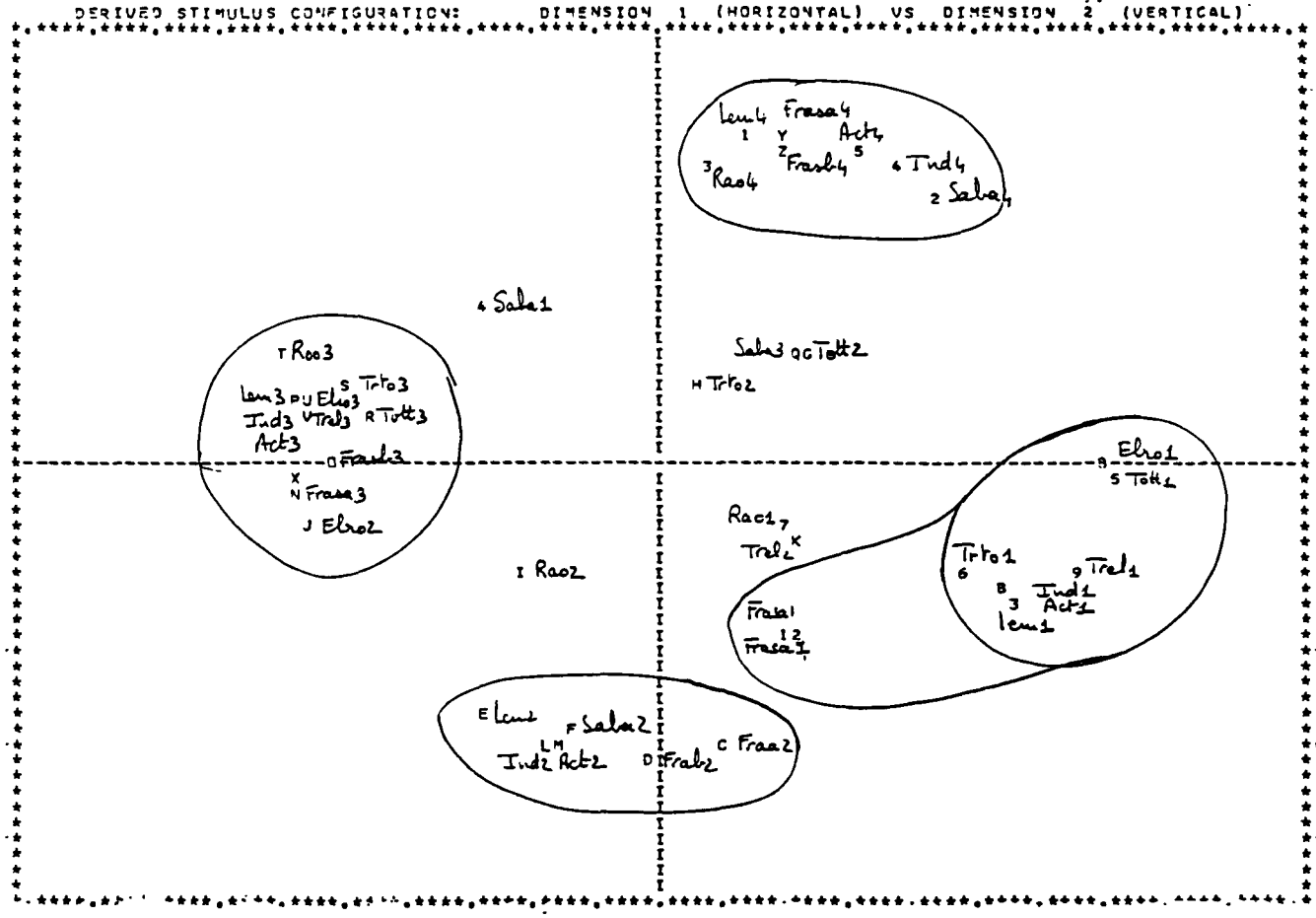
Une telle approche est intéressante, surtout (comme le notent Fornell et Denison) par son aspect complémentaire aux méthodes factorielles.

Le graphique suivant est la projection des indices sur le plan engendré par les deux premiers axes de la représentation obtenue en trois dimensions avec le MDS. Chaque indice est affecté d'un code correspondant au marché sur lequel il a été calculé : 1 pour HUILE, 2 pour BOISSON, 3 pour PARFUM et 4 pour SHAMP. Par exemple SABA3 fait référence à l'indice SABA calculé sur le marché PARFUM. Sur le graphique 2, nous *"visualisons"* bien les problèmes de validité convergente et discriminante. Notamment, nous observons que les indices RAO SABA TOTT et TREL rencontrent des problèmes de convergence sur plusieurs marchés.

Le degré de convergence des indices mesurant le même concept peut être apprécié visuellement par la concentration des points qui le représentent. C'est ainsi que les marchés PARFUM et SHAMP présentent une meilleure validité de traits que les autres. Ils remplissent bien les conditions posées par Fornell et Denison .

La faiblesse de cette approche est que lorsque la position d'un indice n'est pas bonne en terme de validité de traits, nous ne savons pas si cela est le fait de la mauvaise validité de l'indice ou d'une mauvaise estimation de la position de ce point par le modèle.

GRAPHIQUE 2 VISUALISATION DE LA VALIDITE CONVERGENTE A L'AIDE DE L'ANALYSE DES SIMILARITES



En conclusion, nous soulignerons qu'une telle approche doit plutôt être utilisée en complément des analyses factorielles précédentes. Elle n'a pas, en effet, les atouts de clarté, de facilité d'interprétation et de rigueur de l'analyse factorielle. Surtout, elle ne présente pas de cohérence avec la théorie de la mesure, ce qui est le cas pour l'analyse factorielle.

Elle offre par contre une alternative intéressante, notamment par son absence d'hypothèses relativement à la qualité de la mesure et au type de lien des indices avec les concepts qu'ils mesurent (la relation n'a pas besoin d'être linéaire). Son interprétation en termes de distances est très naturelle.

Après avoir étudié la matrice MTMM à partir de méthodes exploratoires, nous allons présenter maintenant une approche confirmatoire .

SS2-2 APPROCHE CONFIRMATOIRE DE LA VALIDITE DE TRAITS

L'utilisation de méthodes de type confirmatoire n'est pas nouvelle pour l'analyse de la matrice MTMM. Joreskog (1970) proposait déjà l'utilisation de l'analyse factorielle confirmatoire et du Path-Analysis pour étudier de telles matrices. La méthode la plus séduisante est le modèle d'équations simultanées à variables latentes et erreur de mesure dont LISREL est une des versions la plus connue (voir Joreskog et Sorböm 1982). Les exemples d'utilisation de ce modèle pour valider les mesures prouvent que la méthode est de plus en plus utilisée en Marketing (voir Anderson 1985, Anderson, Gerbing et Hunter 1987 et Gerbing et Anderson 1988).

SS2-2-1 ANALYSE FACTORIELLE CONFIRMATOIRE

La méthode confirmatoire la plus simple consiste à *"forcer"* les rotations, dans une analyse factorielle, en introduisant une matrice cible. Au lieu d'utiliser un critère de rotation arbitraire (comme le critère Varimax) qui se fonde uniquement sur les données analysées (maximiser la variance des valeurs des loading sur les facteurs), une part de jugement est introduite dans la méthode. Les conditions idéales de validité de traits vont être transcrites sous la forme d'une matrice cible pour la rotation. La rotation s'effectue alors de manière à ce que la structure des facteurs obtenue soit la plus proche possible de celle qui figure sur la matrice cible.

Dans la matrice cible nous allons imposer :

- que le lien d'un indice soit maximum avec le facteur représentant le marché sur lequel il a été calculé (loading égal à 1)
- que le lien d'un indice avec les facteurs correspondant aux marchés sur lesquels il n'a pas été calculé soit minimum (loading égal à zéro).

L'estimation par le maximum de vraisemblance est la mieux adaptée pour une telle approche (voir Harman 1976 chap 10, Heeler et al 1977). Cette procédure permet notamment d'effectuer des tests statistiques sur le nombre de facteurs communs significatifs que l'on peut extraire. D'autre part, avec cette méthode d'estimation, nous tenons compte du fait que la matrice des corrélations analysée n'est pas exempte de fluctuations dues au fait que les données sont collectées sur un échantillon et non sur la population toute entière (voir Babakus, Fergusson et Joreskog 1987 pour une étude à ce sujet). Ici, l'échantillon est constitué par les panélistes de SCAN 5000, nous sommes donc confrontés à ce problème.

Malgré tous ces avantages, nous n'avons malheureusement pas pu utiliser la méthode (disponible dans la bibliothèque SAS), la matrice des corrélations à analyser étant singulière, l'utilisation de la procédure d'estimation par le maximum de vraisemblance devient impossible. Nous avons par contre utilisé la méthode ULS (Unweighted Least Squares) qu'il est conseillé d'utiliser dans de tels cas (voir Joreskog et Sorbom 1986 à ce sujet).

La matrice cible utilisée contraignait la rotation à se faire dans un espace à quatre dimensions (une dimension pour chaque marché). D'autre part, pour atteindre la cible, les facteurs pouvaient devenir obliques.

Nous pouvons reformuler ici les conditions de validité que nous avions développées précédemment à savoir :

C1- Chaque indice doit peser plus fort sur le facteur correspondant au marché sur lequel il a été calculé
C2- Tout indice doit peser moins sur les facteurs correspondant à des marchés sur lesquels il n'a pas été calculé que les indices qui ont été effectivement calculés sur ces marchés.
C3- L'obtention des conditions précédentes ne doit pas se faire au prix de corrélations trop fortes entre les facteurs.

Comme précédemment, nous avons effectué quatre types d'analyses: sur la matrice complète (52 variables) puis simplifiée (40 variables) en utilisant à chaque fois les deux types de coefficients de corrélation (Pearson et Spearman).

ANALYSE 1 52 VARIABLES - COEFFICIENT DE PEARSON

Les résultats figurent à l'Annexe 12 A. La rotation a été effectuée dans un espace à quatre dimensions qui explique 44% de la variance totale (ce qui est remarquable pour un aussi grand nombre de variables). Nous constatons que pour la grande majorité des indices, les conditions de validité sont bien remplies .

BRAA et BRAB présentent toujours les mêmes problèmes de validité convergente, ils ne pèsent significativement, malgré la rotation forcée, sur aucun des marchés sur lesquels ils ont été calculés.

Il en est de même pour MERA et MERR qui eux ne pèsent pas sur leur marché mais sont par contre liés à d'autre marchés, venant ainsi nuire à la validité discriminante globale. Le problème lié à ces indices est manifestement un problème de convergence et non de discrimination.

L'indice SABA rencontre plus des problèmes de validité convergente que de validité discriminante. Il n'est pas très fortement lié aux facteurs correspondant aux marchés sur lesquels il est calculé, mais ces liens sont tout de même plus forts qu'avec les facteurs correspondant aux marchés sur lesquels il n'a pas été calculé; le critère C1 précédent est donc vérifié avec cet indice. Manifestement, cet indice appartient au même domaine que les autres, mais il est affecté d'un fort niveau d'erreur de mesure. C'est ce que nous avions noté à la section précédente relative à l'étude de la fiabilité.

L'examen de la matrice des corrélations entre les facteurs nous montre que cette structure a été obtenue sans que pour autant les facteurs ne deviennent obliques, exceptés les facteurs F1 et F4, avec une corrélation de - 0.33.

ANALYSE 2 52 VARIABLES - COEFFICIENT DE SPEARMAN

La même analyse a été effectuée sur la base du coefficient de corrélation de Spearman (voir Annexe 12B).

Par rapport à l'analyse précédente, nous retiendrons les éléments suivants :

- le niveau général des saturations des indices sur leur facteur est maintenant plus fort. La validité convergente des indices s'est améliorée (notamment pour l'indice SABA calculé sur les marchés BOISSON et PARFUM).

- cette amélioration n'a par contre pas bénéficié aux indices qui présentaient déjà un faible niveau de validité convergente, c'est à dire les indices BRAA, BRAB, MERA et MERR.
- le niveau de validité discriminante a lui aussi évolué favorablement .
- l'examen des corrélations entre les facteurs (voir Annexe 12B) montre que cette amélioration a pu se faire en maintenant une bonne orthogonalité entre les facteurs. Notamment, F1 et F4 sont moins corrélés (.20) que dans l'analyse précédente.

BRAA, BRAB, MERA et MERR s'étant révélés appartenir à un autre domaine, nous avons recommencé les mêmes analyses après les avoir supprimés.

ANALYSE 3 40 VARIABLES - COEFFICIENT DE PEARSON

Les résultats de cette analyse figurent à l'Annexe 12C. Les quatre premiers facteurs représentent maintenant 52 % de la variance totale. Le graphique représentant les pourcentages de variance expliquée par les facteurs montre que ces quatre facteurs se détachent nettement des autres, ils représentent bien la structure interne commune des corrélations. Dans cet espace à quatre facteurs, les rotations nous apportent les éléments suivants concernant la validité des indices.

Sur le marché HUILE (premier facteur, voir annexe 12C), seuls les indice SABA et RAO posent des problèmes de convergence, avec des corrélations de .19 et .22 seulement. D'autre part, les indices TRTO et TREL calculés sur le marché BOISSON, les indices LEM, IND et ACT calculés sur le marché SHAMP, sont significativement corrélés avec ce facteur. Tous les indices calculés sur le marché Boisson sont fortement corrélés avec le facteur 2. Nous pouvons remarquer que les indices IND et ACT calculés sur le marché parfum sont fortement corrélés avec ce facteur, ce qui peut être considéré comme un signe de faiblesse de la validité discriminante. Pour le marché PARFUM, (facteur 3), seul l'indice SABA pose des problèmes de convergence (corrélation de .10 seulement). Enfin, sur le marché SHAMP (facteur 4), les conditions de validité de traits sont bien remplies.

Ainsi, globalement, les conditions de validité discriminante sont correctement remplies. Aucun indice ne rencontre systématiquement de problèmes de validité de traits.

Sur l'annexe 12C, les quatre facteurs apparaissent comme étant orthogonaux, la rotation "forcée" ne s'est pas faite au détriment de l'indépendance des traits (les facteurs), ce qui est un bon signe de validité de traits.

ANALYSE 4 40 VARIABLES - COEFFICIENT DE SPEARMAN

Avec le coefficient de Spearman, les quatre premiers facteurs expliquent 62% de la variance totale (voir annexe 12D). La structure des corrélations est donc plus forte que précédemment.

Par rapport au cas précédent, la validité de traits s'est améliorée. Seul l'indice SABA, calculé sur le marché PARFUM, présente encore un faible niveau de corrélation.

Nous pouvons en conclure sur la *"bonne validité de traits"* des indices étudiés, une fois que les indices BRAA, BRAB, MERA et MERR ont été supprimés. Seul l'indice SABA qui est caractérisé par une fiabilité médiocre, rencontre des problèmes de convergence sur certains marchés. Les indices de Lehmann, de Rao et Sabavala (en version non pondérée), l'indice de Fraser et Bradford calculé à partir de périodes inter-

achats (une fois que l'univers concurrentiel est pris en compte lors du calcul des périodes inter-achats), les indices Défense et Attaque (qui sont des indices de dissimilarité indirecte) mesurent ainsi le même concept et ne mesurent que lui.
Par contre, l'indice proposé par Merunka et Bourgeat d'une part et par Fraser et Bradford d'autre part (dans sa forme originelle, sans prise en compte de l'univers de concurrence pour le calcul des intervalles inter-achats), mesurent chacun *"un autre concept"* que celui qui est mesuré par les indices précédents. De plus, il ne s'agit pas d'une autre facette du même concept, puisque ces indices sont les mesures de concepts qui sont orthogonaux entre eux. Nous retiendrons donc les résultats suivants:

- les indices construits à partir de transferts entre marques et les indices construits à partir de périodes inter-achats (une fois que l'univers de concurrence est pris en compte) sont les mesures du même concept. Cependant, les indices construits à partir des intervalles inter-achats (FRASA et FRASB), de même que les indices RAO et SABA présentent des niveaux de fiabilité moins élevés que l'indice de Lehmann. Par sa simplicité de calcul et sa fiabilité cet indice reste donc une mesure intéressante de la concurrence (selon les critères de validité étudiés ici).
- les indices que nous avons proposés (Défense et Attaque), qui sont des mesures "indirectes" de dissimilarité et prennent en compte la non-symétrie, convergent bien avec les indices de concurrence construits à partir de transferts directs. Comme l'ont montré Totten et Elrod (1988), ce type d'indices est moins sensible que les autres à une définition inexacte du champ concurrentiel. Cependant, leur fiabilité est moins bonne que celle de l'indice de Lehmann. Ils peuvent donc, dans certains cas (notamment quand le chercheur n'est pas certain de l'exactitude de la définition du champ concurrentiel qu'il s'est donnée), être utilisés avec profit.

La section suivante, consacrée à l'étude de la validité externe, permettra d'apprécier ces indices selon un autre point de vue, qui est celui de l'utilisation pratique des indices.
Auparavant, nous présenterons, pour finir cette section, une utilisation du modèle d'équations structurelles appliqué à l'étude de la validité.

SS2-2-2 UTILISATION DU MODELE LISREL POUR ETABLIR LA VALIDITE DE TRAITS DES INDICES DE MESURE DE LA CONCURRENCE

Comme nous l'avons noté, le modèle Lisrel est souvent utilisé pour étudier la validité de traits d'une échelle de mesure multiple, nous avons cité plus haut plusieurs exemples d'application. Nous proposons ici une méthodologie très simple d'utilisation du modèle pour "tenter" d'établir la validité de traits d'une échelle de mesure.

MISE EN FORME DU PROBLEME

Nous avons noté, précédemment, que la matrice des corrélations analysée (à 52 ou à 40 indices) n'est pas définie positive, ce qui ne permet pas d'utiliser la méthode d'estimation par le maximum de vraisemblance, méthode qui est pourtant la seule nous permettant d'effectuer des tests statistiques. Afin d'obtenir une matrice définie positive, nous n'avons, pour chaque type d'indice, conservé qu'une seule version. Ceci nous a amené à supprimer les indices FRASA, SABA, TOTT, TRTO, TREL, IND et ACT. Les quatre indices conservés sont donc FRASB, LEM, RAO, ELRO. D'autre part, afin de ne

pas avoir trop de paramètres à estimer dans le modèle (nous avons pu constater qu'un trop grand nombre de paramètres à estimer aboutit à des estimations de mauvaise qualité), nous n'avons conservé que trois marchés: BOISSON, PARFUM et SHAMP. La matrice des corrélations (coefficient de Spearman) est donc une matrice 15 x 15.

Comme dans les analyses précédentes, chaque facteur correspond à un des marchés sur lequel les indices ont été calculés. Ces facteurs (au nombre de trois) peuvent être contraints à rester orthogonaux ou bien laissés libres de devenir obliques.

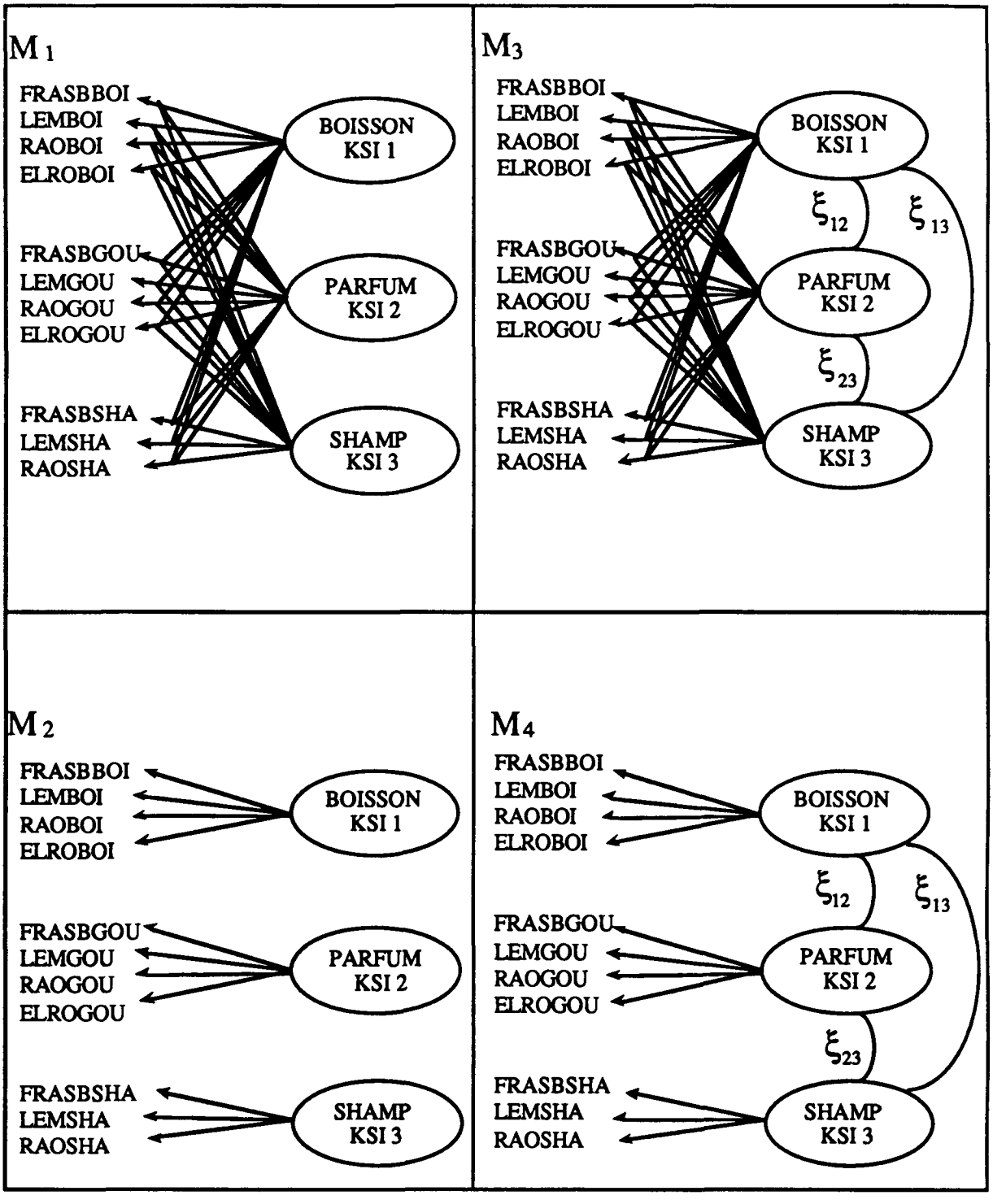
D'autre part, le lien de chaque indice avec chaque facteur (le loading) peut être laissé libre ou être contraint à vérifier "parfaitement" les conditions de validité de traits (chaque indice a alors un loading non-nul sur le seul facteur correspondant au marché sur lequel il a été calculé et nul ailleurs).

Nous avons estimé les quatre modèles qui sont présentés dans le tableau suivant:

		FACTEURS	
		ORTHOGONAUX	OBLIQUES
LOADING	LIBRES	M1	M3
	CONTRAINTS	M2	M4

Les modèles M1 à M4 peuvent être représentés de la manière suivante, en utilisant les conventions classiques en matière de modèles d'équations structurelles (voir Bagozzi 1980 chap 3).

FIGURE 9 REPRESENTATION SCHEMATIQUE DES QUATRE MODELES TESTES



Sur le schéma précédent, chaque indice a pour suffixe les trois lettres caractérisant le marché sur lequel il a été calculé : HUI pour HUILE, GOU pour PARFUM et SHA pour SHAMP. Les résultats de l'estimation de ces quatre modèles figurent à l'annexe 14.

ANALYSE DE LA VALIDITE DE TRAITS

La qualité de l'ajustement relatif à chacun des quatre modèles figure dans le tableau 18 suivant. Figurent en colonnes le nombre de degrés de liberté du modèle (DDL), la valeur du logarithme de la vraisemblance (CHI2), plus cette valeur est élevée et moins la qualité de l'ajustement est bonne, pour un degré de liberté donné. Cette valeur suit la loi du chi2 avec un nombre de degré de liberté égal à DDL. Nous trouvons en outre les valeurs du "goodness of fit" (GFI) et de l'erreur moyenne (RMR). Nous renvoyons le lecteur au manuel d'utilisation de LIREL VI pour une description de ces indices d'appréciation de la qualité globale de l'ajustement.

A la valeur CHI2, pour chaque modèle, correspond en théorie la probabilité P d'observer une valeur plus forte que la valeur CHI2 sous hypothèse H0 que le modèle n'est pas contraint. Plus cette valeur est faible et plus le modèle testé pose une contrainte forte sur les données supposées supporter ce modèle. Une valeur proche de zéro permet de rejeter l'hypothèse H0, ce qui signifie que le modèle testé n'est pas compatible avec les données empiriques utilisées. En fait, ceci n'est, en pratique, pas applicable pour trois raisons (voir Joreskog et al 1981): l'analyse doit être basée sur la matrice des covariances, toutes les variables analysées doivent suivre une distribution normale multivariée, la taille de l'échantillon doit être très élevée. Ces trois conditions ne sont pas remplies ici, ce qui nous amène à effectuer des comparaisons relatives d'ajustements entre des modèles *"emboîtés"* (modèles dont l'un est le cas particulier de l'autre).

TABLEAU 18 RESULTATS RELATIFS A L'AJUSTEMENT DES 4 MODELES ETUDIES

MODELE	DDL	CHI2	GFI	RMR
M1	19	82	.897	.049
M2	44	135	.819	.197
M3	16	82	.897	.049
M4	41	121	.837	.123

Le modèle le plus général est le modèle M3 (loading libres et facteurs obliques). Le modèle M1 (libre et orthogonal) est un cas particulier de M3, de même le modèle M4 (libre et oblique) est un cas particulier de M3. Enfin, le modèle M2 (contraint et orthogonal) est un cas particulier du modèle M1. Pour deux modèles i et j, la différence (CHI2i-CHI2j) suit une distribution du chi deux dont le nombre de degrés de liberté est égal à (DDLi-DDLj). De cette manière, il est possible de tester la contrainte imposée aux données lorsque l'on passe d'une forme de modèle à une autre, cette dernière étant un cas particulier de la première.

Les modèles M1 ou M3 représentent le cas le plus général où aucune contrainte de validité de traits n'est fixée.

Le modèle M2 représente le cas le plus particulier qui correspond à une hypothèse de validité de traits *"parfaite"*. Chaque indice ne peut peser que sur le seul facteur (traits) correspondant au marché sur lequel il a été calculé et les facteurs sont orthogonaux.

Le modèle M4 représente un cas intermédiaire dans lequel chaque indice ne pèse que sur le seul facteur correspondant au marché sur lequel il a été calculé, mais les facteurs

peuvent ne pas être orthogonaux, ce qui facilite l'obtention des contraintes de validité de traits sur chacun des facteurs.

Il est donc possible d'évaluer la qualité de la validité de traits à l'aide des deux tests suivants :

Test 1 : CHI2(M2) - CHI2(M1)
Test 2 : CHI2(M2) - CHI2(M3)

Le test 1 permet d'apprécier la force de la contrainte qui est imposée aux données quand nous passons du modèle libre (M1) au modèle vérifiant la validité de traits (M2). C'est le test le plus dur, puisqu'il prend en compte simultanément les validités convergente et discriminante.
Le test 2 permet d'apprécier dans quelle mesure les facteurs ont tendance à devenir obliques quand la contrainte d'orthogonalité est relâchée. Si l'orthogonalité est une contrainte significativement forte, nous pouvons en conclure à des problèmes de validité discriminante.

RESULTATS

Les résultats figurent respectivement pour les quatre modèles (M1 à M3) dans les annexes 14A à 14D. Nous n'avons fait figurer dans cette annexe que les tableaux des corrélations des indices avec les facteurs (matrice appelée Lambda X) et la matrice des corrélations entre les facteurs, lorsque le modèle est oblique (modèles 3 et 4), avec les autres modèles cette matrice étant égale à la matrice identité. Les valeurs figurant dans ces matrices sont les estimateurs du modèle par le maximum de vraisemblance, dans le cas d'une solution standardisée (la variance de chaque facteur est ramenée à 1), solution qui peut s'interpréter comme une analyse factorielle classique.
De plus, pour le modèle M2 dans lequel les loading sont contraints et les facteurs sont orthogonaux, nous présentons (voir plus bas) le tableau des coefficients de corrélation multiple de chaque variable (indice) avec les facteurs (ici avec le seul facteur correspondant au marché sur lequel l'indice a été calculé puisque la solution est contrainte). Ce coefficient est une mesure de la force avec laquelle l'indice *"capte"* le concept étudié, ici la structure de la concurrence sur chacun des marchés. Ce coefficient, pour l'indice i, a pour valeur:

$$Ci = (1 - (\gamma i / sii))$$

γi est la variance de l'erreur associée à l'indice i
sii est la variance empirique de l'indice i.

Le coefficient Ci est donc une mesure de la *"fiabilité"* de la mesure du concept (structure de la concurrence) à l'aide de l'indice, sur le marché considéré. Cette mesure représente la corrélation de l'indice avec le concept non observable, il s'agit de la corrélation de l'indice avec le score *"vrai"*, au sens de la théorie du *"domaine d'échantillonnage"*. Plus cette valeur est proche de 1 et plus la fiabilité de la mesure est bonne pour l'indice i.

Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-J

TABLEAU 19 RESULTATS CONCERNANT LA FIABILITE DES INDICES

FIABILITE	Ci	MARCHE
FRASBBOI	.268	
LEMBOI	1.00	BOISSON
RAOBOI	.406	
ELROBOI	.300	
FRASBBGOU	.529	
LEMGOU	1.06	PARFUM
RAOGOU	.869	
ELROGOU	.409	
FRABSHA	.767	SHAMP
LEMSHA	.380	
RAOSHA	.095	

Nous pouvons constater que ces valeurs sont "correctes", pour l'ensemble des indices, sur l'ensemble des marchés, excepté l'indice RAO calculé sur le marché SHAMP. Sur les deux premiers marchés, l'indice LEM (comme nous l'a déjà vu par ailleurs) est la mesure la plus "centrale" du concept et présente une fiabilité maximum.

RESULATS RELATIFS A LA VALIDITE DE TRAITS

L'examen de la matrice des loading dans le modèle M1 (loading libres et facteurs orthogonaux) permet de constater que les conditions générales de validité de traits sont bien remplies:

- chaque indice est bien corrélé sur le facteur correspondant au marché sur lequel il a été calculé.
- chaque indice pèse plus sur son facteur que sur n'importe quel autre
- chaque indice pèse moins sur un facteur correspondant à un marché sur lequel il n'a pas été calculé que n'importe quel indice calculé sur ce marché.

Nous allons maintenant utiliser les tests présentés plus haut pour vérifier de manière plus "rigoureuse" ce qui est constaté visuellement sur la matrice des loading. Dans le tableau 18 précédent, figurent les caractéristiques de l'ajustement des quatre modèles qui sont étudiés. A partir de la valeur de GFI (une valeur proche de 1 signifie que l'ajustement statistique est excellent), nous constatons que tous les modèles présentent une bonne qualité globale d'ajustement.

TEST 1

CHI2(M2)-CHI2(M1) suit un CHI2 à (44-19)=25 degrés de liberté. La valeur observée est égale à (135-82)=53. Joreskog et al (1981) proposent d'accepter l'hypothèse H0 :*"il n'y a pas de différence significative entre les niveaux d'ajustement des deux modèles M2 et M1"* lorsque la valeur CHI2(M1)-CHI2(M2) est comparable à la différence de degré de liberté entre les deux modèles (ici 25).

Nous pouvons observer que l'augmentation du CHI2 lorsque l'on passe du modèle libre (M1) au modèle contraint (M2) n'est pas très importante relativement à l'augmentation du nombre de degrés de liberté: elle est de 52 par rapport à 25.

L'observation de la distribution théorique du CHI2 à 25 degrés de liberté indique que la valeur théorique pour le CHI2 à 25 ddl, pour un niveau de risque alpha de 1% est égale à 44.3. Cette valeur est inférieure à la valeur observée, mais elle en est très proche ce qui nous conduit à accepter l'hypothèse H0. Nous en conclurons que le modèle M2 (contraint) présente une qualité d'ajustement qui n'est que légèrement inférieure à celle du modèle libre. Les données s'ajustent bien aux contraintes posées par le modèle M2, ce qui signifie que la validité de traits est *"bien"* remplie (M1 correspond, rappelons le, au cas où la validité de traits est parfaite).

TEST 2

La même démarche est utilisée pour tester la différence d'ajustement entre M2 (modèle contraint orthogonal) et M4 (modèle contraint oblique) afin d'étudier la validité discriminante.

CHI2(M2)-CHI2(M4) suit un CHI2 à (44-41)=3 degrés de liberté. CHI2(M2)-CHI2(M4)=(135-121)=14, est la valeur observée du CHI2 à 3 degrés de liberté. La valeur théorique, pour 3 degrés de liberté et un seuil alpha de 1% est égale à 11.3, valeur qui est inférieure à la valeur 14 observée, mais elle en est peu différente. Nous accepterons encore, sans prendre trop de risques, l'hypothèse H0: "on n'améliore pas la qualité de l'ajustement en permettant aux facteurs de devenir obliques". La validité discriminante est donc confirmée par ce test.

La méthodologie que nous venons de proposer permet de tester facilement et de manière rigoureuse la validité de traits d'une mesure multiple, une fois que les résultats sont présentés sous la forme d'une matrice MTMM. Sa limite essentielle réside dans le fait que la matrice ne doit pas être de taille trop importante, ceci afin de limiter le nombre de paramètres à estimer et de limiter les chances que la matrice devienne singulière (non-inversible). Ainsi, dans notre exemple, nous avons du ramener la taille de cette matrice de 52 à 15, pour pouvoir l'analyser.

SECTION 4 ETUDE DE LA VALIDITE EXTERNE

Une fois étudiée la validité interne de la mesure, il reste à la sortir de sa "logique de mesure" et à la confronter aux théories en vigueur, aux faits et aux points de vue du spécialiste, c'est ce que nous avons appelé "validité externe".

SS1 PRESENTATION DU PROBLEME

Nous l'avons déjà noté, trois voies s'offrent à nous pour étudier la validité externe des mesures. Par ordre croissant selon ce caractère externe, nous trouvons:

SS1-1 LA VALIDITE NOMOLOGIQUE

Elle se définit comme "*la capacité de l'indice à s'inscrire correctement dans un tissu théorique d'hypothèses et de relations déjà existant et reconnu par la communauté scientifique*". Un tel tissu, à notre connaissance, n'existe pas, ou en tout cas serait très difficile à formaliser. Nous avons donc abandonné cette voie pour la validation externe des indices de concurrence étudiés.

SS1-2 LA VALIDITE PREDICTIVE

A partir des mesures de la concurrence entre les marques, il est possible d'effectuer des prévisions, par exemple des prévisions de part de marché. Si les prévisions s'avèrent justes, c'est un gage en faveur de la validité externe de la mesure du concept. Il est possible aussi de procéder à la manoeuvre inverse qui consiste à prévoir la mesure à partir d'autres variables observées. Elrod et Winer (1988) ont utilisé une telle démarche pour comparer cinq méthodes d'analyse "d'obtention de cartes à partir de données de comportements", à savoir :

- une méthode dérivée de l'analyse des correspondances (Carroll, Green et Shaffer 1987)
- le modèle de Levine de décomposition des matrices de transferts en valeurs singulières (1979), qui est aussi une méthode proche de l'analyse des correspondances.
- Choice Map, modèle à point idéal (Elrod 1988)
- le modèle de choix de Green et DeSarbo (1981)
- le modèle de choix de DeSarbo et Cho (1988)

Un des critères d'évaluation utilisé par Elrod et Winer est la validité prédictive qui est évaluée en comparant les probabilités d'achat des marques obtenues à l'aide du modèle avec les parts de marchés réelles de ces marques.

Nous n'utiliserons pas ici de telles prévisions, pour les raisons suivantes :

- l'objectif des indices de mesure de la concurrence n'est pas de faire des prévisions de parts de marché. Ils n'ont pas été conçus dans cette optique, il n'y a donc pas de raisons à les évaluer selon ce critère.
- pour effectuer des prévisions, il faut utiliser les mesures comme intrant dans un modèle de prévision de parts de marchés. Quel modèle choisir ?

  Pour faire des prévisions de parts de marché à partir des deux modèles d'analyse des correspondances et du modèle de DeSarbo (1988), Elrod et Winer ont utilisé les distances des marques au point idéal. La mauvaise qualité de la prévision avec ces modèles peut provenir tout simplement de cette manière d'opérationnaliser la prévision.
- en fin de compte, comment savoir si la bonne ou mauvaise qualité des prévisions est le fait du modèle de prévision utilisé, des données introduites dans le modèle ou encore de l'adéquation des données au modèle ?

SS1-3 LA VALIDITE FACIALE

Nous la définissons comme la *"valeur de crédibilité que représente la mesure auprès d'une population d'experts et utilisateurs de cette mesure"*.

Certains auteurs (Nunnally 1978, Shocker et al 1987) ont insisté sur l'importance de la confrontation de la mesure à de tel avis. Dans une perspective de diffusion des outils (modèles) issus de la recherche Marketing, il est important que l'utilisateur potentiel ait un droit de regard sur la valeur de l'outil que le chercheur lui propose.

Nous avons choisi dans cette recherche d'établir la validité externe des indices sur la base de leur validité faciale. La recherche étant effectuée à partir des données du panel Scan 5000 de la société Nielsen, il semblait logique d'utiliser l'avis de responsables d'études et conseils en clientèle des marchés correspondants.

SS2 ETUDE DE LA VALIDITE FACIALE

SS2-1 MISE EN OEUVRE DE L'ETUDE

Afin que les indices soient évalués sur les bases les plus comparables, nous les avons homogénéisés au maximum.

C'est ainsi que tous les indices ont été rendus symétriques (quand ce n'était pas le cas) en calculant les moyenne de Sij et Sji pour chaque couple de marque (i, j).

Les marques aux parts de marché les plus faibles ont été supprimées afin de n'utiliser (éventuellement) que les versions réduites du marché, si les versions complètes se révélaient difficile à étudier (par exemple si les stress obtenus dans une analyse des similarités sont trop importants).

Nous avons utilisé la typologie hiérarchique associée à l'analyse des similarités (MDS) pour effectuer des représentations graphiques des marchés, pour chacun des indices, excepté pour les deux d'indices IND et ACT.

Pour ces deux indices nous disposions uniquement des données de transferts bruts (en individus ou en actes). Pour les analyser, nous avons utilisé l'analyse factorielle des correspondances. La méthode se prête naturellement à l'analyse de ce type de données. Elle permet notamment de prendre en compte l'effet de taille des observations, c'est à dire la part de marché des marques. L'importance du changement entre deux marques est pondérée par l'inverse de la part de marché des marques. Pour que deux marques soient proches (fortement concurrentes) il faut que les transferts observés entre ces deux marques soient plus forts que les niveaux moyens de transferts de ces marques avec les autres. Nous pourrons comparer les cartes obtenues avec l'analyse des correspondances, avec les représentations obtenues pour les indices TOTT TRTO ELRO et TREL qui utilisent une distance du chi2 plus sophistiquée (voir chapitre IV, Section 3).

De plus, l'analyse factorielle des correspondances (AFC) permet, comme pour les quatre indices précédents, d'obtenir les représentations du marché en deux versions : "Attaque" et "Défense", ce qui permet, entre autre, de prendre en compte la non-symétrie des relations de concurrence. La représentation selon une logique "Attaque" correspond à la représentation des profils colonnes. La représentation selon une logique "Défense" est obtenue avec la représentation des profils lignes. Le plan de traitement suivant a été mis en oeuvre:

Indice Calculé	HUILE	BOISSON	PARFUM	SHAMP
RAA	-	x	-	
BRAB	-	x	-	
FRASA	x	x	x	x
FRASB	x	x	x	x
LEM	x	x	x	x
MERA	x	x	-	x
MERR	x	x	x	x
SABA	x	x	x	x
TOTT	x	x	x	-
TRTO-	x	x	x	-
RAO	x	x	x	x
ELRO	x	x	x	-
TREL	x	x	x	-
IND	x	x	x	x
ACT	x	x	x	x

Un "x" signifie que la représentation du marché a été soumise à l'avis des experts, pour l'indice correspondant.

Les cartes du marché ont été obtenues en deux versions: le marché complet et une version simplifiée du marché, obtenue après suppression des marques aux plus faibles parts de marché; ceci afin de pouvoir comparer les qualités des représentations dans ces deux cas. La suppression des marques aux parts de marché les plus faibles n'améliorant pas de manière sensible la qualité des représentations, nous n'avons retenu pour la suite que les représentations des marchés "complets".

SS2-2 QUALITE DES REPRESENTATIONS OBTENUES

La qualité des représentations qui sont obtenues pour chaque indice est un élément très important. Elle montre avec quel degré les indices calculés se prêtent à une représentation cartographiée de qualité et facilement interprétable, ce qui est un critère de première importance pour le professionnel. Nous faisons l'hypothèse que la qualité d'une représentation (à structure de marché identique) sera meilleure pour une mesure fiable (contenant peu d'erreur). Ce point a, en effet, été étudié par MacKay et Zinnes, à l'aide du modèle de MDS stochastique Multiscale (voir MacKay et Zinnes 1979, 1982, 1986, Zinnes et MacKay 1983). Plus la mesure est entachée d'erreur et plus la valeur du stress est importante, à nombre de dimensions identique. La qualité des représentation avec le MDS est donnée par les valeurs du stress (S) et du R carré (R2). Nous avons donné la signification de ces indicateurs de qualité de l'ajustement à la section 2 précédente.

Pour les AFC, nous avons retenu le pourcentage de variance expliquée par les premiers facteurs. Bien sûr, ces deux types de critères ne sont pas comparables.

Les résultats figurent dans les tableaux suivants.

TABLEAU 20 MARCHE HUILE STRESS ET R2 (sauf pour l'AFC où figure le pourcentage d'inertie expliquée)

	NOMBRE DE DIMENSIONS	
	2	3
INDICE	S R2	S R2
FRASA	.22 (.73)	.07 (.95)
FRASB	.18 (.83)	.10 (.91)
LEM	.17 (.92)	.08 (.97)
MERA	.21 (.74)	.11 (.86)
MERR	.25 (.67)	.13 (.84)
SABA	.21 (.75)	.11 (.88)
TOTT	.11 (.94)	.07 (.97)
TRTO	.12 (.94)	.08 (.97)
RAO	.21 (.73)	.12 (.84)
ELRO	.17 (.86)	.09 (.94)
TREL	.10 (.96)	.05 (.99)
IND	.47	.65
ACT	.43	.61

La structure générale des indices est, malgré le faible nombre de marques (10 marques), mal représentée en 2 dimensions. Nous observons une nette amélioration en trois dimensions. Le marché se structure donc essentiellement selon trois grandes dimensions. Nous avons retenu, pour l'évaluation de la validité faciale, les cartes des représentations selon les deux premières dimensions de l'espace à trois dimensions, ceci afin de faciliter le diagnostic des experts (nous sommes cependant conscients que ceci constitue une faiblesse au plan méthodologique).

Nous remarquons que les indices ne présentent pas le même niveau de qualité de représentation en trois dimensions. Des indices comme MERA, MERR, SABA, RAO ont

des qualités de représentation moyennes, alors que les indices TOTT, TRTO, ELRO et TREL obtiennent une très bonne qualité de représentation.

Ceci peut s'expliquer de deux manières différentes :

- soit les indices de concurrence mal représentés sont affectés d'erreur de mesure : la qualité de la représentation est alors détériorée.
- soit la structure du marché sous-jacente aux indices mal représentés comporte un nombre de dimension supérieur

Pour les indices, Attaque et Défense, la qualité des représentations est attribuable aux propriétés métriques de ces indices ainsi qu'au fait qu'ils prennent en compte la non symétrie des relations de concurrence, à la différence des autres.

TABLEAU 21 MARCHE BOISSON STRESS ET R2 (sauf pour l'AFC où figure le pourcentage d'inertie expliquée)

INDICE	NOMBRE DE DIMENSIONS 2 S R2	3 S R2
BRAA	.06 (.98)	.03 (.99)
BRAB .	.25 (.59)	.15 (.76)
FRASA	.26 (.70)	.14 (.83)
FRASB	.30 (.41)	.19 (.63)
LEM	.21 (.74)	.12 (.89)
MERA	.25 (.85)	.16 (.90)
MERR	.24 (.82)	.17 (.88)
SABA	.30 (.69)	.18 (.84)
TOTT	.17 (.85)	.10 (.92)
TRTO	.15 (.86)	.10 (.96)
RAO	.23 (.68)	.15 (.78)
ELRO	.10 (.98)	.05 (.99)
TREL	.18 (.82)	.10 (.93)
IND	.36	.46
ACT	.30	.40

La qualité de la représentation s'améliore nettement lorsqu'on passe de deux à trois dimensions. Nous avons retenu les deux premières dimensions dans l'espace à trois dimensions, pour l'évaluation de la validité faciale.

Concernant la qualité de la représentation des indices, nous observons les mêmes phénomènes que sur le marché HUILE.

L'apparente bonne qualité de la représentation de BRAA est en réalité le fait d'une solution de type "dégénérée" (voir Shepard dans Davis et Coxon 1981, section 1) pour une explication à ce problème.

TABLEAU 23 MARCHE PARFUM STRESS ET R2 (sauf pour l'AFC où figure le pourcentage d'inertie expliquée)

	NOMBRE DE DIMENSIONS	
INDICE	2	3
	S R2	S R2
BRAA	.06 (.98)	.03 (.99)
BRAB .	25 (.59)	.15 (.76)
FRASA	.26 (.70)	.14 (.83)
FRASB	.30 (.41)	.19 (.63)
LEM	.21 (.74)	.12 (.89)
MERA	.25 (.85)	.16 (.90)
MERR	.24 (.82)	.17 (.88)
SABA	.30 (.69)	.18 (.84)
TOTT	.17 (.85)	.10 (.92)
TRTO	.15 (.86)	.10 (.96)
RAO	.23 (.68)	.15 (.78)
ELRO	.10 (.98)	.05 (.99)
TREL	.18 (.82)	.10 (.93)
IND	.36	.46
ACT	.30	.40

Ici encore, et de façon plus nette que précédemment, la structure du marché se caractérise bien en trois dimensions. Nous observons toujours le même pattern concernant la qualité de la représentation par type d'indices.

TABLEAU 23 MARCHE SHAMP: STRESS ET R2 (sauf pour l'AFC où figure le pourcentage d'inertie expliquée)

	NOMBRE DE DIMENSIONS		
	3	4	5
INDICE	S R2	S R2	S R2
FRASA	.25 (.69	.19 (.77) .	.14 (.81)
FRASB	.25 (.40)	.18 (.51)	.45 (.61)
LEM	.19 (.75)	.14 (.79)	.10 (.86)
MERA	.25 (.38)	.18 (.47)	..25 (.38)
MERR	.25 (.37)	.18 (.47)	.14 (.57)
SABA	.37 (.17)	.19 (.65)	.15 (.70)
RAO	.23 (.52)	.18 (.60)	.13 (.69)
IND	.34 .	.40	.46
ACT	.25	.32	.38

Avec un nombre de marques plus important (24 marques), il est normal d'observer que la qualité des représentations s'est nettement détériorée par rapport aux marchés précédents.

Cependant, la qualité de la représentation de l'indice LEM, à la différence des autres, reste encore acceptable et ne s'améliore pas fondamentalement en passant à des dimensions supérieures. Les indices TOTT, TRTO, ELRO et TREL n'ont pas été calculé sur ce marché, mais nous pouvons supposer qu'il en aurait été de même pour ces quatre indices qui obtiennent de meilleures qualités de représentation que les autres indices sur les trois marchés précédents. Pour les autres indices, il faut cinq dimensions pour que la qualité de la représentation devienne acceptable. Pour certains comme MERA et FRASB elle reste mauvaise. L'évaluation de la validité faciale a été faite en retenant la représentation selon les deux premiers axes de l'espace à cinq dimensions .

Nous retiendrons les élément suivants de cette analyse de la qualité des représentations obtenues:

- les trois premiers marchés sont caractérisés par une structure en trois dimensions. Pour le marché SHAMP, il est plus difficile de répondre, à cause du nombre élevé de marques. Cependant, avec l'indice LEM, l'analyse du stress permet encore d'envisager une structure en trois dimensions.
  Nous retiendrons que les analyses portant sur un nombre important d'objets, comme c'est le cas sur le marché SHAMP, sont difficilement interprétables et donc de peu d'intérêt.
- sur les quatre marchés, la qualité des représentations obtenue est variable selon les indices qui sont utilisés. Les mêmes indices produisent toujours les mêmes qualités de représentation.

Le tableau suivant résume de manière qualitative l'aptitude de chaque indice à fournir des représentations de bonne qualité (à nombre égal de dimensions). IND et ACT ne figurent pas dans ce tableau car ils ont été représentés avec une autre méthode. Quant à BRAA et BRAB, n'ayant pas été représentés sur un nombre suffisant de marchés, ils n'y figurent pas non plus.

TABLEAU 24 EVALUATION GLOBALE DE LA QUALITE DES REPRESENTATIONS
Stress Moyen et R carré *

FRASA	.12 (.88)	Moyenne
FRASB	.16 (.72)	Moyenne
LEM	.10 (.93)	Bonne
MERA	.14 (.88)	Moyenne
MERR	.15 (.84)	Moyenne
SABA	.15 (.81)	Moyenne
TOTT	.07 (.96)	Très Bonne
TRTO	.08 (.97)	Très Bonne
RAO	.13 (.85)	Moyenne
ELRO	.06 (.97)	Très Bonne
TREL	.06 (.97)	Très Bonne

* Stress moyen en trois dimensions obtenu sur les trois premiers marchés .

Nous retiendrons que les indices qui ont révélé une mauvaise fiabilité à la section 2, ne permettent d'obtenir que des représentations de qualité médiocre, à savoir (BRAA, BRAB, MERA, MERR et SABA). Ceci confirme notre hypothèse selon laquelle les représentations de mauvaise qualité sont dues (à marché identique) à de l'erreur de mesure

SS2-3 PROTOCOLE D'EVALUATION DES INDICES

L'ensemble des cartes représentant les marchés (voir Annexe 15A) ont été présentées à un groupe d'experts de la société Nielsen. Avec ces cartes, figuraient les instructions nécessaires à leur évaluation (voir Annexe 15B). Ces instructions portaient:

- Sur la nature de l'indice dont était issue la carte. Les experts ont formulé en effet le souhait de savoir de quel type de concurrence les cartes sont représentatives. Ceci nous indique bien, qu'au moins au niveau des praticiens de l'analyse de la structure des marchés, le concept de concurrence n'est pas univoque. Les praticiens parlent ainsi de concurrence en terme "de transferts", en termes de "parts de choix", etc. Il n'y a pas, dans leur esprit, unicité du concept de concurrence.
- Sur la manière d'évaluer la carte. Il était demandé d'effectuer deux types d'évaluations :
    a- une évaluation relative où l'expert devait fournir un pré-ordre complet des cartes en terme de crédibilité, pour chacun des marchés.
    b- une évaluation absolue où l'expert devait donner une note (sur dix) à chaque carte.

SS2-4 LES RESULTATS OBTENUS

Les résultats de cette évaluation ne sont malheureusement pas disponibles actuellement dans leur intégralité, la société Nielsen n'ayant pas pu effectuer ces évaluations en temps voulu. Nous escomptons cependant obtenir des résultats plus complets prochainement. Nous ne présenterons donc que les premiers résultats relatifs à cette étude de validité faciale. A partir de ces résultats, nous pouvons caractériser les éléments suivants.

INDICES CONSTRUITS A PARTIR DES INTERVALLES INTER-ACHATS

Les indices BRAA et BRAB sont apparus comme n'apportant pas d'éléments intéressants en matière d'analyse de la structure du marché.

Les indices FRASA et FRASB ont apporté des éléments d'analyse de la concurrence, mais pas de façon systématique. Sur le marché des boissons où nous disposions d'un nombre moyen de marques (14) et où toutes les marques ont des parts de marché suffisantes, il apparait que cet indice peut apporter des éléments de diagnostic intéressants sur la concurrence entre les marques. Les proximités apparues semblent être liées au "format du produit".

Cependant, sur un marché comme SHAMP où les marques sont nombreuses et les parts de marché faibles, l'indice ne donne pas de résultats interprétables. Nous attribuons ceci au fait que les estimations individuelles, pour être fiables, doivent être effectuées sur des historiques assez longs, condition qui n'était pas remplie avec cette catégorie de produits.

La faiblesse de cet indice semble ainsi provenir de la fiabilité avec laquelle peuvent être estimés les paramètres (individuels) nécessaires à son calcul.

INDICES CONSTRUITS A PARTIR DES QUANTITES ACHETEES

Les deux indices MERA et MERR n'ont pas donné de résultats interprétables, ceci sur les quatre marchés sur lesquels ils ont été calculés.

INDICES CONSTRUITS A PARTIR DE TRANSFERTS DIRECTS ENTRE MARQUES

Selon l'avis des experts, les indice LEM et RAO présentent des résultats qui sont qualifiés de "moyennement interprétables", notamment pour l'indice RAO. Des trois indices LEM, RAO et SABA, il apparait que c'est l'indice LEM qui fournit les éléments les plus intéressants.

INDICES CONTRUITS A PARTIR DE SIMILARITE DES PROFILS DE TRANSFERTS (indices Attaque et Défense).

Ce sont, de tous les indices, ceux qui ont le plus intéressé les experts, les éléments d'analyse qu'ils apportent sont considérés comme instructifs. Des deux versions, pondérée (TOTT et TRTO) et non-pondérée (ELRO et TREL), il apparait que c'est la version pondérée qui apporte le plus d'éléments d'analyse. Selon l'avis des experts,

l'utilisation d'un indice non-pondéré (en nombre d'individus ayant effectué un transfert) "gommerait une partie du phénomène concurrentiel".
Quant au cartes obtenues en utilisant l'analyse des correspondance sur les indices de transferts bruts, elles apparaissent comme manquant de précision et semblent exagérer la structure des relations, rendant les interprétations confuses. D'autre part, nous devons insister sur le faible pouvoir explicatif de ces analyses lorsque l'on n'extrait que trois dimensions (moins de 50% de la variance totale).

Tous ces résultats demandent bien entendu d'être confirmés par des analyses plus approfondies. Pour notre part, nous retiendrons que les avis d'experts qui ont été donnés sont en forte correspondance avec la qualité statistique des ajustements que nous avons étudiée précédemment. Les indices caractérisés par les plus faibles stress sont ceux qui ont été considérés comme étant les plus pertinents et inversement. Ceci montre, s'il en était besoin, l'intérêt qu'il y a a obtenir des mesures de la concurrence possédant des caractéristiques mathématiques intéressantes, comme c'est le cas pour les indices TOTT, TRTO, ELRO et TREL.

SECTION 5 ETUDE DES FAMILES D'INDICES

La manipulation des facteurs REVENU, CSP, MAGASIN, TAILLE du FOYER, QUANTITE CONSOMMEE et DATE (voir Chapitre IV section 2) a donnée lieu à la création de 17 Univers correspondant aux modalités de ces facteurs. Ces Univers de calculs vont nous permettre de compléter nos analyses sur la similarité des indices et leur regroupement en familles .

Nous avons déjà étudié la similarité des indices dans la section 2 précédente lors de l'étude de leur fiabilité. L'objet de cette section sera de comparer les familles d'indices obtenues à partir de différents types de corrélations :

A - Corrélations des indices calculés sur l'univers GLOBAL, c'est ce qui a été fait dans la section 2.
B - Corrélations moyennes entre les indices calculés sur les mêmes univers (il s'agit de corrélations intra-univers, case 4 du tableau 3, Chapitre IV, section 4).
C - Corrélations moyennes entre les indices calculés sur des Univers complémentaires (il s'agit de corrélations intra-facteurs, case 5 du tableau 3).

Nous pourrons ainsi observer dans quelle mesure la comparaison des indices sur des bases différentes permet d'obtenir des structures de relations identiques.
Afin de mener cette analyse de similarité entre les indices, nous avons construit trois matrices A, B et C de corrélations inter-indices pour chaque marché. Ces matrices figurent dans l'annexe 16.

Les matrices de type A sont constituées des corrélations entre les indices calculés sur l'univers GLOBAL.

Dans les matrices de type B, la similarité entre deux indices est constituée par la moyenne des corrélations intra-univers entre ces deux indices. Pour 17 univers, chaque indice de similarité entre deux indices de concurrence est constitué par la moyenne des 17 corrélations intra-univers entre ces deux indices.

Dans les matrices C, chaque indice de similarité entre deux indices de concurrence est constitué par la moyenne de leurs corrélations intra-facteurs. C'est dans cette condition que l'étude de la structure des indices sera la plus difficile car nous introduisons alors un effet de structure dans l'analyse. Cet effet de structure correspond au fait que la corrélation est calculée entre des mesures effectuées sur deux univers différents. De plus, pour cette raison, les éléments de la diagonale principale ne sont pas égaux à un, ce qui pose problème pour l'utilisation de certaines méthodes, notamment l'analyse factorielle.

Nous pourrons vérifier si les familles obtenues à partir des trois types de matrices de similarités (types A, B et C) sont identiques. Les avantages de cette triple approche sont les suivants :

- elle permet de confronter des résultats obtenus à partir de données différentes
- elle permet de contrôler si la structure des relations entre les indices reste identique quand on inclut un effet de structure, ce qui est fait avec la matrice C.

Pour analyser ces matrices de similarité (de types A, B et C) nous utiliserons conjointement trois types de méthodes qui se prêtent bien à ce type de données:

- l'analyse des similarités
- l'analyse typologique
- l'analyse factorielle en composantes principales

SS1 ETUDE DES FAMILLES D'INDICES A PARTIR DE L'ANALYSE DES SIMILARITES

Nous avons utilisé le modèle ALSCAL dans sa version métrique et agrégée (une seule matrice analysée). La qualité des représentations obtenues (Stress et R carré), figure dans le tableau suivant.

TABLEAU 25 : VALEUR DES STRESS ET R CARRE

	DIM	HUILE S/R	BOISSON S/R	PARFUM S/R	SHAMP S/R
MATRICE A	5	.05/98	.03/.99	-	-
	4	.06/.96	.06/.97	.04/.99	-
	3	.10/.90	.11/.93	.09/.96.	02/.99
	2	.22/.80	.22/.78	.22/.80.	13/.90
MATRICE B	5	.07/.96	.10/.91	-	-
	4.	12/.89.	14/.86.	11/.94	-
	3.	17/.83.	19/.77.	17/.86.	12/.89
	2.	33/.54.	31/.56.	27/.68.	16/.83
MATRICE C	5.	14/.68.	14/.59	-	-
	4.	20/.60.	18/.51.	18/.77	-
	3.	26/.50.	24/.50.	28/.69.	21/.51
	2.	35/.36.	34/.35.	36/.51.	36/.28

Dans ce tableau, figurent en colonnes les quatre marchés étudiés et en lignes les trois types de matrices de similarités inter-indices et dans chaque case la valeur du stress et du R carré de l'analyse, pour la dimension indiquée.

Ces deux indicateurs de la qualité des représentations nous permettent de contrôler :

- le degré de fiabilité de l'analyse
- en dessous de quelle dimension nous observons une dégradation considérable des indicateurs de qualité, ceci permettant de cerner la dimension de la structure étudiée.

L'examen de ce tableau révèle les éléments suivants :

- nous obtenons les mêmes structures de stress et R carré sur les trois marchés HUILE, BOISSON et PARFUM. Le quatrième (SHAMP), qui ne comporte que sept indices, n'est pas comparable aux autres.
- nous obtenons les mêmes structures de stress et R carré pour les trois types de matrices de corrélations analysées (A, B et C). D'autre part, les représentations à

partir des matrices A (calculées sur l'univers GLOBAL) sont toujours les meilleures. Ceci est le signe que la structure est plus forte et qu'elle est moins sujette à erreur (variance) dans la mesure des indices de similarité entre les indices. Les représentations à partir des corrélations de type B ont une qualité qui, bien que moins bonne que précédemment, reste intéressante. La qualité de la représentation se dégrade nettement avec les matrices de type C, ce qui était prévisible compte tenu du fait que ces corrélations prennent en compte deux types d'effets : la similarité des indices d'une part, mais aussi celle des univers (qui sont différents) sur lesquels sont a calculés les indices.

Cependant, quelque soit le cas A, B ou C, il apparait que la structure comprend trois dimensions, le point d'inflexion dans l'augmentation du stress se situant au passage entre trois et deux dimensions. Nous en concluons que les indices se structurent en trois familles.

La comparaison entre les différentes représentations obtenues caractérise les éléments suivants (les cartes obtenues figurent à l'Annexe 17). La structure des relations obtenue à partir des matrices de corrélations de types A et B est très similaire. A partir des matrices de type C, nous trouvons l'essentiel de la structure observée sur A ou B, mais de façon moins nette. Les structures obtenues sont les suivantes.

α MARCHE HUILE

Trois familles naturelles (F1, F2 et F3) se détachent nettement:

F1 FRASA et FRASB
F2 TOTT TRTO ELRO TREL
F3 MERA et MERR

A ces trois familles viennent se rapprocher les indices restants, à savoir LEM avec F1 et RAO et SABA avec F3.

ß MARCHE BOISSON

Nous retrouvons les trois familles de base (F1, F2, F3) mais avec quelques différences dans la manière dont les autres indices s'en rapprochent. LEM, RAO et SABA sont maintenant regroupées ensembles, avec FRASA et FRASB.

Nous observons que TREL et TRTO sont fortement associés, montrant par là que les indices Attaque se ressemblent plus entre eux, qu'ils ne ressemblent aux indices Défense, même si ils sont calculés sur deux types de données différentes.

γ MARCHE PARFUM

L'indice MERA n'étant pas calculé sur ce marché, la famille F3 précédente disparait. L'indice MERR se trouve isolé. Nous retrouvons toujours les familles F1 (FRASA et FRASB) et F2 (TOTT, TRTO, ELRO et TREL). LEM et RAO sont bien associés, alors que SABA est nettement isolé de tous les autres. Il existe ici un lien très fort entre les deux indices Défense : TOTT et ELRO.

π MARCHE SHAMP

En l'absence des indices de la famille F2, nous trouvons toujours une séparation bien nette entre les familles F1 et F3. D'autre part, LEM et RAO sont biens associés, SABA l'est aussi mais à un degré moindre.

En résumé, trois familles de base qui ont été caractérisées:

F1 FRASA FRASB
F2 TOTT TRTO ELRO TREL
F3 MERA MERR

A ces trois familles se rattachent les indices restant, à savoir LEM RAO et SABA. LEM apparait en fait comme étant un indice central, à l'inverse SABA se rattache en général assez difficilement aux trois familles précédentes. Quand il se rattache, nous pouvons constater que c'est par l'intermédiaire de RAO.

SS2 ETUDE DES FAMILLES D'INDICES A PARTIR DE L'ANALYSE TYPOLOGIQUE

Nous avons effectué une série d'analyses typologiques dans les mêmes conditions que les analyses de similarité précédentes. La méthode utilisée est la typologie hiérarchique (Johnson 1967). La similarité entre deux groupes est mesurée par la distance moyenne entre les éléments de ces groupes. Le choix de ce critère se justifie par le fait que les mesures de similarité inter-indices qui sont utilisées sont métriques.

Les résultats des analyses figurent à l'annexe 18. Ils confirment bien les résultats obtenus précédemment, pour les matrices de similarités de type A et B. Par contre, pour les analyses sur les matrices de type C, les typologies ne sont pas claires. La structure étant moins forte, elle ne ressort pas à partir de ce type d' analyse (les indices se rattachent successivement les uns aux autres sans qu'il y ait constitution de groupes).

D'un point de vue méthodologique, ces deux types d'analyses (MDS et typologie) suggèrent les remarques suivantes (qui bien sûr nécessiteraient une validation) qui valent la peine d'être notées dans le cadre de cette recherche, car ces deux méthodes sont souvent utilisées conjointement en marketing, dans le cadre des études produits:

- Lorsque les données possèdent une structure forte, exempte de variance (variance qui peut être considérée comme de l'erreur de mesure), MDS et Typologie convergent remarquablement bien. C'est le cas notamment avec les matrices de type A qui correspondent à des corrélations inter-indices calculées sur le même univers GLOBAL.
- Lorsque les données ont une structure moins nette (présence de variance ou d'erreur de mesure dans l'estimation des similarités), la typologie semble moins apte à restituer la structure d'origine, alors que celle-ci reste encore assez nette avec l'analyse des similarités. C'est ce que nous avons pu observer sur la matrice C.

SS3 ETUDE DES FAMILLES D'INDICES A PARTIR DE L'ANALYSE FACTORIELLE

Pour ce troisième type d'analyse, nous avons utilisé l'analyse factorielle classique. L'estimation des saturations (loading) est obtenue avec l'algorithme de l'analyse en composantes principales. Nous avons effectué ensuite une rotation orthogonale (critère Varimax) afin de faciliter l'interprétation des résultats.

DIMENSION DE LA STRUCTURE DES INDICES

Pour sélectionner le nombre de facteurs à retenir, nous avons étudié les pourcentages de variance expliquée par les premiers facteurs. Ces valeurs figurent dans le tableau suivant. En colonnes figurent les quatre marchés étudiés et en lignes les trois types de corrélations analysées (A, B et C). Dans chaque case, figure le pourcentage de variance cumulée expliquée par le nombre de facteurs correspondants à la ligne.

TABLEAU 26 POURCENTAGES CUMULES DE VARIANCE EXPLIQUEE

	Dim	HUILE	BOISSON	PARFUM	SHAMP
MATRICE A	1	.33	.35	.49	.39
	2	.51	.54	.66	.65
	3	.68	.70	.75	.82
	4	.77	.78	.84	.94
MATRICE B	1	.30	.30	.37	.37
	2	.45	.45	.54	.64
	3	.59	.57	.67	.83
	4	.70	.67	.76	.95
MATRICE C	1	.25	.22	.26	.23
	2	.36	.33	.37	.38
	3	.45	.41	.48	.52
	4	.53	.49	.57	.65

L'examen des graphiques représentant le pourcentage de variance expliquée ainsi que l'utilisation du critère de la valeur propre supérieure à un, permet de caractériser sur les quatre marchés une structure en trois dimensions (voir tableau 26). Même sur le marché SHAMP où seulement 7 indices sont analysés, cette structure en trois dimensions apparait.

L'analyse factorielle confirme les résultats obtenus avec les méthodes précédentes. Ici encore, la qualité de la représentation est la moins bonne avec les corrélations de type C. L'examen de las structure des facteurs (corrélations des indices avec les facteurs après rotation Varimax) nous apporte les éléments suivants.

1 Les analyses de type A et B sont toujours très similaires. Les analyses de type C sont, bien qu'à un degré moindre, représentatives de la même structure. Il y a une bonne convergence entre les trois types d'analyses.

2 Les indices se structurent en trois grandes familles :

F1 : FRASA et FRASB
F2 : MERA et MERR
F3 : TOTT, TRTO, ELRO et TREL

Sur ces trois familles, viennent se raccrocher, de manière pouvant différer d'un marché à l'autre, les trois autres indices LEM, RAO et SABA. Les indices LEM et RAO sont souvent regroupés ensembles sur l'un des trois facteurs précédents (essentiellement F1 ou F3). Encore une fois, l'indice LEM apparait comme un indice central. Il est souvent lié à plus d'une famille à la fois. Par exemple, sur HUILE, il est relié aux familles F1 et F3.

L'indice SABA a beaucoup plus de difficultés à se raccrocher à l'une des trois familles. Lorsqu'il y parvient, il apparait que c'est toujours par l'intermédiaire de l'indice RAO.

Les valeurs des communalités (voir annexe 19) caractérisent le degré d'appartenance globale de chaque indice aux trois familles F1, F2 et F3. L'observation de ces communalités montre le faible degré d'appartenance de l'indice SABA à ces trois familles.

3 Les rotations obliques permettent de vérifier si la structure obtenue est bien orthogonale (voir annexe 19). Les structures obtenues après rotation oblique ne sont pas modifiées mais, nous pouvons observer que les facteurs portant les familles F1 et F3 sont plus corrélés entre eux qu'avec la troisième famille F2 (excepté sur le marché HUILE où toutes les corrélations sont faibles). Nous vérifions là ce que nous avions constaté précédemment à la section 2, à savoir que les indices MERA et MERR sont bien différents de tous les autres (excepté sur le marché HUILE). Les autres indices, même si ils n'appartiennent pas à la même famille, présentent des degrés d'association non négligeables.

Pour conclure cette section consacrée à l'étude de la similarité entre les indices, nous retiendrons les éléments suivants :

- l'ensemble des méthodes utilisées (MDS, typologie et analyses factorielles) convergent pour caractériser trois grandes familles :

- F1 : FRASA, FRASB
- F2 : MERA, MERR
- F3 : TOTT, TRTO, ELRO, TREL

Les trois indices restants, LEM, RAO et SABA, se rattachent alternativement aux familles F1 et F3. LEM apparait comme étant un indice central, il est relié simultanément aux familles F1 et F3. L'indice SABA, quand il est relié aux autres, présente un lien privilégié avec l'indice RAO dont il est proche par la forme du calcul. Cependant, cet indice est certainement affecté d'un fort degré d'erreur de mesure tant son degré de lien avec les autres apparait fluctuant. Nous avons montré à la section 2, que l'utilisation d'un indicateur de mesure du lien non-linéaire (coefficient de corrélation de Spearman)

améliore le degré d'association de cet indice avec LEM et RAO, mais pas de manière systématique (uniquement sur deux des quatre marchés).

Les indices MERA et MERR apparaissent comme étant d'un type différent à tous les autres, excepté sur le marché HUILE où ils s'associent bien avec les indices RAO et SABA. Nous insisterons, pour terminer, sur le fait que différentes techniques statistiques appliquées à différents niveaux de lien entre les indices (matrice de corrélations du type A, B ou C), sur des marchés différents, ont convergées vers ces résultats. Les familles obtenues peuvent être considérées comme "solides".

CONCLUSION AU CHAPITRE V

Les analyses de validité que nous venons de mener laissent apparaître les éléments suivants :

α CONCLUSIONS RELATIVES A LA MESURE DU CONCEPT DE CONCURRENCE

1 Tous les indices ne convergent pas vers le même concept. Ceci peut être expliqué de deux manières différentes :

- soit ils ne sont pas une mesure pertinente du concept de concurrence, ils mesurent alors *"autre chose"*.
- soit, ils mesurent une autre dimension du concept de concurrence. Dans ce cas, il faut conclure à l'indépendance (orthogonalité) de cette dimension de la concurrence par rapport à la dimension vers laquelle convergent les autres indices étudiés. Notamment, les indices MERA et MERR calculés à partir des *"parts de choix"* se révèlent être la mesure d'une dimension indépendante de celle qui est mesurée par les autres indices.

Dans cette optique, nous pourrons envisager la concurrence comme un concept comportant plusieurs dimensions:

- une concurrence de type *"dynamique"*, mesurée à partir de l'observation des comportements de transferts entre marques. Dans ce cas le changement d'une marque pour une autre exprime *"de fait"* une concurrence qui se manifeste au niveau du marché.
- une concurrence d'un point de vue *"statique"*, à partir de l'observation des parts de choix des consommateurs. Le postulat est alors que deux marques qui ont été consommées dans les mêmes proportions par un consommateur au cours d'une période donnée sont concurrentes l'une de l'autre. Quoiqu'il en soit, ces deux définitions de la concurrence ne convergent pas vers des mesures identiques.

2 Pour les indices qui convergent, c'est à dire tous les indices sauf BRAA BRAB MERA MERR, nous pouvons observer qu'il n'y a pas de problème de validité discriminante. Les mesures sur des marchés différents sont bien indépendantes les unes des autres. Nous retiendrons tout de même que, contrairement à l'esprit originel de la méthode, les traits mesurés sont nettement différents, ce qui a facilité le test de

discrimination. Cependant, ce résultat se vérifie avec suffisamment de force pour que puissions considérer la validité de traits comme étant acquise.

3 Parmi les indices qui convergent, certains sont affectés d'une erreur de mesure (aléatoire ou systématique) assez importante. Il s'agit notamment de l'indice SABA et dans une moindre mesure des indices RAO, FRASA et FRASB.

4 Les indices construits sur la base des intervalles inter-achats (FRASA et FRASB), tels que nous les avons définis à la section 3 du chapitre IV, convergent assez bien avec les indices construits sur la base des transferts entre marques. Nous avions d'ailleurs noté, d'un point de vue théorique, les points communs entre ces deux types d'indices. En effet, l'intervalle inter-achats est bien une mesure de transfert (dynamique) entre deux marques, à laquelle est ajoutée la notion de durée entre les deux achats (pondérée par la quantité consommée). Cet indice est très séduisant d'un point de vue théorique, cependant sa sophistication parait se payer en terme de difficulté d'estimation et en fin de compte, en erreur de mesure. Nous avons montré que pour les marques à faible part de marché, les coefficients de forme de la loi gamma (les ri) sont très mal estimés; leurs valeurs sortent des intervalles de variations communément admis pour ce paramètre (voir en Annexe 2). Pour être estimés de façon fiable, les paramètres de cet indice nécessitent :

- des marques à part de marché suffisante (ce qui est problématique pour l'étude des petites marques ou des produits nouveaux).
- des périodes d'observation très longues pour chaque ménage pour lequel sont estimés les paramètres, avant agrégation (Fraser et Bradford, 1983 ont noté ce point). Or, nous l'avons dit, des périodes d'observation supérieures à deux années ne nous semblent pas raisonnables.

5 Les résultats obtenus concernant la validité interne des indices, nous amènent à penser que les indices estimés au plan individuel puis ensuite agrégés pour la population toute entière sont moins valides que les indices calculés directement au niveau agrégé. Ceci rejoint ce que nous disions précédemment à savoir que pour de tels indices, l'amélioration de la précision de leur estimation nécessite de disposer de périodes d'observation plus longues, ce qui n'est pas toujours très réaliste dans le contexte de marchés en évolution permanente. Nous avons constaté le faible degré de fiabilité des indices MERR et MERA et aussi FRASA et FRASB quand ils sont estimés sur des périodes de six mois.
Par contre, l'amélioration de la qualité d'indices de type *"agrégé"* passe par l'augmentation du nombre d'actes d'achats observés. Ceci peut être obtenu en augmentant la taille de l'échantillon des panélistes, ce qu'a fait récemment la société Nielsen, en passant de 5000 à 7000 panélistes avec SCAN 7000 (10000 panélistes prévus pour 1992).

6 Les indices SABA et RAO présentent un niveau de fiabilité et de validité de traits très moyen, comparé à celui de l'indice LEM. Nous avons, dans les paragraphes précédents, essayé d'en expliquer les raisons. Cette faiblesse est certainement le fait de fluctuations attribuables à la forme multiplicative du dénominateur. Notons que la validité

de SABA s'améliore nettement sur deux marchés (BOISSON et SHAMP) quand nous avons utilisé un coefficient de corrélation non-linéaire (coefficient de Spearman).

7 Les indices que nous avons proposés et qui sont construits sur la base d'une pseudo- distance du chi2 (entre lignes ou entre colonnes) sont, à quelques cas près, bien liés avec les autres indices de transferts directs (LEM, RAO, SABA). Nous avons pu constater que les deux versions (Attaque et Défense) de ces deux indices peuvent produire des mesures de la concurrence assez différentes, caractérisant ainsi un manque de symétrie des relations de concurrence entre les marques. L'indice TOTT peut fournir des mesures différentes de celles de TRTO et de même pour ELRO et TREL. Par contre, l'utilisation de données de transferts en *"individus"* ou en *"actes"* n'affecte pas le calcul de ces indices. Les deux versions TOTT et ELRO d'une part et TRTO et TREL d'autre part sont toujours très similaires. Sur certains marchés, nous avons pu constater que la fiabilité dans le temps de ces indices est très faible. Ceci peut être attribué à la finesse de leur calcul (comparaison de profils), qui les rend plus sensibles à de légères modifications de ces profils.
Néanmoins, il est très important de retenir que des mesures *"indirectes"* de la concurrence (par comparaison des profils d'échanges des deux marques avec toutes les autres, sauf elles-mêmes) convergent bien vers des mesures directes (le transfert entre ces deux marques). Ces mesures présentent donc un réel intérêt, comme nous l'avons décrit à la section 3 du chapitre IV.

8 Les transferts bruts (IND et ACT) se sont révélées être extrêmement fiables dans le temps (en test-retest) et être bien corrélées en moyenne avec les autres indices. Ceci constitue un gage de fiabilité pour toutes les mesures qui utilisent *"à la base"* ces transferts comme intrant pour la mesure de la concurrence.

9 Du point de vue de la validité faciale, les indices *"Attaque"* et *"Défense"*, dans leur version pondérée (TOTT et TRTO), semblent fournir les mesures les plus intéressantes. Il convient néanmoins de rester prudent puisque ces résultats découlent d'une analyse qui reste pour l'instant "exploratoire".

10 Utilisation du concept "d'achats simultanés" pour mesurer la concurrence entre deux marques.
Un quatrième concept pouvant servir de base à la mesure de la concurrence n'a pas été abordé dans cette recherche parce qu'il n'a pas encore, à notre connaissance, fait l'objet de recherches et d'applications dans le domaine de l'analyse de la structure des marchés. Les achats simultanés constituent en fait une masse d'information qui n'a été que très peu exploitée dans l'étude de la concurrence à partir des actes d'achats. Pourtant, le fait que deux marques appartenant à un univers de concurrence défini au préalable soient présentes de façon simultanée dans le panier d'achat d'un consommateur en une occasion d'achat particulière, est révélateur du lien qui peut unir ces deux marques.
Une telle mesure de la concurrence serait équivalente, au niveau du comportement d'achat du consommateur, au concept d'appartenance au même ensemble évoqué, concept sur lequel est construit l'indice de substituabilité proposé par Chandon et Strazziéri (1986, voir première partie, chapitre II, section 4).
La mesure pourrait alors être opérationnalisée de la manière suivante:

Soit n_{ij}^l le nombre de co-présences des marques i et j lors des n.. occasions d'achat du consommateur l. La matrice S de terme général:

$$N_{ij} = \sum_{l=1}^{L} n_{ij}^l$$

serait une mesure agrégée au niveau des N consommateurs, du lien de concurrence existant entre les marques i et j.

De nouvelles recherches sont donc à attendre dans cette direction, recherches qui permettraient d'apprécier la validité de cette mesure et ce qu'elle apporte de spécifique par rapport aux autres mesures. Par exemple, une telle mesure pourrait être une manière de prendre en compte la recherche de variété de type inter-temporel, à partir de données de comportement d'achat puisque l'analyse se place alors dans le cas où un consommateur achète, en une occasion particulière, des produits qu'il juge "complémentaires" et qu'il compte consommer pour satisfaire son besoin de recherche de variété. Un tel type de comportement peut être, par exemple, observé au sein de catégories de produit comme les boissons aux fruits.

10 Conclusions relatives aux propriétés souhaitables d'un indice de mesure de la concurrence:

Un indice de mesure de la concurrence peut finalement être considéré, d'un point de vue méthodologique comme étant constitué de deux composantes:

1- Une matrice des données servant de base au calcul de l'indice, nous l'appellerons *"matrice des intrants de base"*. Ces intrants peuvent être qualifiés de "mesure brute de la relation entre deux marques". Nous pouvons considérer qu'ils reposent sur quatre types de concepts:

- les transferts entre marques
- les intervalle inter-achats
- les parts de choix
- les achats simultanés

Les indices ACT et IND, transferts bruts entre marques (pondérés et non pondérés par les quantités achetées), sont deux exemples de telles matrices d'intrants pouvant servir de base au calcul d'indices de mesure de la concurrence.

2 Une *"formule de calcul"*, qui à partir d'une matrice d'intrants de base, permet de calculer des indices de concurrence entre les marques. Par exemple, l'indice de Lehmann procède à partir de la matrice N de terme général n_{ij}: le nombre d'individus ayant effectué au moins un transfert entre les marques i et j sur la période étudiée. Une telle matrice peut être ensuite utilisée pour servir d'intrant au calcul de la matrice d'indices de mesure de la concurrence S de terme général s_{ij}:

$$s_{ij} = \frac{n_{ij}}{n_{i.}+n_{.j}}$$ où $n_{i.}$ et $n_{.j}$ sont les totaux marginaux en lignes et en de la matrice N.

Il apparait ainsi, qu'à partir d'une même matrice d'intrants, il est possible de calculer différents types d'indices.
Nous noterons encore, pour illustrer l'intérêt d'un découpage du calcul d'un indice en deux phases, l'utilisation possible des indices que nous avons développés dans cette recherche (les indices Attaque et Défense), sur une matrice d'intrants dont les éléments seraient les achats simultanés entre de deux marques. Le même type d'indices (Attaque et Défense) qui ont l'avantage de prendre en compte la non-symétrie des relations de concurrence, d'avoir les propriétés d'une distance, d'être robuste à une définition incorrecte du champ concurrentiel, pourraient alors être calculés en utilisant deux types d'intrants différents: les transferts entre marques et les achats simultanés. Une telle mise en oeuvre aurait l'avantage de fournir des représentations du marché selon différents points de vue; ici les points de vue reposant sur les transferts entre marques et les achats simultanés. Ces différents points de vue reposeraient sur différentes opérationnalisations de la concurrence mais ne différeraient pas pour autant dans leur mode de calcul. C'est ainsi que des cartes obtenus à partir d'intrants différents traités à l'aide de la même formule de calcul pourraient être utilisées avec bénéfice car ils seraient alors directement comparables.

Le double niveau d'analyse d'un indice de mesure de la concurrence que nous venons de proposer peut servir de base à une double analyse des propriétés que l'on pourrait attendre d'un "bon indice de mesure de la concurrence".

• Tout d'abord, un bon indice devrait être bâti sur une matrice d'intrants possédant des caractéristiques conceptuelles (théoriques) satisfaisantes.

• D'autre part, les caractéristiques de cet indice, que nous qualifions de statistiques ou encore méthodologiques, sont liées à sa formule de calcul. Les indices Attaque et Défense permettent par exemple de prendre en compte la non-symétrie des relations de concurrence. La même propriété, qui est de type méthodologique, sera obtenue avec cet indice, qu'il soit calculé à partir des concepts de transferts entre marques ou à partir de l'appartenance simultanée au même panier d'achat. Nous allons donc successivement présenter les caractéristiques attendues pour un indice, au plan conceptuel puis au plan méthodologique.

<u>Caractéristiques attendues au plan conceptuel (matrice d'intrants):</u>
Les propriétés suivantes pourraient être attendues, au plan conceptuel, d'une matrice d'intrant servant ensuite de base au calcul d'un "bon indice":

- permettre de savoir si deux marques sont substituables ou si elle sont achetées pour satisfaire des besoins correspondant à des situations d'usage différentes (complémentarité). Si cela était possible, il serait non seulement possible de qualifier deux marques comme concurrentes, à partir de données de comportement, mais en plus de savoir si cette concurrence correspond, au niveau du consommateur, à de la substituabilité ou de la complémentarité.
- permettre de faire la différence entre deux marques achetées parce quelles sont substituables ou parce qu'elles satisfont un besoin anticipé de recherche de variété (complémentarité).
- permettre de savoir si deux marques sont achetées parce qu'elles sont substituables au sein du ménage ou parce qu'elles satisfont aux différents besoins des individus composant ce ménage.

- prendre en compte la dynamique des actes d'achats et l'intégrer dans la mesure de la concurrence
- s'intégrer au sein d'une théorie sur la modélisation du comportement du consommateur

Le tableau suivant permet de caractériser les quatre concepts à la base du calcul d'un indice de concurrence selon les différentes qualités que nous venons de citer.

TABLEAU 27 PROPRIETES DES INDICES AU PLAN CONCEPTUEL (THEORIQUE)

	Transferts entre marques	Intervalles inter-achats	Parts de choix	Achats simultanés
Prise en compte de la situation d'usage	Non	hypothétique	non	non
Prise en compte de la recherche de variété	non	hypothétique	non	non
Prise en compte de l'hétérogénéité des goûts au sein du ménage	non	hypothétique	non	non
Prise en compte de la dynamique du comportement	oui	oui	non	non
Calcul direct au plan agrégé	oui	non	non	oui
Nécessité de définir un univers concurrentiel	oui	oui	non	oui
S'intégrer au sein d'une théorie du choix du consommateur	Oui	Non	Non	Non encore étudié

Caractéristiques attendues au plan statistique (méthodologique)

Un "bon indice" de mesure de la concurrence devrait posséder, au plan statistique (méthodologique) les propriétés suivantes:

- rendre compte du caractère non-symétrique des relations de concurrence
- tenir compte des niveaux de fidélité aux marques dans l'évaluation du niveau de concurrence existant entre deux marques sous hypothèse d'indépendance
- posséder les caractéristiques d'une distance, à savoir la distance minimale (la distance d'un objet à lui même est plus petite que n'importe quelle autre distance), la symétrie de la mesure, l'inégalité du triangle
- prendre en compte les différences de niveau d'achat (parts de marché) des marques.En effet, le niveau de relation entre deux marques ne doit pas être fonction des niveaux d'achats de ces marques. Par exemple, dans l'indice proposé par Rao et Sabavala, le niveau des parts de marché est pris en compte en standardisant la valeur du transfert entre deux marques (n_{ij}) par le produit des volumes d'achat de chaque marque: $n_{i.} \times n_{.j}$ (le transfert attendu sous hypothèse d'absence de concurrence entre ces deux marques)
- être facilement calculé
- être robuste à une diminution de la taille de l'échantillon des actes d'achats sur lequel il est calculé, de manière à rester fiable lors qu'il est calculé, pour les besoins d'une étude, auprès d'une sous-population particulière
- ne pas être trop sensible à des différences des caractéristiques des individus figurant dans le panel ou du magasin dans lequel sont observés les actes d'achats
- enfin, il ne devrait pas, dans son estimation, être trop fortement dépendant d'une définition particulière de l'univers de concurrence. En effet, nous avons vu qu'à partir de données de comportement, il n'existe pas de méthode satisfaisante permettant de délimiter les frontières d'un marché. Cette définition ne peut alors se faire qu'avec une part importante d'arbitraire. Dans ces conditions, il est bon que l'estimation d'un indice de concurrence ne souffre pas trop de cet arbitraire.

Le tableau suivant permet de caractériser les différentes formes de calcul d'un indice de concurrence que nous avons étudié dans cette recherche selon les caractéristiques attendues d'un bon indice de mesure de la concurrence.

TABLEAU 28 PROPRIETES DES INDICES AU PLAN STATISTIQUE (METHODOLOGIQUE)

	LEM	RAO	Attaque Défense	MERA MERR	FRASA FRASB
Prise en compte de la non-symétrie	non	non	oui	non	non
Prise en compte du niveaux de fidélité à la marque	non	non	oui	non	non
Propriété d'une distance	non	non	oui	oui	oui (en théorie)
Prise en compte des différences de niveau de consommation des marques	non	oui	oui	non	non
Simplicité du calcul	oui	oui	oui	oui	non
Robustesse à la diminution de la taille échantillon	oui	non	non	non	non
Pas trop sensible à une mauvaise définition de l'univers concurrentiel	non	non	oui	non	non
Pas trop sensible aux caractéristiques des individus	oui	non	non	non	non

ß CONCLUSIONS RELATIVES A LA METHODOLOGIE

Nous avons pu réunir des éléments intéressants concernant les études de validité :

1- Dans une étude de validité, le choix d'une mesure d'association (coefficient de corrélation) entre les mesures mérite peut-être plus d'attention que celle qui lui est donnée à l'heure actuelle dans la recherche marketing, où le coefficient de corrélation de Pearson est utilisé de manière systématique. Evaluer le niveau du lien entre deux mesures ne signifie pas forcément qu'il faille évaluer un lien de type *"linéaire"*. Nous avons constaté (de façon très empirique), que l'utilisation d'un indice de lien *"monotone"* améliore les niveaux apparents de validité interne.

2 - L'utilisation des modèles d'équations simultanées à variables latentes (modèle Lisrel) permet d'effectuer en une seule fois, l'analyse de la fiabilité et de la validité de traits des indices. Cette méthode permet d'utiliser la rigueur des tests statistiques pour évaluer la validité interne. La faiblesse de ce type d'analyse, par rapport aux méthodes plus conventionnelles que nous avons utilisées, se place à un niveau technique: la matrice des corrélations peut ne pas être définie positive, ce qui, dans notre cas, nous a interdit l'utilisation de la procédure d'estimation par le maximum de vraisemblance.

D'autre part, ce type de méthode est peu adaptée quand la taille de la matrice des corrélations est importante, ce qui était notre cas au début de l'analyse. Finalement,

l'association de plusieurs méthodes (analyse factorielle classique, modèle Lisrel), se révèle tout à fait profitable.

3 - Dans les études sur la validité faciale, nous avons pu constater l'existence d'un lien entre la qualité de l'ajustement statistique que permet d'obtenir un indice (lors de la construction de cartes à l'aide des techniques de scaling) et la *"pertinence"* que représente cet indice aux yeux des experts qui en sont les utilisateurs potentiels. Notons encore, cependant, le caractère approximatif de cette analyse qui est du au fait que des représentations en trois dimensions étaient, par mesure de simplification, évaluées sur des cartes en deux dimensions.

CHAPITRE VI ANALYSE DE L'EFFET DES FACTEURS

L'objet de ce chapitre est d'analyser l'influence des variables manipulées (Revenu, CSP, Magasin,Taille du Foyer, Quantité Consommée, Période d'observation) sur la structure d'un marché étudiée à partir de données de comportement.

Mieux connaître l'effet de ces facteurs correspond selon nous à deux types de motivations: des motivations d'ordre théorique et des motivations d'ordre méthodologique.

1 - MOTIVATIONS D'ORDRE THEORIQUE

La question que l'on se pose est la suivante : "la structure d'un marché obtenue au plan global reste-t-elle encore valide quand L'analyse ne porte plus que sur un segment particulier de la population totale" (par exemple les consommateurs aux gros revenus) ?

La structure d'un marché est elle spécifique à des segments de consommateurs particuliers ou bien à l'inverse, est-ce un concept stable, valable quel que soit le segment ?

Nous avons noté, dans la première partie (chapitre IV, section 1), que les recherches ont fait apparaître une bonne stabilité de la structure des marchés lorsqu'elle est établie à partir de données perceptuelles. Par contre, d'autres études ont montré que les préférences des individus sont plus hétérogènes, moins stables, que leurs perceptions. Les cartes de marchés construites à partir de données de préférence sont moins stables que celles qui sont construites à partir de données de perception.

Si nous considérons que l'acte d'achat est l'expression d'une préférence, nous pouvons nous attendre alors à ce que les structures obtenues à partir de données de comportement soient moins stables que celles qui sont établies à partir de données de perception.

D'autre part, nous avons vu dans la première partie que plusieurs recherches ont montré le faible degré de lien entre préférences et comportements...

Comme nous le voyons, les éléments apportés par la théorie du comportement du consommateur ne facilitent pas le débat. Une étude empirique dans ce sens présente un intérêt évident.

2- MOTIVATIONS D'ORDRE METHODOLOGIQUE

Nous étudierons ici la robustesse du concept de "Mesure de la Concurrence entre les Marques" à la manipulation de variables d'influence. Si nous caractérisons des effets importants, de la part de certaines variables, sur la structure d'un marché, cela signifiera que la structure des échantillons de panélistes doivent être particulièrement fiable selon ces variables.

D'autre part, il sera nécessaire, lors du calcul des indices de concurrence entre les marques, de contrôler sérieusement la répartition des panélistes selon ces variables. Dans le cas contraire, le praticien ou chercheur court le risque d'obtenir des structures non

représentatives. Si certaines variables montrent une forte influence sur la structure d'un marché, il faudra vérifier si ces influences persistent d'un marché à l'autre.

Peu d'études ont analysé l'influence des variables classiques (comme la CSP, l'âge, la quantité consommée, etc..) sur la mesure de la concurrence. Nous citerons cependant Chance et French (1972) qui ont étudié le lien des transferts entre marques avec les variables revenu et niveau d'éducation. Les auteurs ont montré que les niveaux des transferts entre marques dépendent fortement de la classe de revenu des ménages observés. Cependant, nous ne pouvons pas retenir un tel résultat car la recherche étudiait en fait des transferts fictifs en observant les achats des ménagères avant de connaître le prix des marques et après en avoir pris connaissance. Dans de telle conditions, il est logique que la variable revenu apparaisse comme étant très influente. D'autre part, dans cette étude, les achats sont "simulés" et les échantillons sont de taille très faible (soixante ménagères par groupe, au maximum).

Plus récemment, Parry et Gengler (1988) ont étudié la fréquence d'achat sur les estimations des paramètres du modèle de Grover et Srinivasan (voir partie I, chapitre II). Les auteurs concluent à l'influence de cette variable sur les valeurs des paramètres estimés.

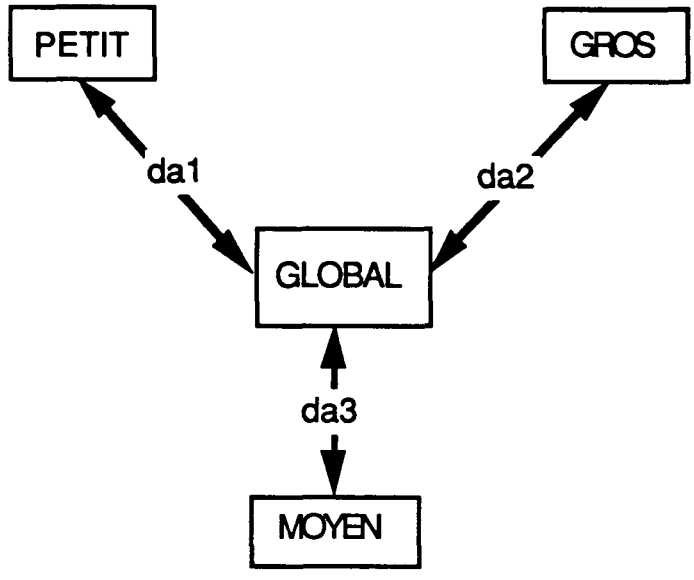
SECTION 1 LA MISE EN FORME DU PROBLEME :

Telles que les données se présentent, l'analyse de l'influence des facteurs est délicate . Deux voies sont possibles pour analyser ces influences.

SS1 ETUDE A PARTIR DES DISTANCES A L'UNIVERS GLOBAL

Il est possible d'étudier le lien entre chacun des univers issus de la manipulation d'un facteur et l'univers GLOBAL. L'éloignement entre les deux univers est une mesure de l'influence du facteur manipulé. Le graphique suivant permet de visualiser le problème.

GRAPHIQUE 3A EFFET DES FACTEURS : DISTANCES A L'UNIVERS GLOBAL



Sur ce graphique, nous visualisons (de façon tout à fait informelle) l'effet attribuable à chacune des modalités du facteur Revenu.

soit da1 la distance entre les modalités REVG et GLOBAL

d_{a1} = d (REVG, GLOBAL)

De même:

soit da2 la distance entre les modalités REVM et GLOBAL
soit da3 la distance entre les modalités REVP et GLOBAL

Les distances d_{a1}, d_{a2} et d_{a3} sont des fonctions inverses des corrélations entre chaque univers issu de la manipulation du facteur Revenu et l'univers GLOBAL .

Nous noterons:

pour tout indice i , i = 1 - - - I
pour tout facteur manipulé f , f = 1 - - F
pour toute modalité l du facteur, l = 1 - - L

U_{fl} est l'univers issu de la manipulation de la modalité l du facteur f .

Ainsi:

U_{f1}, U_{f2} et U_{f3} sont les univers issus de la manipulation des modalités 1, 2 et 3 du facteur f (pour le facteur date, il n'y a que deux modalités).

Notons:

U_G l'univers GLOBAL, aucun facteur n'a été manipulé .

U_{ifl}, le vecteur composé des mesures de concurrence effectuées avec l'indice i sur l'univers U_{fl}.

Nous pouvons alors exprimer l'effet du facteur f à partir de la distance des univers U_{f1}, U_{f2} et U_{f3} à l'univers U_G .

$$a_{ifl} = f^{-1}(d(a_{ifl})) = corr\ (U_{ifl} , U_{iG})$$

a_{ifl} est la corrélation entre les vecteurs de mesure U_{ifl} et U_{iG} , c'est une fonction inverse (f^{-1}, non spécifiée ici) de la distance entre les deux univers correspondants.

Les corrélations a_{if1}, a_{if2} et a_{if3} permettent de mesurer l'effet de la manipulation du facteur f. Si les a_{if3} sont proches de 1, les distances entre les univers issus de la manipulation du facteur f et l'univers GLOBAL sont faibles, nous pouvons alors en déduire que l'effet du facteur f est lui même faible.

Nous pouvons ainsi calculer pour chaque indice i et chaque facteur f, la similarité moyenne entre ses modalités et la modalité Global. Pour un facteur f à trois modalités. Soit:

$$a_{if.} = 1/3 \sum_{l=1}^{3} a_{ifl}$$

$a_{if.}$ est l'écartement moyen des modalités du facteur f au facteur Global. C'est une mesure de l'effet de ce facteur sur l'indice i.

D'autre part,nous pouvons calculer:

$$a_{.f.} = 1/I \sum_{i=1}^{I} a_{if.}$$

$a_{.f.}$ est la similarité moyenne des mesures de la concurrence sur les univers issus de la manipulation du facteur f.
$a_{.f.}$ mesure l'effet du facteur f, tous indices confondus .

De façon similaire, nous calculerons:

$$a_{i..} = 1/F \sum_{f=1}^{F} a_{if.}$$

$a_{i..}$ est la similarité moyenne, pour le seul indice i, de l'ensemble des univers issus de la manipulation de tous les facteurs.
$a_{i..}$ représente la capacité de l'indice à être influencé par les facteurs. Considéré relativement aux autres indices, $a_{i..}$ est une mesure de la sensibilité de l'indice i à la manipulation des différents facteurs.

Enfin, nous calculerons:

$$a_{...} = 1/(I.F.3) \sum_{i=1}^{I} \sum_{f=1}^{F} \sum_{l=1}^{3} a_{lif}$$

$a_{...}$ est un indicateur de cohésion moyenne avec l'univers Global, sur le marché étudié. Il servira notamment de base de comparaison aux mesures $a_{.f.}$ et $a_{i..}$
$a_{...}$ mesure la stabilité du concept de "structure des marchés" étudié à partir de la concurrence entre les marques (données de comportement).

Nous retiendrons que :

$(a_{.f.} - a_{...})$ est une mesure de l'effet du facteur f sur la mesure de la concurence, tous indices confondus .
$(a_{i..} - a_{...})$ mesure la propension de l'indice i à être influencé, tous facteurs confondus.
$(a_{if.} - a_{i..})$ mesure l'importance de l'effet du facteur f sur l'indice i
$(a_{if.} - a_{.f.})$ mesure la propension de l'indice i à être plus ou moins influencé que la moyenne des indices, par le facteur f

L'approche de la mesure des facteurs que nous venons de présenter a pour caractéristique de dépendre des deux éléments suivants :

1- L'effet du facteur manipulé, ce que l'on cherche à étudier ici mais aussi
2- L'effet de la taille de l'échantillon (en terme du nombre d'actes d'achats) car la distance d'un univers particulier à l'univers GLOBAL est directement fonction de la taille de l'échantillon de cet univers (plus la taille de l'échantillon relatif à un univers particulier est importante, plus il comporte d'observations communes avec l'univers GLOBAL et donc plus leur corrélation est importante)..

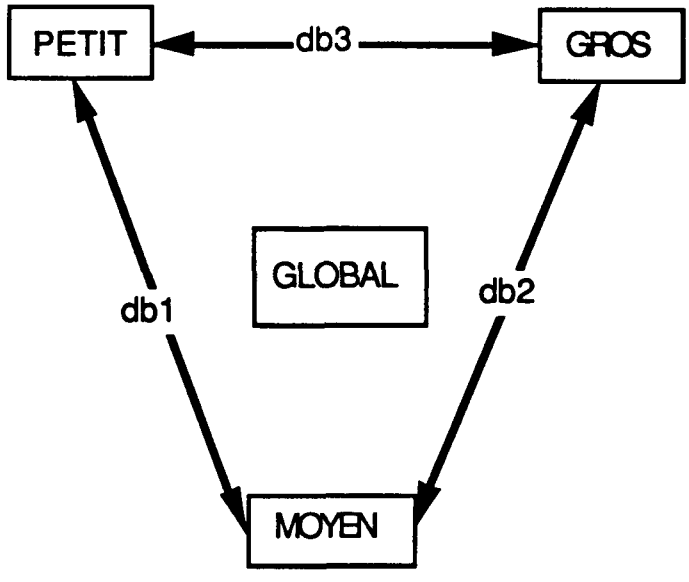
La taille de l'échantillon de chaque univers est (environ) égale au tiers de celle de l'univers GLOBAL, puisque nous avons effectué le découpage des variables manipulées

de manière à avoir le même nombre d'actes d'achats dans chaque Univers. Seuls les univers issus de la manipulation du facteur CSP ne suivent pas cette règle (nous avons expliqué pourquoi au chapitre IV, section 1). Ces univers comportant moins du tiers de l'échantillon, leur corrélation avec l'Univers Global en sera diminuée, ce qui augmente l'effet apparent du facteur CSP. Nous en tiendrons compte lors des interprétations.

SS2 ETUDE A PARTIR DES DISTANCES ENTRE LES UNIVERS ISSUS DE LA MANIPULATION DES FACTEURS

Nous pouvons étudier l'effet d'un facteur à partir du lien entre chacun des univers issus de sa manipulation. En effet, s'il est important d'étudier le pouvoir *'séparateur'* d'un facteur (à partir des valeurs des distances aifl) il est aussi intéressant d'étudier la similarité entre les univers issus de la manipulation d'un facteur. Le graphique suivant permet de visualiser ce point de vue .

GRAPHIQUE 3B EFFET DES FACTEURS : DISTANCES ENTRE LES UNIVERS



Sur le graphique 3B, d_{b1}, d_{b2} et d_{b3} sont les distances entre les trois modalités du facteur manipulé. Ces distances sont une fonction inverse (non précisée) de la corrélation entre les univers REVG,REVM et REVP pris deux à deux.
Pour l'indice i et le facteur f nous noterons:

$b_{if1} = f^{-1}(d(b_{if1})) = corr\ (U_{if1}\ ,\ U_{if2})$
b_{if1} est la corrélation entre U_{if1} et U_{if2} . C'est une fonction inverse de la distance entre les deux univers.

De même :

b_{if2} est la corrélation entre U_{if1} et U_{if3} .
b_{if3} est la corrélation entre U_{if2} et U_{if3} .

b_{if1}, b_{if2} et b_{if3} permettent de mesurer l'effet de la manipulation du facteur f. Si les bifl sont proches de 1, c'est que les distances entre les univers issus de la manipulation du facteur f sont faibles. Nous pouvons alors en déduire que l'effet du facteur f est lui même faible.

Nous calculerons ainsi, pour chaque indice i et chaque facteur f, la similarité moyenne entre les trois univers issus de la manipulation du facteur f. Pour un facteur f à trois modalités.
Notons:

$$b_{if.} = 1/3 \sum_{l=1}^{3} b_{ifl}$$

$b_{if.}$ est l'écartement moyen des modalités du facteur f, prises deux à deux . C'est une mesure de l'effet de ce facteur sur l'indice i .

D'autre part, nous pouvons calculer:

$$b_{.f.} = 1/I \sum_{i=1}^{I} b_{if.}$$

$b_{.f.}$ est la similarité moyenne des mesures de la concurrence sur les univers issus de la manipulation du facteur f.
$b_{.f.}$ mesure l'effet du facteur f, tous indices confondus.

De façon similaire, nous pouvons calculer:

$$b_{i..} = 1/F \sum_{f=1}^{F} b_{if.}$$

$b_{i..}$ est la similarité moyenne, pour le seul indice i, entre l'ensemble des univers issus de la manipulation de tous les facteurs.
bi.. représente la capacité de l'indice à être influencé par les facteurs manipulés. Considéré relativement aux autres indices, $b_{i..}$ est une mesure de la sensibilité de l'indice.

Enfin, nous calculerons:

$$b_{...} = 1/(I.F.3) \sum_{i=}^{I} \sum_{f=1}^{F} \sum_{l=1}^{3} b_{lif}$$

$b_{...}$ est un indicateur de la cohésion moyenne entre les univers d'un même facteur, pour tous les facteurs, sur le marché étudié . Il servira notamment de base de comparaison des mesures $b_{.f.}$ et $b_{i..}$.

$b_{...}$ mesure la stabilité du concept de "structure des marchés" étudié à partir de la concurrence entre les marques (données de comportement).

Nous retiendrons que:

$(b_{f..} - b_{...})$ est une mesure de l'effet du facteur f sur la mesure de la concurrence, tous indices confondus.
$(b_{i..} - b_{...})$ mesure la propension de l'indice i à être influencé, tous facteurs confondus.
$(b_{if.} - b_{i..})$ mesure l'importance de l'effet du facteur f sur l'indice i.
$(b_{if.} - b_{.f.})$ mesure la propension de l'indice à être plus ou moins influencé que la moyenne des indices, par le facteur f.

SECTION 2 ANALYSE COMPARATIVE DE L'EFFET DES FACTEURS

Dans cette section, nous allons comparer l'effet global des facteurs manipulés, tous indices confondus. Pour cela, nous utiliserons la double logique correspondant aux sous-sections précédentes.

SS1 COMPARAISON DE L'EFFET DES FACTEURS A PARTIR DES SIMILARITES MOYENNES AVEC LA MODALITE GLOBAL (les a.f.)

Nous allons comparer ici, tous indices confondus, l'effet des facteurs sur la similarité à la modalité GLOBAL. Pour cela, nous comparerons les valeurs des $a_{.f}$, lorsque f varie sur les facteurs, en utilisant le modèle suivant :

$$ACORR = f(Manip_A)$$

$Manip_A$ est la variable explicative du modèle . C'est une variable qualitative dont chacune des modalités correspond à la manipulation d'un facteur particulier. La codification de cette variable figure dans le tableau suivant

TABLEAU 29A CODIFICATION DU FACTEUR MANIPA

F1	F2	Facteur Manipulé	CODE	$Manip_A$
REVENU	GLOBAL	REVENU	1	aREV
GLOBAL	REVENU	REVENU	1	aREV
CSP	GLOBAL	CSP	2	aCSP
GLOBAL	CSP	CSP	2	aCSP
MAGASIN	GLOBAL	MAGASIN	3	aMAG
GLOBAL	MAGASIN	MAGASIN	3	aMAG
TAYF	GLOBAL	TAYF	4	aTAY
GLOBAL	TAYF	TAYF	4	aTAY
QUANTITE	GLOBAL	QUANTITE	5	aCON
GLOBAL	QUANTITE	QUANTITE	5	aCON
DATE	GLOBAL	DATE	6	aDAT
GLOBAL	DATE	DATE	6	aDAT

Rappelons que F1 et F2 correspondent aux variables indicatrices de la manipulation des facteurs sur les univers de calcul des deux indices corrélés. D'autre part, sur le marché SHAMP, les facteurs REV et CSP n'ont pas été manipulés.

L'analyse de variance à un facteur est la technique qui convient le plus naturellement à la résolution du problème posé. De plus, elle est très facile à interpréter.

Pour supprimer l'effet de l'indice sur chaque corrélation, nous n'avons retenu que les corrélations intra-indice. Ceci revenait à poser la condition I1 = I2 pour la sélection des observations participant au modèle.

Les résultats pour les quatre marchés figurent dans le tableau suivant.

TABLEAU 30A EFFET DES FACTEURS MANIPULES (a.f.)

	HUILE	BOISSON	PARFUM	SHAMP
NOBS	185	186	170	74
DDL	5	5	5	3
FISHER	7.74	5.37	5.58	0.74
PROB	0.0001	0.0001	0.0001	0.46

TEST DE MOYENNE sur les a.f. α=10%

	.46 B	.40 BC	.58 AB	-
CSP	.31 C	.33 C	.35 C	-
MAG	.30 C	.32 C	.36 AB	.36
TAYF	.55 AB	.59 A	.56 BC	.38
CONS	.54 AB	.49 A	.43 A	.49
DAT	.66 A	.46 A B	.67	.43
MOYE (a...)	.46	.43	.48	.41

Dans les lignes de ce tableau figurent respectivement :

NOBS : Nombre d'observations dans l'analyse
DDL : Nombre de degrés de liberté du facteur étudié
Fisher : Valeur du Fisher
Prob : Signification du test de l'hypothèse H0: $a_{.f.} = a_{.f'.}$ pour tous les facteurs f, f'.
H0 signifie que les distances moyennes à l'univers global sont toutes égales c'est à dire que tous les facteurs ont le même effet sur les relations de concurrence. La valeur de Prob correspond au risque que l'on prend à rejeter l'hypothèse H0. Plus cette valeur est faible et plus les facteurs ont des effets différenciés sur les relations de concurrence.
Nous trouvons ensuite, dans ce tableau, les valeurs de a.f. pour chacun des facteurs ainsi que la valeur moyenne a... (notée MOYE). A chacune de ces valeurs est jointe une lettre (A pour la moyenne la plus forte, B pour les valeurs significativement inférieures à A, etc). Cette lettre représente le résultat du test de différence entre les moyennes. Deux facteurs avec la même lettre n'ont pas d'effets significativement différents au seuil de risque alpha de 5% .

Sur les trois marchés HUILE, BOISSON et PARFUM, nous observons de fortes différences entre l'effet des facteurs manipulés (les a.f.). Ceci signifie que tous les facteurs n'ont pas le même effet sur la structure du marché.

ANALYSE DES ECARTS

Le tableau suivant fournit les écarts ($a_{.f.} - a_{...}$), pour chaque facteur, sur chaque marché. Une valeur positive signifie que le facteur a moins d'effet que la moyenne et inversement (les différences significatives au seuil de 5% sont notées par un *).

TABLEAU 31A : EFFETS DIFFERENCIES DES FACTEURS : ($a_{.f.} - a_{...}$)

	HUILE	BOISSON	PARFUM	SHAMP
REV	.00	-.03	.10*	
CSP	-.15*	-.10*	-.13*	
MAG	-.16*	-.11*	-.12*	-.05
TAY	.09*	.16*	.08	-.03
CONS	.08	.06	-.05	.08
DATE	.20*	.03	.19*	.02

A partir de ces résultats, nous pouvons faire les commentaires suivants pour chacun des facteurs.

REVENU

Ce facteur, sur les trois marchés où il a été manipulé, a un effet "moyen" (relativement aux autres facteurs) .Nous retiendrons que la structure des relations entre les marques se modifie moyennement (relativement aux autres facteurs) quand elle est étudiée à partir des achats de panélistes appartenant aux trois catégories de revenus.

CSP

Sur les trois marchés HUILE, BOISSON, PARFUM, l'effet de ce facteur est plus fort que la moyenne (sur trois marchés cet effet est significativement plus fort) que la moyenne). Cependant, avant d'en conclure que l'effet du facteur CSP est fort, il est nécessaire de tenir compte de la remarque faite précédemment (les trois univers CSP5, CSP6 et CSP7 comportent moins d'actes d'achats que les autres univers). L'étude à partir des $b_{.f.}$ nous permettra d'avoir des résultats qui ne tiennent pas compte de ces différences d'effectifs entre les Univers.

MAGASIN

C'est le facteur qui se révèle comme ayant le plus d'effet. Sur trois marchés, son effet est significativement plus fort que la moyenne des effets.

TAILLE DU FOYER

Sur tous les marchés, sauf SHAMP, ce facteur apparait comme ayant un effet inférieur aux autres.

QUANTITE CONSOMMEE

Ce facteur a un effet moyen sur les quatre marchés. La quantité consommée ne parait pas jouer très fortement sur la similarité des univers à l'univers GLOBAL.

DATE

L'étude de l'influence de ce facteur est d'importance, nous l'avons souligné dans l'introduction de ce chapitre. Apparemment, le facteur DATE est celui qui influence le moins la structure d'un marché.

Cependant, ce facteur n'a été découpé qu'en deux modalités (les deux périodes d'observations), les deux univers correspondant comportent plus d'actes d'achats

que les autres (qui sont le résultat d'un découpage en trois classes), il est donc logique que les similarités de ces univers avec l'univers GLOBAL soient plus fortes. C'est pourquoi l'analyse à partir des b.f. sera indispensable pour porter un jugement définitif concernant l'importance relative de ce facteur.

SS2 COMPARAISON DE L'EFFET DES FACTEURS A PARTIR DES SIMILARITES MOYENNES ENTRE LES UNIVERS (les $b_{.f.}$)

Nous allons comparer ici, tous indices confondus, l'effet des facteurs sur la similarité entre les univers issus de la manipulation d'un même facteur. Pour cela, nous comparerons les valeurs des $b_{.f.}$, lorsque f varie sur les facteurs, en utilisant le modèle suivant :

$$ACORR = f(MANIP_B)$$

$MANIP_B$ est la variable explicative du modèle. C'est une variable qualitative dont chacune des modalités correspond à la manipulation d'un facteur particulier. La codification de cette variable figure dans le tableau suivant.

TABLEAU 29B CODIFICATION DU FACTEUR MANIPB

F1	F2	Facteur Manipulé	CODE	$Manip_A$
REVENU	REVENU	REVENU	1	bREV
CSP	CSP	CSP	2	bCSP
MAGASIN	MAGASIN	MAGASIN	3	bMAG
TAYF	TAYF	TAYF	4	bTAY
QUANTITE	QUANTITE	QUANTITE	5	bCON
DATE	DATE	DATE	6	bDAT

Il s'agit d'analyser des corrélations d'indices calculés sur des univers issus de la manipulation d'un facteur particulier, c'est à dire des *"corrélations intra-facteur"*.

Pour supprimer l'effet de l'indice sur chaque corrélation, nous n'avons retenu dans l'analyse que les corrélations intra-indice. Ceci revenait à poser la condition I1 = I2 pour la sélection des observations participant au modèle. L'analyse porte finalement sur des corrélations intra-facteur et intra-indice (la case 2 du tableau 3, au chapitre IV).

Les résultats pour les quatre marchés figurent dans le tableau suivant.

TABLEAU 30B EFFET DES FACTEURS MANIPULES ($b_{.f.}$)

	HUILE	BOISSON	PARFUM	SHAMP
NOBS	173	176	160	70
DDL	5	5	5	3
FISHER	1.01	4.7	3.46	0.40
PROB	0.41	0.0005	0.0005	0.75

TEST DE MOYENNE sur les b.f. α=10%

REV	.35 AB	.29 AB	.36 AB	-
CSP	.30 AB	.24 ABC	.26 BC	-
MAG.	27 B	.17 C	.22 C	.21 A
TAYF	.34 AB	.35 A	.39 A	.19 A
CONS	.27 B	.20 BC	.22 C	.25 A
DAT.	43 A	.33 A	.39 A	.20 A
MOYE (b...)	.31	.26	.30	.21

La composition de ce tableau est la même que celle du tableau 30A précédent.
L'hypothèse testée est ici:

H0 $b_{.f.} = b_{.f'.}$ pour tous facteurs f, f'

H0 signifie que les similarités moyennes entre les univers issus de la manipulation des facteurs sont toutes égales, ce qui signifie que tous les facteurs ont le même effet sur les relations de concurrence.
La valeur de Prob correspond au risque que l'on prend à rejeter H0. Plus cette valeur est faible et plus les facteurs ont des effets différenciés sur les relations de concurrence.
Nous trouvons ensuite les valeurs de $b_{.f.}$ pour chacun des facteurs ainsi que la valeur moyenne $b_{...}$ (notée MOYE). Deux facteurs avec la même lettre n'ont pas d'effets significativement différents au seuil de risque alpha de 5%.

RESULTATS

Le niveau moyen des facteurs ($b_{.f.}$) est de l'ordre de .30, témoignant que globalement (tous indices confondus) la manipulation des facteurs a un effet important sur le calcul des indices de concurrence.
Sur les trois marchés HUILE, BOISSON et PARFUM, nous observons de fortes différences entre les effets des facteurs manipulés (les $b_{.f.}$). Tous les facteurs n'ont pas le même effet sur la structure du marché. La mesure de la concurrence (à partir de données de comportement) peut ainsi être qualifiée de mesure sensible.

ANALYSE DES ECARTS

Le tableau suivant présente les écarts ($b_{.f.} - b_{...}$), pour chaque facteur f, sur chaque marché. Une valeur positive signifie que le facteur a moins d'effet que la moyenne (ce qui ne signifie pas qu'il n'a pas d'effet) et inversement, une valeur négative signifie que le

facteur a plus d'effet que la moyenne des facteurs. Les différences significatives au seuil de 5% sont notées par un astérisque.

TABLEAU 31B EFFETS DIFFERENCIES DES FACTEURS ($b_{.f.} - b_{...}$)

	HUILE	BOISSON	PARFUM	SHAMP
REV	.04	.03	.06	-
CSP	-.01	-.02	-.04	-
MAG	-.04	-.09*	-.08	.02
TAY	.03	.09*	.09*	-.02
CONS	-.04	-.06	-.08	.04
DATE	.12	.07	.09*	.01

A partir de ce tableau, nous pouvons faire les commentaires suivants pour chacun des facteurs.

REVENU:

Les univers issus de la manipulation de ce facteur sont plus similaires que la moyenne. L'effet *"séparateur"* de l'indice est plus faible que la moyenne.

CSP:

Sur les trois marchés HUILE BOISSON PARFUM, l'effet de ce facteur est maintenant moyen, ce qui confirme nos réflexions précédentes. Nous considérerons que, globalement, ce facteur (pour les CSP étudiées à savoir cadres moyens et supérieurs, employés et retraités) a moins d'effet que la moyenne .

MAGASIN:

L'effet de ce facteur est nettement plus prononcé que celui des autres. Le MAGASIN apparait comme le facteur ayant le plus d'effet.

TAILLE DU FOYER:

L'effet de ce facteur est plutôt faible (sauf sur le marché des shampooings), il fait partie des facteurs qui ont le moins d'influence.

QUANTITE CONSOMMEE:

Ce facteur a un effet important quand dans l'analyse à partir des $b_{.f.}$, excepté sur le marché SHAMP.

DATE:

Ce facteur confirme ici sa faible influence relative (en moyenne sur tous les indices). C'est un résultat important puisque l'hypothèse de stabilité des préférences dans le temps, dont les achats observés sont supposés dépendre (voir Bass 1974), est à la base des nombreuses modélisations (voir Massy, Montgomery et Morrison 1969). Il faut néanmoins relativiser ce résultat par le fait que la longueur des périodes d'observation n'est que de six mois. Qu'en serait il si ces périodes étaient plus longues (par exemple un an) ?

Nous retiendrons des deux séries d'analyses que nous venons d'effectuer (dans les paragraphes SS2-1 et SS2-2), les points suivants :

- à partir des niveaux moyens de corrélations observés (les $a_{...}$ et $b_{...}$) nous pouvons constater que les indices étudiés sont "fortement" influencés par la manipulation des facteurs .
- les facteurs manipulés présentent des niveaux d'influence significativement différents sur la structure des relations entre marques.
- l'influence de chaque facteur est assez similaire d'un marché à l'autre, excepté pour le marché des Shampooings .
- le facteur DATE, quel que soit le marché étudié, apparait comme un facteur dont le niveau d'influence est moins fort que les autres facteurs (sans pour autant être négligeable, comme nous l'avons vu au chapitre précédent quand nous avons effectué des test-retest).
- sur les trois marchés, HUILE, BOISSON et PARFUM les facteurs REVENU et TAILLE du FOYER apparaissent comme ayant une faible importance.
- le facteur MAGASIN est celui qui a l'effet le plus fort sur la structure des marchés étudiés, ce qui constitue un résultat très important.

SECTION 3 ANALYSE DE L'EFFET GLOBAL DES FACTEURS SUR CHACUN DES INDICES ET ETUDE DE LA SENSIBILITE DES INDICES

Toutes les analyses précédentes ne sont valables que pour les indices considérés dans leur ensemble. Or, nous avons montré, dans le chapitre précédent, que les indices ne sont pas tous liés et qu'il existe des familles bien différenciées d'indices. Lors de l'étude de la Fiabilité, à la section 2, nous avons vu que les indices sont influencés de manière très différente par la manipulation du facteur Date. Les indices bruts de transferts entre marques, IND et ACT, se sont révélés être beaucoup moins influencés que les autres par la manipulation de ce facteur. Nous pouvons supposer qu'il en va de même avec la manipulation des autres facteurs. Ces deux indices en données brutes, nous serviront donc de base comparative pour les autres indices étudiés.
Il est naturel d'envisager :

- que tous les indices ne présentent pas les mêmes niveaux de sensibilité à la manipulation des facteurs
- que tous les indices ne réagissent pas à l'identique à la manipulation de chacun des facteurs

A ces deux points correspondent les deux parties de cette section. Tous les calculs qui suivent seront effectués uniquement à partir des distances entre les univers (les bifl).

SS1 ETUDE DE L'EFFET GLOBAL DES FACTEURS SUR LES INDICES

Nous allons étudier et comparer l'effet global des facteurs sur chacun des indices, c'est à dire la sensibilité des indices à la manipulation de ces facteurs.

Rappelons, que $b_{i..}$ représente la similarité moyenne des vecteurs de concurrence, calculés pour un indice i particulier, sur l'ensemble des univers issus de la manipulation des différents facteurs.

$$b_{i..} = 1/(F.3) \sum_{f=1}^{F} \sum_{l=1}^{3} b_{ifl} \quad \text{(pour un facteur à trois modalités)}$$

$b_{i..}$ est une mesure de la propension de l'indice a être affecté par la manipulation des facteurs: sensibilité de l'indice.

Nous pouvons apprécier cette propension par la différence :

$$b_{i..} - b_{...}$$

Pour tester si les indices présentent des niveaux moyens de sensibilité différents, nous avons utilisé le modèle suivant (qui sera estimé par Analyse de la Variance à un facteur) :

ACORR = f (INDICE)

INDICE est une variable qualitative dont chacune des modalités correspond à l'indice utilisé pour effectuer la mesure. Les observations qui participent à l'analyse sont des corrélations intra-indice (condition I1=I2).

L'hypothèse testée H0 est la suivante :
"les similarités moyennes inter-univers sont identiques d'un indice à l'autre", c'est à dire:

$b_{i..} = b_{...}$, pour tout indice i.

Les résultats, marché par marché, figurent à l'Annexe 20. Le tableau suivant résume les résultats obtenus pour les quatre marchés.

TABLEAU 32 SENSIBILITE MOYENNE DES INDICES A LA MANIPULATION DES FACTEURS : CORRELATIONS MOYENNES ENTRE LES UNIVERS POUR CHAQUE INDICE ($b_{i..}$)

INDICE	HUILE	BOISSON	PARFUM	SHAMP
FRASA	.40	.20	.14	.23
FRASB	.41	.21	.20	.22
LEM	.66	.54	.58	.43
MERA	.18	.23	-	.12
MERR	.22	.22	.29	.10
SABA	.23	.16	.10	.15
TOTT	.29	.13	.32	-
TRTO	.31	.29	.33	-
RAO	.19	.19	.26	.23
ELRO	.25	.36	.36	-
TREL	.30	.29	.38	-
IND	.83	.74	.74	.65
ACT	.90	.72	.84	.68
MOYE b...	.31	.26	.30	.21

Sur les quatre marchés, nous pouvons constater des différences très significatives des niveaux de similarité, d'un indice à l'autre (voir Annexe 20 pour les tests de moyennes correspondants). Les indices présentent donc des niveaux de sensibilité très différents à la manipulation des facteurs.
Nous retiendrons les éléments suivants:

1 Les données de transfert *"bruts"* (indices IND et ACT) sont beaucoup moins sensibles à la manipulation des facteurs que tous les autres indices. Les patterns de transferts entres les marques sont donc très stables. L'indice LEM, par la nature de sa formule de calcul (proche des données brutes), est d'ailleurs le seul indice qui présente un degré de stabilité acceptable. Par contre, les indices SABA et RAO, calculés sur la base

des mêmes données de transfert, se révèlent beaucoup plus sensibles. Nous attribuons ce résultat, nous l'avons déjà souligné, à la forme multiplicative de leur dénominateur.

2 Le fort niveau de sensibilité de certains indices (comme MERA, MERR, SABA, RAO) peut s'expliquer de deux manières différentes: soit ces indices sont très sensibles aux facteurs manipulés, soit ces mêmes indices sont sensibles à la diminution de la taille de l'échantillon sur lequel ils ont été calculés (puisqu'ils ont été calculés sur des univers d'actes d'achat représentant le tiers de la taille de l'échantillon total). Pour répondre à cette interrogation, il aurait fallu partager l'échantillon, par le hasard, en trois et effectuer les mêmes calculs que précédemment. Nous aurions alors observé l'effet dû à la "seule" diminution de la taille de l'échantillon. Pour des raisons de disponibilité des bases SCAN 5000, nous n'avons pas pu effectuer ces vérification.

3 Les niveaux relatifs de sensibilité de chaque indice sont similaires d'un marché à l'autre.

A partir des tableaux de l'annexe 20, nous avons construit le tableau 33 des écarts (bi.. - b...) qui permet d'apprécier les niveaux de sensibilité des différents indices. Un écart négatif signifie que l'indice correspondant, sur le marché correspondant, est plus sensible à la manipulation des facteurs que la moyenne des indices.

TABLEAU 33 : INFLUENCE DIFFERENCIEE DES FACTEURS SUR LES INDICES
($b_{i..} - b_{...}$)

INDICE	HUILE	BOISSON	PARFUM	SHAMP
FRASA	09	-.06	-.16	.02
FRASB	.10	-.05	-.10	.01
LEM	.35	.28	.28	.22
MERA	-.13	-.03	-	-.09
MERR	-.09	-.02	-.01	-.11
SABA	-.08	-.10	-.20	-.06
TOTT	-.02	-.07	.02	-
TRTO	.00	.03	-.03	-
RAO	-.12	-.07	-.04	.02
ELRO	-.06	.10	.06	-
TREL	-.01	.03	.08	-

LEM est l'indice le moins sensible, quel que soit le marché. Il est le seul indice qui présente un niveau de sensibilité correct. Nous trouvons ensuite les indices TOTT, TRTO, ELRO et TREL dont les niveaux de sensibilité sont moyens. Tous les autres indices sont très sensibles à la manipulation des facteurs, notamment MERA, MERR et SABA. L'indice SABA est l'indice le plus sensible.

SS2 ETUDE DE L'EFFET DE CHAQUE FACTEUR SUR CHAQUE INDICE

Nous allons mener ici le même type d'analyse que dans la section précédente, mais cette fois pour chacun des indices. L'objectif est d'établir, pour chaque indice, le diagnostic des facteurs qui l'influencent le plus. Le même modèle d'analyse de variance que dans la section 2 est utilisé ici à savoir (pour chaque indice):

ACORR = f(MANIPB)

MANIPB est la variable explicative du modèle. C'est une variable qualitative dont chacune des modalités correspond à la manipulation d'un facteur particulier. La codification de cette variable figure dans le tableau suivant.

TABLEAU 34 CODIFICATION DU FACTEUR "MANIPB"

F1	F2	FACTEUR MANIPULE	CODE	MANIPB
REVENU	REVENU	REVENU	1	bREV
CSP	CSP	CSP	2	bCSP
MAGASIN	MAGASIN	MAGASIN	3	bMAG
TAYF	TAYF	TAYF	4	bTAY
QUANTITE	QUANTITE	QUANTITE	5	bCON
DATE	DATE	DATE	6	bDAT

Les observations retenues dans cette analyse sont des corrélations intra-facteur (F1 = F2). L'hypothèse H0 testée dans le modèle d'analyse de variance est la suivante: "l'indice i est affecté à l'identique par chacun des facteurs", soit:

H0 : $b_{if.} = b_{i..}$ pour tout facteur f

Les résultats figurent dans les tableaux 35 (α à π) suivants, pour les quatre marchés, ainsi qu'à l'annexe 21. Nous constatons que pour la plupart des analyses, il n'y a pas d'effet significatif du facteur. Ce résultat suggère que les facteurs ont des effets non significativement différents, sur chacun des indices étudiés. A notre avis, il n'en est rien. En effet, pour mener ces analyses, nous avons dû réduire considérablement les tailles d'échantillons. Chaque analyse, ne compte que 17 observations (cinq facteurs à trois modalités et un facteur à deux modalités). Le test de moyennes est alors très rigoureux et il y a peu de chances d'observer des effets significatifs (plus le nombre d'observations est faible et plus le test de différence de deux moyennes devient rigoureux).

A défaut d'utiliser la statistique classique (qui repose sur l'hypothèse de normalité des distributions des corrélations), nous aurions pu utiliser ici un test non-paramétrique. En particulier, le test de Kruskal-Wallis (voir Lecoutre et Tassi 1987, chapitre 8) qui consiste à comparer k échantillons serait bien adapté à ce problème. Ne disposant pas du programme statistique nécessaire, nous n'avons pas utilisé ces tests.

Les tableaux 35 suivants présentent les niveaux moyens des corrélations intra-facteur, pour chaque indice et chaque marché (les bif.).

TABLEAU 35α MARCHE HUILE : TABLEAU DES bif.

	REV	CSP	MAG	TAY	CONS	DAT	MOYE ($b_{i..}$)
FRASA	.50	.14	.51	.58	.25	-	.40
FRASB	.54	.28	.45	.52	.28	-	.41
LEM	.73	.62	.27	.79	.81	.84	.66
MERA	.08	.22	.32	.12	.18	-	.18
MERR	.17	.25	.34	.19	.17	.17	.22
SABA	.30	.33	.08	.23	.12	.50	.23
TOTT	.29	.27	.24	.33	.21	.70	.29
TRTO	.46	.43	.23	.18	.32	.15	.31
RAO	.26	.20	.17	.13	.06	.49	.18
ELRO	.22	.19	.20	.35	.23	.43	.24
TREL	.30	.41	.18	.29	.36	.12	.30
IND	.84	.78	.67	.93	.91	.93	.83
ACT	.98	.97	.71	.96	.86	.90	.90

TABLEAU 35ß MARCHE BOISSON : TABLEAU DES bif.

	REV	CSP	MAG	TAY	CONS	DAT	MOYE ($b_{i..}$)
FRASA	.08	.30	.11	.28	.17	.38	.20
FRASB	.26	.12	.20	.30	.13	.43	.21
LEM	.66	.55	.37	.63	.40	.78	.54
MERA	.16	.16	.20	.47	.23	.03	.23
MERR	.30	.18	.20	.21	.23	.19	.22
SABA	.18	.22	.07	.14	.12	.32	.16
TOTT	.18	.11	.19	.06	.04	.38	.13
TRTO	.33	.36	.11	.40	.30	.14	.29
RAO	.23	.20	.11	.22	.11	.40	.19
ELRO	.51	.24	.24	.65	.12	.45	.36
TREL	.32	.23	.11	.50	.32	.13	.29
IND	.86	.76	.50	.80	.74	.92	.74
ACT	.84	.76	.44	.73	.74	92	.72

TABLEAU 35Γ MARCHE PARFUM : TABLEAU DES bif.

	REV	CSP	MAG	TAY	CONS	DAT	MOYE ($b_{i..}$)
FRASA	.21	.16	.17	.08	.17	.08	.14
FRASB	.12	.26	.23	.32	.12	.02	.20
LEM	.72	.50	.36	.74	.51	.77	.58
MERR	.26	.31	.32	.23	.35	.28	.29
SABA	.15	.16	.06	.11	.04	.01	.10
TOTT	.37	.27	.29	.53	.15	.31	.32
TRTO	.41	.30	.11	.28	.36	.87	.33
RAO	.29	.17	.21	.34	.29	.29	.26
ELRO	.51	.24	.24	.65	.12	.45	.36
TREL	.51	.24	.24	.65	.12	.85	.38
IND	.88	.75	.38	.92	.73	.93	.74
ACT	.86	.69	.29	.86	-	.92	.84

TABLEAU 35π MARCHE SHAMP : TABLEAU DES bif.

	MAG	TAY	CONS	DAT	MOYE ($b_{i..}$)
FRASA	.18	.13	.36	.31	.23
FRASB	.23	.13	.33	.17	.22
LEM	.35	.43	.47	.52	.43
MERA	.20	.09	.09	.01	.12
MERR	.17	.08	.08	.01	.10
SABA	.11	.23	.10	.18	.15
RAO	.19	.23	.30	.17	.23
IND	.49	.67	.73	.81	.65
ACT	.54	.74	.73	.82	.68

Le résultat le plus marquant dans ces tableaux est la différence de sensibilité à la manipulation des facteurs, entre les deux indices IND et ACT d'une part et les autres indices d'autre part. Nous constatons que les données brutes de transferts entre marques sont très peu sensibles à la manipulation des facteurs, à l'exception du facteur MAGASIN qui se révèle être nettement plus influent que les autres. Il n'apparait pas, par contre, de différence de sensibilité entre les indices IND (changements en individus) et ACT (changements en actes d'achats).

Ceci confirme ce que nous avons vu au chapitre V, à savoir que la sensibilité d'indices comme SABA, RAO, TOTT, TRTO, ELRO et TREL n'est pas due (excepté pour le facteur MAGASIN) à la sensibilité des données de transfert utilisées pour les calculer, mais à leur mode de calcul.

A défaut d'utiliser des tests non-paramétriques et afin de mieux visualiser la structure des effets relatifs des facteurs manipulés sur chaque indice, nous avons procédé de la manière suivante (tout à fait empirique):

- nous avons calculé les termes ($b_{if.} - b_{i..}$)
- nous avons codifié la valeur de ces termes de différences de la manière suivante :

++ lorsque la différence est supérieure à 10% de la valeur de $b_{i..}$
+ lorsque la différence est comprise entre 5% et 10% de la valeur de $b_{i..}$
= lorsque la différence est comprise entre -4 et +4% de la valeur de $b_{i..}$
- lorsque la différence est comprise entre -5% et -10% de la valeur de $b_{i..}$
-- lorsque la différence est inférieure à -10% de la valeur de $b_{i..}$

Malgré son caractère arbitraire, une telle approche nous permet de mieux visualiser la structure des effets des facteurs sur chaque indice. Nous allons maintenant présenter les résultats obtenus pour chaque indice.

INDICE FRASA :

Les résultats des tableaux précédents montrent que les similarités entre les univers calculées à partir de cet indice sont nettement plus faibles que la moyenne. L'analyse des écarts nous fournit le tableau suivant :

TABLEAU 36 STRUCTURE DES EFFETS SUR L'INDICE FRASA

	HUILE	BOISSON	PARFUM	SHAMP
REV	++	--	+	
CSP	--	++	=	
MAG	++	--	=	-
TAY	++	++	-	--
CONS	--	-	=	+
DATE		++	-	++

Nous observons que cet indice est sujet à des effets variés de la part des facteurs. Il n'y a pas de structure stable de l'effet des facteurs d'un marché à l'autre. D'autre part, le niveau des corrélations est trop faible, avec cet indice, pour que nous puissions effectuer des commentaires précis. L'indice apparait surtout comme étant *"volatile"*, ce qui est certainement la conséquence de la diminution de la taille de l'échantillon des univers de calcul.

INDICE FRASB :

Sur les tableaux 35, nous pouvons observer que l'indice est assez fortement influencé par les facteurs manipulés.

TABLEAU 37 STRUCTURE DES EFFETS SUR L'INDICE FRASB

	HUILE	BOISSON	PARFUM	SHAMP
REV	++	+	--	
CSP	--	--	+	
MAG	=	=	=	=
TAY	++	++	++	--
CONS	--	--	--	++
DATE		++	--	-

Le facteur qui influence le plus cet indice est la quantité consommée. Après analyse plus fine des corrélations entre les univers (à partir des tableaux des bifl qui figurent à l'Annexe 22), nous constatons que c'est à l'univers des petits consommateurs qu'est attribuable l'importance de l'effet du facteur. En effet, gros et moyens consommateurs apparaissent bien corrélés. Ce résultat est logique dans la mesure où les intervalles inter-achats sont estimés au plan individuel. Ces estimations, pour être fiables, nécessitent un nombre suffisant d'actes d'achats pour *"chaque"* individu (ménage), ce qui n'est pas le cas avec de petits consommateurs qui ne permettent pas d'obtenir des estimations fiables.

INDICE LEM :

Les tableaux 35 montrent que les niveaux moyens de corrélations sont élevés pour cet indice. Seul le facteur MAGASIN influence cet indice de façon considérable. L'analyse des écarts figure dans le tableau suivant :

TABLEAU 38 STRUCTURE DES EFFETS SUR L'INDICE LEM

	HUILE	BOISSON	PARFUM	SHAMP
REV	+	+	++	
CSP	=	=	-	
MAG	--	--	--	-
TAY	+	+	+	=
CONS	++	--	-	=
DATE	++	++	++	+

Nous pouvons observer que le magasin a l'effet le plus fort sur les quatre marchés. Inversement, le facteur DATE a très peu d'effet sur cet indice, ce que nous avions déjà constaté lors de l'analyse de la fiabilité des indices (section 2 du chapitre V précédent). Les facteurs REVENU et TAILLE du FOYER ont des effets peu prononcés, ceci sur tous les marchés.

Comme pour l'indice FRASB précédent, mais à un degré moindre, (voir Annexe 22), l'univers des petits consommateurs est très différent des deux autres, ce qui donne son importance à l'effet du facteur QUANTITE CONSOMMEE. Les univers Gros et Moyens consommateurs sont bien corrélés. La mesure de la concurrence à partir des achats des petits consommateurs conduit à l'obtention de structures de concurrence différentes de celles qui sont obtenues avec l'échantillon total, structures qui sont certainement peu fiables.

Nous retiendrons que l'indice de Lehmann est globalement peu sensible aux effets des facteurs (et aussi à la taille de l' échantillon). Un facteur a cependant beaucoup d'influence sur cet indice, il s'agit du magasin.

INDICE MERA :

Les niveaux de similarité pour cet indice sont faibles ($b_{i..}$ = .18) comme nous le constatons sur les tableaux 35. Contrairement à la plupart des indices, le magasin est ici le facteur qui a le moins d'importance sur le calcul de cet indice. Comme nous l'avons vu au chapitre V, lors de l'étude de la fiabilité des indices, le facteur DATE apparait comme ayant un effet important sur cet indice. Ceci est certainement attribuable à un effet de saisonnalité des achats du produit (comme c'est le cas sur le marché BOISSON).

TABLEAU 39 STRUCTURE DES EFFETS SUR L'INDICE MERA

	HUILE	BOISSON	PARFUM	SHAMP
REV	--	--		
CSP	=	--		
MAG	++	=		+
TAY	-	++		=
CONS	=	=		=
DATE		--		--

INDICE MERR :

Le niveau moyen de similarité entre les univers est très faible pour cet indice (voir tableaux 35). A partir du tableau d'analyse des écarts, nous observons, comme pour MERA, que le facteur MAGASIN a moins d'influence sur cet indice que sur la majorité des autres. Ceci constitue une propriété intéressante de ces indices puisque le facteur MAGASIN s'avère être celui qui a le plus d'influence sur la majorité des indices. Ceci peut être expliqué par la faible influence du point de vente sur les indices MERA et MERR par le fait qu'ils ne nécessitent pas, pour être calculés, d'avoir effectué une définition préalable du marché. De ce fait, lors de leur calcul, ces indices ne sont pas sensibles à l'assortiment présent dans le magasin où le panéliste effectue ses achats. Ils ne dépendent pas, comme c'est le cas des transferts entre marques, de l'univers concurrentiel fixé par l'offre en magasin.

D'autre part, contrairement à la plupart des indices, la période d'observation est le facteur qui a le plus d'importance.

TABLEAU 40 STRUCTURE DES EFFETS SUR L'INDICE MERR

	HUILE	BOISSON	PARFUM	SHAMP
REV	-	+	=	
CSP	=	-	=	
MAG	++	=	=	++
TAY	=	=	-	=
CONS	-	=	+	=
DATE	-	-	=	--

INDICESABA

Cette analyse confirme encore le niveau très élevé de sensibilité de cet indice (voir tableaux 35). L'examen du tableau suivant montre que les facteurs ayant le moins d'influence sont le REVENU, la CSP et le TEMPS.

TABLEAU 41 STRUCTURE DES EFFETS SUR L'INDICE SABA

	HUILE	BOISSON	PARFUM	SHAMP
REV	++	=	+	
CSP	++	+	+	
MAG	--	--	-	-
TAY		=	=	-
CONS	--	-	-	=
DATE	++	++ -	--	=

INDICE RAO :

Cet indice est lui aussi très sensible à la manipulation des facteurs. Le niveau des bif. est faible (voir tableaux 35). L'examen de l'influence des facteurs caractérise un fort effet du MAGASIN et de la QUANTITE CONSOMMEE.

TABLEAU 42 STRUCTURE DES EFFETS SUR L'INDICE RAO

	HUILE	BOISSON	PARFUM	SHAMP
REV	++	+	=	
CSP	=	=	--	
MAG	=	--	-	-
TAY	-	=	++	=
CONS	--	--	=	++
DATE	++	++	=	-

INDICES TOTT et TRTO

Ces indices ont des niveaux de sensibilité assez forts. Il est intéressant de constater que les patterns d'influence sur ces deux indices ne sont pas identiques. Notamment, l'effet sur l'indice TRTO du facteur DATE est fort (sur deux marchés HUILE et BOISSON) alors qu'il ne l'est pas sur l'indice ELRO.

TABLEAU 43 STRUCTURE DES EFFETS SUR L'INDICE TOTT

	HUILE	BOISSON	PARFUM
REV	=	+	++
CSP	=	=	=
MAG	-	+	--
TAY	+	-	-
CONS	--	--	=
DATE	++	++	++

TABLEAU 44 STRUCTURE DES EFFETS SUR L'INDICE TRTO

	HUILE	BOISSON	PARFUM
REV	++	=	++
CSP	++	+	=
MAG	--	--	--
TAY	--	++	-
CONS	=	=	=
DATE	--	--	++

Ces différences d'effets des facteurs tendent à montrer que les deux logiques d'analyse de la concurrence: Défense (en termes de transferts en provenance de la marque i) et Attaque (en termes de transferts vers la marque i) ne permettent pas de caractériser les mêmes types d'effets des facteurs. Ceci vient appuyer les arguments en faveur d'un concept de concurrence pouvant être, dans sa dynamique, non-symétrique.

INDICES ELRO et TREL

Il est intéressant de confronter la structure des influences de ces deux indices à celle des deux indices précédents. Il apparait ici encore que les deux versions: "Attaque" et "Défense" présentent des patterns d'effets des facteurs assez différents. Notamment, ici encore, le facteur DATE n'influence pas de la même manière les deux versions de l'indice. Son effet est important pour TRTO (sur les deux marchés HUILE et BOISSON) alors qu'il est peu important pour l'indice ELRO. Par contre, les deux versions "Attaque" (TRTO et TREL), qui sont calculées sur des données de transfert en actes ou en individus présentent des structures d'influence assez semblables. Il en est de même pour les deux versions "Défense" (TOTT et ELRO). Ceci vient appuyer encore nos remarques précédentes quant au caractère non-symétrique de la concurrence, du moins tel qu'elle est mesurée à partir des transferts entre marques.

TABLEAU 45 STRUCTURE DES EFFETS SUR L'INDICE ELRO

	HUILE	BOISSON	PARFUM
REV	+	=	++
CSP	+	=	--
MAG	-	--	--
TAY	-	++	++
CONS	--	-	--
DATE	++	+	++

TABLEAU 46 STRUCTURE DES EFFETS SUR L'INDICE TREL

	HUILE	BOISSON	PARFUM
REV	+	++	++
CSP	++	=	--
MAG	--	--	--
TAY	--	++	++
CONS	=	+	--
DATE	--	--	++

SECTION 4 ANALYSE COMPARATIVE DES EFFETS DE MESURE ET DES EFFETS DE STRUCTURE

Nous avons pu constater, notamment dans la section précédente, que les niveaux de similarité entre deux mesures dépendent au moins autant de l'indice qui est utilisé pour effectuer la mesure que des univers sur lesquels sont effectuées les mesures.

Nous appelons *"effet d'instrument"* (ou effet de mesure), la part dans une mesure qui est due à l'instrument utilisé pour effectuer la mesure. De la même manière, nous appelons *"effet de structure"*, la part dans la mesure qui est attribuable à l'objet mesuré (ici l'univers sur lequel la mesure est pratiquée).

SS1 MISE EN FORME DU PROBLEME

L'objet de cette section est d'analyser et de comparer l'importance de chacun des deux effets cités précédemment: "l'effet dû à l'instrument est-il plus ou moins fort que l'effet dû à la structure". Afin d'aller plus loin dans cette analyse, nous pouvons considérer que l'effet dû à l'instrument de mesure est la conjonction de deux éléments :

- le type de données qui est utilisé pour calculer l'indice (par exemple des transferts entre les marques, calculés en actes).que nous avons appelé "matrice des intrants de base" dans la conclusion du chapitre précédent.
- la formule de calcul elle même, par exemple, pour l'indice LEM, nous avons:

$$S_{ij} = \frac{N_{ij}}{N_{i.} + N_{.j}}$$

Il est donc intéressant de séparer, à l'intérieur de ce que nous appelons *"effet d'instrument"*, ce qui est dû au type de données utilisées de ce qui est dû à la formule de calcul. Pour dissocier ces deux éléments et analyser leurs effets respectifs sur la mesure, nous avons créé deux variables, DONNE (qui caractérise le type d'input de base utilisé) et CALCUL (qui caractérise la formule de calcul utilisée), codifiées de la manière suivante :

TABLEAU 47 CODIFICATION DES FACTEURS DONNE ET CALCUL

VARIABLE	MODALITES	CODE
DONNE	Périodes inter-achats	1
	Transferts en individus	2
	Transferts en actes	3
	Parts de choix en actes	4
	Parts de choix en individus	5
CALCUL	FRASB	1
	LEM	2
	RAO SABA	3
	MERA MERR	4
	TOTT TRTO ELRO TREL	5

Le tableau suivant donne la codification et la correspondance entre les trois variables: INDICE, DONNE et CALCUL

TABLEAU 48 CORESPONDANCE ENTRE LES FACTEURS INDICE DONNE ET CALCUL

	INDICE	DONNE	CALCUL
FRASB	2	1	1
LEM	3	2	2
MERA	4	4	3
MERR	5	5	3
SABA	6	3	4
TOTT	7	3	5
TRTO	8	3	5
RAO	9	2	4
ELRO	10	2	5

D'autre part, nous pouvons codifier la variable descripteur de l'univers sur lequel est calculé l'indice de la manière suivante :

TABLEAU 49 CODIFICATION DUFACTEUR UNIVERS

UNIVERS	CODE
REVG	1
REVM	2
REVP	3
CSPS	4
CSP7	5
CSP8	6
MAGG	7
MAGM	8
MAGP	9
TAYG	10
TAYM	11
TAYP	12
CONSG	13
CONSM	14
CONSP	15
DATE1	16
DATE2	17

Les 17 modalités de cette variable correspondent aux 17 univers issus de la manipulation des facteurs (tous ces univers ne sont pas disponibles pour le marché SHAMP sur lequel tous les facteurs n'ont pas été manipulés).

Nous pouvons alors analyser l'effet de chacune des quatre variables que nous venons de présenter à l'aide d'un modèle d'analyse de variance comportant un ou plusieurs facteurs, selon le nombre d'effets qui sont testés simultanément.
La variable expliquée dans le modèle est la variable ACORR1 (valeur absolue du coefficient de corrélation de Pearson). ACORR1 peut être ainsi exprimée comme une fonction des caractéristiques de chacune des deux mesures qui sont corrélées. Pour les quatre variables précédentes, qui figurent chacune dans les deux éléments du couple de corrélation, nous adopterons la codification suivante :

PREMIER ELEMENT DU COUPLE	DEUXIEME ELEMENT DU COUPLE
I1 INDICE1	I1 INDICE1
D1 DONNE1	D1 DONNE1
C1 CALCUL1	C1 CALCUL1
U1 UNIVERS1	U1 UNIVERS1

Chaque corrélation peut être alors exprimée comme suit :

$$ACORR1 = f(D1, C1, U1, D2, C2, U2)$$

Afin d'éviter d'estimer des modèles ayant trop de paramètres, nous avons utilisé une codification simplificatrice. Pour chaque facteur étudié (INDICE, DONNE, UNIVERS, CALCUL) nous n'avons considéré que deux cas: soit les deux éléments du couple de la corrélation sont identiques relativement au facteur étudié, soit ils sont différents. Nous obtenons alors la codification suivante :

INDICE = 1 si I1 = I2 (corrélations intra-indice)
= 0 si I1 ≠ I2 (corrélations inter-indices)

DONNE =1 si D1 = D2
=0 si D1 ≠ D2

CALCUL = 1 si C1 = C2
= 0 si C1 ≠ C2

UNIVERS = 1 si U1 =U2 (corrélations intra-univers)
= 0 si U1 ≠U2 (corrélations inter-univers)

Les modèles étudiés sont finalement les suivants :

M1 ACORR = f (UNIVERS)
M2 ACORR = f (DONNE)
M3 ACORR = f (CALCUL)
M4 ACORR = f (INDICE)
M5 ACORR = f (UNIVERS, DONNE, CALCUL, INDICE,U.C, U.D, C.D)

U.C est l'interaction entre les facteurs UNIVERS et CALCUL
U.D est l'interaction entre les facteurs UNIVERS et DONNE
C.D est l'interaction entre les facteurs CALCUL et DONNE.

Les modèles M1 à M4 sont à un seul facteur. Le modèle M5 comprend simultanément les quatres facteurs ainsi que les interactions possibles entre ces facteurs (U.C, U.D, C.D).

SS2 RESULTATS

Les résultats figurent, pour les quatre marchés, dans les tableaux suivants.

TABLEAU 50α MARCHE HUILE

MODELE	ACORR	DDL	F	PROB	R^2
M1	UNIVE	1	32.5	.0001	.011
M2	DONNE	1	167.2	.0001	.055
M3	CALCUL	1	443	.0	.139
M4	INDICE	1	84	.0001	.029
M5	UNIVE	1	69	.0001	.221
	INDICE	1		.0	.74
	DONNE	1		54	.0001
	CALCUL	1		.325	.0001
	U.D	1		15	.0001
	U.C	1		.123	.0001
	D.C	1		.18	.0001

TABLEAU 50ß MARCHE DES BOISSONS

ACORR	DDL	F	PROB	R^2
UNIVE	1	112	.0001	.038
DONNE	1	104	.0001	.17
CALCUL	1	251	.0001	.08
INDICE	1	55	.0001	.019
UNIVE	1	174	.0001	.18
INDICE	1	17	.0001	
DONNE	1	23	.0001	
CALCUL	1	171	.0001	
U.D	1	16	.007	
U.C	1	102	.0001	
D.C	1	0.11	.73	

TABLEAU 50γ MARCHE DES PARFUMS

NOBS = 2393

ACORR	DDL	F	PROB	R^2
UNIVE	1	80	.0001	.032
DONNE	1	85	.0001	.035
CALCUL	1	607	.0	.20
INDICE	1	34	.0001	.014
UNIVE	1	167	.0001	.343
INDICE	1	2.5	.11	
DONNE	1	31	.0001	
CALCUL	1	699	.0	
U.D	0	3	.08	
U.C	1	6	.0001	
D.C	0	3	.0001	

TABLEAU 50π MARCHE DES SHAMPOOINGS

NOBS = 739

ACORR	DDL	F	PROB	R^2
DONNE	1	103	.0001	.123
CALCUL	1	99	.0001	.119
INDICE	1	8	.0042	.011
UNIVE	1	121	.0014	.47
INDICE	1	2	.15	
DONNE	1	101	.0001	
CALCUL	1	80	.0001	
U.D	1	40	.0001	
U.C	1	120	.0001	
D.C	1	6	.013	

Dans le modèle M5, nous constatons que le facteur INDICE n'a aucun effet. Ceci est normal car sa variance est parfaitement expliquée par les deux facteurs DONNE et CALCUL qui en sont la décomposition.

Tous ces modèles ont été estimés en utilisant l'Analyse de Variance à un ou plusieurs facteurs et interactions. Il est important de noter qu'en analyse de la variance à plusieurs facteurs, quand le plan d'analyse n'est pas orthogonal (c'est notre cas ici puisque les facteurs étudiés ne sont ni indépendants, ni orthogonalisés), l'importance relative de chaque facteur (ce qu'apporte chaque facteur dans l'explication de la variance totale) peut être appréciée à partir de plusieurs critères. Nous avons utilisé ici un critère permettant de mesurer l'importance de chaque facteur indépendamment de son *"ordre"* d'introduction dans le modèle. Ce critère équivaut à comparer la somme des carrés des erreurs pour deux sous-modèles: le modèle complet sans le facteur étudié et le modèle complet avec le facteur étudié (voir SAS Statistics, chapitre 20 p 466).

A partir de l'étude des quatres tableaux précédents nous ferons les commentaires qui suivent.

Dans les modèles à un seul facteur, nous pouvons observer que sur les quatre marchés, les quatre facteurs ont un effet très significatif. L'importance de ces quatre facteurs n'est cependant pas identique. Le facteur qui apparait comme ayant le plus fort effet (excepté sur SHAMP) est le facteur CALCUL. Ceci signifie que les niveaux des corrélations observées dépendent principalement de la formule de calcul utilisée. Cet effet est, en particulier, supérieur à celui dû au facteur DONNE. Le type de données utilisées comme intrant pour calculer un indice de concurrence a ainsi "moins d'importance" que la formule de calcul de cet indice, pour expliquer la similarité entre deux vecteurs de mesure de concurrence. Ces résultats sont confirmés dans le modèle à quatre facteurs et interactions. Excepté sur le marché SHAMP, le facteur CALCUL apparait comme étant le plus significatif. Par importance relative décroissante, nous trouvons ensuite le facteur UNIVE (l'Univers de calcul). Ceci confirme bien que la manipulation des facteurs a un effet sur le niveau des corrélations observé mais aussi et c'est rassurant, que la mesure est significativement influencée par l'objet mesuré. Il reste cependant que l'effet dû a la

formule de calcul est bien le premier facteur explicatif de la corrélation entre deux mesures. Sur tous les marchés sauf PARFUM, il y a une forte interaction entre l'univers de calcul et la formule de calcul. Ceci reflète sans doute deux résultats que nous avons mis à jour dans la section précédente, à savoir:

1 Les indices ne sont pas influencés à l'identique par la manipulation des facteurs. Par exemple, les indices construits à partir de "transferts entre marques" sont plus influencés par le facteur MAGASIN que les indices construits à partir des parts de choix.

2 Ce ne sont pas tant les données qui sont influencées par la manipulation des facteurs mais plutôt les estimations qui sont faites à partir de ces données. Nous avons mis en évidence la remarquable stabilité des données de transfert et conjointement la mauvaise stabilité d'indices comme SABA et RAO qui sont pourtant calculés sur les mêmes intrants.

Nous conclurons cette section en insistant sur le résultat suivant: le premier facteur expliquant les différences entre les vecteurs de mesure de la concurrence obtenus est le facteur CALCUL, c'est à dire la formule de calcul de l'indice. Nous avions déjà noté ce point en constatant que des indices comme LEM et RAO, TOTT et TRTO, bien que calculés sur les mêmes données, peuvent considérablement diverger. Ce résultat nous permet d'insister sur l'intérêt de voir un indice de mesure de la concurrence comme étant en fait une formule de calcul appliquée sur ce que nous avons appelé à la conclusion du chapitre précédent une matrice d'intrants. Il apparait ici que les indices proposés dans la littérature différent essentiellement par leurs modes de calcul et non par les données sur la base desquelles ils sont estimés. Il serait intéressant, dans l'avenir, d'appliquer des formules de calcul identiques sur différentes matrices d'intrants de manière à effectuer des comparaisons qui ne seraient pas dues à des différences de type méthodologique (par exemple la formule de l'indice de Lehmann appliquée sur différentes matrices d'intrants: transferts entre marques, achats simultanés).

CONCLUSION DU CHAPITRE VI

Bien que la diminution de la taille des échantillons liée à la manipulation des facteurs ait rendu assez difficile l'analyse des résultats, nous pouvons caractériser les résultats essentiels suivants:

• 1 Les données de transfert bruts entre marques (indices IND et ACT) se révèlent très peu sensibles à la manipulation des facteurs excepté le facteur MAGASIN.

•2 La majeure partie des indices sont fortement influencés par la manipulation des facteurs, les niveaux de corrélations obtenus sont faibles. L'indice LEM apparait cependant significativement moins influençable que les autres, c'est le seul indice qui présente des niveaux de stabilité *"satisfaisants"*. Ceci s'explique par le fait qu'il est très proche, dans son mode de calcul, des indices de transferts bruts (IND et ACT) qui eux, se révèlent très peu sensibles. Cet indice, ainsi que les indices IND et ACT, présentent des qualités de robustesse à la diminution de la taille de l'échantillon mais aussi à la non-représentativité de l'échantillon des actes d'achats (la manipulation des facteurs équivaut en quelque sorte à constituer des échantillons non représentatifs selon la variable manipulée). Ceci constitue une propriété très intéressante de ces indices.

Par contre, les indices SABA et RAO apparaissent nettement plus sensibles à l'effet des facteurs. Il s'agit là d'un effet dû, non pas aux intrants qui servent à les calculer (les transferts entre marques), mais à leur mode de calcul. Nous expliquons cette sensibilité par la forme multiplicative du dénominateur qui, rappelons le, représente les changements attendus sous hypothèse d'absence de concurrence entre les deux marques.

Les autres indices construits à partir de données de transfert (ELRO, TOTT, TREL et TRTO) présentent un niveau de sensibilité moins fort que celui des indices SABA et RAO. Cependant, ce niveau de sensibilité reste tout de même assez élevé. Ceci peut être attribué à leur trop grande finesse de calcul qui repose, rappelons le, sur des comparaisons de profils d'échanges pondérés par l'inverse des niveaux moyens d'échanges entre les marques.

Les autres familles d'indices, construites à partir des intervalles inter-achats ou à partir des parts de choix, apparaissent dans cette étude comme étant très sensibles.

• 3 Tous les facteurs que nous avons manipulés n'influencent pas les indices dans les mêmes proportions. Certains facteurs ont une influence qui, en moyenne, est significativement plus forte que d'autres.

Le facteur MAGASIN s'est révélé être, en moyenne et sur tous les marchés, le facteur qui est significativement le plus influent. Il s'agit là, vraisemblablement, d'un *"Effet d'Offre"*, offre qui varie d'un type de magasin à un autre (Grande, Moyenne ou Petite surface) mais aussi d'un point de vente à un autre. Cette constatation nous inspire deux réflexions importantes.

La première concerne le choix de l'échantillon des points de vente, pour la constitution d'un panel de consommateurs.

Il faut sélectionner avec beaucoup de soins l'échantillon des points de vente qui sont retenus dans le panel. Ceci est visiblement un des soucis majeurs de la société Nielsen qui vient, avec SCAN 7000, de passer d'un échantillon de dix magasins à un échantillon de quinze magasins (grandes surfaces alimentaires). L'objectif de cette société, pour 1993, est d'avoir un échantillon de dix mille panélistes répartis dans soixante dix magasins. A

cette époque, le rapport du nombre de panélistes par magasin sera ainsi passé de 500 (en 1988 avec SCAN 5000) à 143. L'objectif de la société, qui est plus d'augmenter le nombre de points de vente que le nombre de panélistes, est donc tout à fait cohérent avec nos résultats .

La deuxième réflexion a trait au problème du calcul de la concurrence entre les marques.

Il apparaît nécessaire, pour mesurer la concurrence entre deux marques, de prendre en compte cet effet d'offre. Prenons l'exemple de deux marques A et B qui sont rarement présentes de façon simultanée en magasin. La concurrence calculée à partir des indices étudiés ici va conclure à l'absence de concurrence entre ces deux marques. Pourtant, que se passerait-il si les deux marques A et B se retrouvaient (au hasard des référencements) présentes sur le même linéaire ? Il s'agit là d'un problème de différence entre la concurrence manifeste, telle qu'elle s'observe réellement sur le marché et la concurrence potentielle, qui elle, découle de la substituabilité de ces marques dans l'esprit du consommateur (voir partie 1, chapitre 1, section 4) et revient à considérer que les influences des points de vente sont identiques.

Notons par ailleurs que le degré de *"présence simultanée"* de deux marques en rayon, dans l'échantillon des magasins étudié, n'est peut être pas représentatif du degré de co-présence de ces deux marques à l'échelle du marché tout entier. Deux marques pour pouvoir être concurrentes doivent être présentes ensemble en rayon. Ceci signifie qu'il est nécessaire, pour mesurer convenablement la concurrence entre deux marques, de prendre en compte le degré de distribution de ces deux marques (la distribution numérique). Ceci parait être une préoccupation de la société Nielsen pour l'utilisation de SCAN 7000 que nous pouvons citer: *"il est nécessaire de faire la synthèse, à chaque moment, de la valeur d'Offre d'une marque par rapport à ses concurrentes .. SCAN7000 est un panel de consommateurs explicatif du comportement d'achat et du lien entre achat et Offre en magasin "* (SCAN 7000 Manuel d'utilisation 1990). La valeur d'offre d'un produit comprend les éléments du mix de ce produit (le produit lui même, son prix, son image ..), mais aussi les caractéristiques de l'offre en magasin de ce produit, à savoir, son facing, sa promotion, sa PLV s'il y a lieu, etc ..

Selon ce point de vue, les variables de *"situation d'achat"* deviennent *"partie intégrante"* de l'offre du produit", offre qui influence le comportement d'achat du consommateur. Elles représentent ainsi autant de composantes du mix du produit et c'est, selon nous, à la force de vente du producteur qu'incombe la responsabilité de la gestion de cette partie du mix.

La prise en compte de la valeur d'offre d'un produit revient à intégrer (tout au moins en partie) les facteurs situationnels dans l'explication du comportement du consommateur. Nous avons vu, dans la première partie de cette recherche, que l'influence des facteurs situationnels est une des raisons majeures avancées pour expliquer la différence entre les préférences déclarées d'un consommateur et ses comportements réels observés. SCAN 7000 permet ainsi, grâce à la connaissance des principales variables représentatives de l'offre en magasin, de prendre effectivement en compte cet effet d'Offre.

Nous pouvons formaliser ce point de vue en exprimant la concurrence entre deux marques de manière à rendre compte:

- du linéaire des marques
- des promotions dans le magasin
- de la présence de PLV

Le modèle suivant exprimerait alors la concurrence observée :

$$S_{ij} = C_{ij} + a\,P_{ji} + b\,L_{ji} + cPLV_j + e$$

Où :

S_{ij} est la concurrence réelle observée entre les marques i et j
C_{ij} est une constante du modèle, représente la concurrence théorique (potentielle, une fois que l'effet des variables du Point de Vente est pris en compte). P_{ji} est la promotion relative de la marque j par rapport à la marque i.
L_{ji} est le linéaire relatif de j part rapport à i.
PLV_{ji} est la PLV relative de la marque j par rapport à la marque i. e est un terme d'erreur aléatoire

D'autres facteurs manipulés apparaissent moins influents que la moyenne. Il s'agit des variables CSP, REVENU et de la variable TEMPS.

4 Tous les indices ne sont pas influencés par les mêmes facteurs. Nous avons pu notamment constater que le facteur MAGASIN, s'il a une influence globalement forte, est le facteur qui influence le moins les indices MERA et MERR. Ceci constitue une propriété intérressante de ces indices puisque la variable MAGASIN s'est avérée être la plus influente pour la plus part des indices. La faible influence de cette variable sur les indices MERA et MERR s'explique par le fait qu'ils ne nécessitent pas, pour être calculés, d'avoir effectué une définition préalable du marché. Ainsi, ces indices ne sont pas sensibles à l'assortiment présent dans le magasin où le panéliste effectue ses achats. Ils ne dépendent pas de l'univers de concurrence qui est fixé par l'offre magasin.

D'autre part, nous constatons que le facteur TEMPS influence beaucoup les indices MERR et MERA, ce qui reflète peut-être un effet de saisonnalité des quantités achetées.

5 Le facteur QUANTITE CONSOMMEE, au premier abord apparait comme un facteur de forte influence. En fait, l'examen des corrélations entre les trois univers issus de la manipulation de ce facteur: CONSG, CONSM, CONSP, corrélations qui figurent à l'Annexe 22, montre que CONSP est opposé aux deux autres. Nous en conclurons que les petits consommateurs d'un produit se prêtent mal à la mesure des relations de concurrence à partir d'historiques d'achats. Il est en effet nécessaire de disposer de chroniques d'achats (pour chaque panéliste) suffisamment fournies. Les petits consommateurs n'offrent pas de chroniques suffisantes à cet égard. Il est donc important de veiller à avoir des panélistes suffisamment acheteurs et des périodes d'observations suffisamment longues pour effectuer des mesures de concurrence dans de bonnes conditions. Il importe moins d'avoir beaucoup de panélistes que d'avoir des chroniques d'achats suffisamment longues, pour effectuer des estimations convenables. Ceci est très net pour les indices FRASA et FRASB pour lesquels il est nécessaire d'estimer des périodes moyennes inter-achats au plan individuel.

6 Notre dernier point de conclusion portera sur les problèmes de non-symétrie des relations de concurrence. Nous avons pu constater que les indices *"Défense"* ne sont pas influencés de la même manière par les facteurs que les indices *"Attaque"*. Au sein d'un même type d'indice, ("Attaque" ou bien "Défense"), nous observons par contre des structures d'effets assez similaires. Ceci suggère que la dynamique de la concurrence, telle qu'elle est mesurée à l'aide des données de transfert, n'est pas symétrique. Le profil de ce qu'une marque perd au profit de ses concurrentes n'est pas identique à ce qu'elle gagne sur ces mêmes concurrentes (ce qui traduit une évolution du marché) et surtout n'est pas influencé de la même manière par les facteurs que nous avons manipulés. Notamment, nous avons pu constater que malgré le fait que les transferts entre marques soient très stables dans le temps, TRTO et TREL sont fortement influencés par le facteur TEMPS, alors que TOTT et ELRO ne le sont pas.

CONCLUSION GENERALE

Cette recherche nous a permis de progresser sur deux fronts: la mise au point d'une méthodologie au service de l'analyse de la structure des marchés et la définition des concepts qu'elle met en jeu ainsi que l'étude du problème de la mesure de la concurrence entre les marques.

Dans la première partie de la recherche, nous avons insisté sur l'importance de l'analyse de la structure des marchés et nous l'avons définie comme étant un des concepts fondamentaux des "analyses produits" aux côtés de la segmentation des consommateurs et du positionnement des marques (voir figure 1 page 5).
Nous avons proposé une démarche en trois étapes:

- définition du marché
- choix d'une mesure de la relation entre les marques
- étude et représentation graphique de la structure des relations entre marques

La première étape, la définition du marché, est un préalable à toute "étude produit", qu'il s'agisse d'analyser la structure d'un marché, d'en segmenter les consommateurs ou encore de faire des études de positionnement des marques. La définition du marché permet de fixer le degré de généralité de l'analyse, degré qui peut évoluer sur un continuum dont les deux extrémités sont l'analyse tactique et l'analyse stratégique.

A des analyses tactiques correspondent des définitions étroites du marché alors qu'à des analyses stratégiques doivent correspondre des définitions plus larges.

L'objet d'analyses tactiques est de mieux connaître un marché, à une période donnée, de manière à contrôler l'effet des actions marketing qui ont été mises en oeuvre.

Les analyses de type stratégique permettent d'anticiper l'évolution d'un marché de manière à pouvoir effectuer des choix fondamentaux relativement aux marques: lancement, repositionnement, abandon.

Une fois le marché défini, il est alors possible de s'interroger sur le type de relation entre marques que l'on souhaite étudier, ce qui constitue la deuxième phase de la méthodologie.
Nous avons vu (voir figure 2 page 10) que la relation entre deux marques peut être étudiée à deux niveaux fondamentalement différents: le consommateur et le marché.

Sur le marché, les relations entre marques sont étudiées à partir des comportements effectifs des individus. Nous avons montré que ces relations correspondent alors au concept de concurrence. Ainsi, par concurrence, nous faisons référence *"aux relations entre deux marques telles qu'elles se sont réellement exprimées sur le marché, au niveau agrégé"*.

Nous avons effectué une revue des différents indices de mesure de la concurrence qui ont été présentés dans la littérature en les distinguant notamment selon leur nature déterministe ou stochastique. D'autre part, nous avons vu que le concept de concurrence a été opérationnalisé à partir de trois types de concepts: les transferts entre marques, les intervalles inter-achats, les parts de choix. Nous avons noté qu'il est encore possible d'utiliser un quatrième concept: les achats simultanés entre marques.

MESURE DE LA CONCURRENCE
TRANSFERTS ENTRE MARQUES
INTERVALLES INTER-ACHATS
PARTS DE CHOIX
ACHATS SIMULTANES

Au niveau du consommateur, la relation entre deux marques correspond au concept de substituabilité perçue. Elle doit être mesurée à partir des jugements individuels des consommateurs. Nous avons vu que le modèles qui ont été proposés dans la littérature procèdent à partir de quatre types de concepts: les substituabilités perçues, les attitudes et les préférences, les ensembles évoqués et enfin les choix.

MESURE DE LA SUBSTITUABILITE
SUBSTITUABILITE PERCUE
ATTITUDES - PREFERENCES
ENSEMBLES EVOQUES
CHOIX

Nous avons, d'autre part, insisté sur le caractère non-symétrique de la relation entre deux marques. La grande majorité des modèles de concurrence existant dans la littérature considère cependant cette relation comme étant symétrique. Nous avons présenté l'assymétrie de la relation entre deux marques au plan conceptuel (première partie, chapitre I, section 4) et au plan statistique (première partie, chapitre III, section 2). Nous avons proposé un indice de mesure de la relation entre deux marques pouvant s'appliquer à différents types de données (jugements et comportements) et qui prend en compte cette assymétrie.

Une fois qu'un indice de mesure de la relation entre deux marques a été défini au plan théorique et opérationnalisé correctement, il est alors possible d'analyser la structure qui est à l'origine des mesures de concurrence effectuées. Nous avons vu qu'une telle analyse peut être effectuée en adoptant une démarche exploratoire ou confirmatoire. Dans ce dernier cas, il est nécessaire de définir une structure, au plan théorique, avant de tester si les données collectées confirment ou infirment cette structure. Enfin, nous avons présenté de nouvelle méthodes pouvant permettre de caractériser des structures de marché sans utiliser une analyse spatiale de la matrice d'indices De telles méthodes méritent encore recherches et validations, mais elle nous semblent très prometteuses.

Les trois étapes de cette méthodologie constituent à notre avis un cadre de réflexion intéressant qui oblige le chercheur ou le praticien à considérer les points fondamentaux lors de la mise en oeuvre d'une méthodologie d'analyse de la structure des marchés, ce qui conditionne l'obtention de structures valides.

Dans la deuxième partie de la recherche, nous nous sommes centrés sur la deuxième phase de cette méthodologie: la meure de la relation entre les marques. Considérant qu'il s'agit avant tout du problème de la mesure d'un concept non observable: la relation entre deux marques, nous avons utilisé les outils et les techniques qui ont été développés en psychométrie. Ceux-ci n'ont été appliqués, jusqu'à présent, qu'à l'étude de phénomènes sociologiques ou psychologiques. Cette recherche montre que la théorie de la mesure peut être utilisée dans d'autres domaines et notamment celui de la mesure de la relation entre les marques.à partir de données de panels (la concurrence). La théorie de la mesure permet de considérer le concept de concurrence avec plus de rigueur et notamment aux différents niveaux suivants:

- la définition du concept de concurrence
- l'opérationnalisation de sa mesure
- la vérification que la mesure opérée est valide, robuste et donne lieu à des interprétations pertinentes, etc ..

Cette recherche a ainsi permis de caractériser un nombre important de résultats relatifs à la mesure de la concurrence entre les marques à partir de données de panels.

L'étude de validité des indices de mesure de la concurrence que nous avons menée à partir d'un échantillon des indices proposés dans la littérature marketing (concurrence au sens où nous l'avons définie dans la première partie de la recherche), nous a permis de caractériser les résultats suivants:

1- Tous les indices étudiés ne convergent pas vers la mesure d'un même concept de concurrence. Ce résultat empirique est très important. Il confirme les remarques que nous avons faites concernant le manque d'effort de la part des chercheurs pour définir précisément, au plan théorique, le concept qu'ils mesurent à savoir la concurrence. La concurrence, telle qu'elle est mesurée par les indices que nous avons étudiés, comporte plusieurs dimensions qui sont apparues comme étant indépendantes. En particulier, nous avons caractérisé trois dimensions correspondant à trois catégories bien différentes d'indices.

A la première catégorie, appartient l'indice proposé par Fraser et Bradford, indice qui est calculé à partir des périodes inter-achats entre marques, sans tenir compte de l'univers concurrentiel auquel appartiennent les marques.

Dans la deuxième catégorie, nous trouvons l'indice proposé par Merunka et Bourgeat, qui est estimé à partir des quantités achetées et que nous avons calculé selon deux versions.

La troisième catégorie est constituée par les autres indices, qui sont calculés à partir de transferts entre marques (indices de Lehmann, de Rao et Sabavala, indices "Attaque" et "Défense") ainsi que les indices calculés à partir des périodes inter-achats en tenant compte cette fois de l'univers concurrentiel auquel appartiennent les marques.

Au sein de ce troisième groupe d'indices, nous avons pu vérifier, en utilisant plusieurs méthodes, que la validité de traits est bien respectée. Tous ces indices mesurent le même concept et ne mesurent pas autre chose. En particulier, nous retiendrons que les mesures indirectes de similarité procédant à partir de la comparaison des profils de transferts entre

marques (indices "Attaque" et "Défense") convergent avec les mesures procédant à partir des transferts directs entre marques (indices de Lehmann, de Rao et Sabavala). Nous retiendrons aussi que les mesures utilisant les intervalles inter-achats des marques convergent avec les indices utilisant les transferts, ceci quand l'univers concurrentiel est pris en compte lors de l'estimation de ces intervalles.

Nous avons proposé un autre concept à partir duquel pourrait être établies des mesures de la concurrence: les achats simultanés. De nouvelles recherches sont donc nécessaires pour valider ce concept et établir son degré de convergence avec les trois catégories de mesure précédentes qui ont été caractérisées.

D'un point de vue méthodologique, nous avons montré que les résultats d'une analyse de validité à l'aide de méthodes dérivées de la matrice multi-traits multi-méthodes dépendent pour partie du choix d'un indice d'association entre les mesures et notamment du choix d'un indice de mesure d'un lien linéaire (coefficient de corrélation linéaire de Pearson) ou d'un lien monotone (coefficients de Spearman ou de Kendall). Un tel résultat devrait amener le chercheur à s'interroger plus longuement, lors d'une étude de validité, sur le choix d'un indice d'association particulier.

2 - L'étude de la fiabilité a montré que, même au sein de la troisième catégorie d'indices, catégorie pour laquelle nous avons établi une bonne validité de traits, il existe des niveaux de "fiabilité" fort différents. En particulier, les deux versions de l'indice de Rao et Sabavala que nous avons étudiées sont affectées par beaucoup d'erreur de mesure, erreur au sens de la théorie du domaine d'échantillonnage. Nous avons proposé une explication à ce manque de fiabilité. A l'opposé, l'indice de Lehmann a présenté, dans la majorité des analyses, une fiabilité supérieure à celle des autres indices.

L'étude de la stabilité dans le temps des indices (test-retest) a elle aussi caractérisé de fortes différences dans les niveaux de fiabilité des indices. Les données de transfert bruts (indices IND et ACT) et à un moindre niveau l'indice de Lehmann, présentent une très bonne stabilité dans le temps et (ou) au partage en deux de l'échantillon des actes d'achats, partage qui a été rendu nécessaire par la mise en oeuvre du test-retest. A l'opposé, les indices de Fraser et Bradford ainsi que les deux versions de l'indice proposé par Merunka et Bourgeat présentent des niveaux de stabilité dans le temps qui sont très faibles.

A ce niveau encore, nous avons pu observer que le niveau de fiabilité d'une échelle dépend du type d'indice d'association choisi. Les indices de lien monotones (coefficients de Kendall et Spearman) étant par nature plus généraux, ils permettent d'obtenir des niveaux de lien plus élevés.

3- L'étude de la validité faciale des indices a confirmé les résultats de l'analyse de la validité de traits et notamment le fait que le concept de concurrence peut recouvrir, selon l'avis des experts, des réalités bien différentes. Les praticiens parlent ainsi de concurrence en termes de prix, de format du produit ou de toute autre variable d'action marketing. L'analyse de la concurrence à partir de données d'historiques d'achats permet alors au responsable marketing d'analyser un marché et les relations entre les marques appartenant à ce marché selon plusieurs points de vue: les prix, le format du produit, les actions promotionnelles, etc... Selon cette perspective, la mesure de la concurrence conçue en tant que "concept unique" ne conserve plus beaucoup de sens. Les praticiens lui substituent des analyses "diagnostic", ces analyses étant

conduites par rapport à des variables spécifiques et en particulier les variables d'action marketing.

Il est donc opportun que de nouvelles recherches soient menées afin de préciser le contenu du concept de concurrence et dans le but d'en définir les différentes dimensions s'il s'avère qu'il s'agit d'un concept multidimensionnel.

Parmi les indices que nous avons étudiés, ce sont les indices "Attaque" et "Défense", en version pondérée, qui permettent de construire les cartes de marché les plus intéressantes, du point de vue des experts utilisateurs de ces cartes. Nous notera de plus, que ces deux indices permettent d'obtenir les meilleurs ajustements statistiques avec l'utilisation de l'analyse des similarités, pour construire les cartes du marché. Nous attribuons, pour partie, la qualité de l'ajustement obtenu avec ces indices, au fait que ces indices sont capables de prendre en compte la non-symétrie des relations de concurrence. Nous avons en effet insisté, tout au long de cette recherche, sur le caractère non-symétrique (selon notre point de vue) du concept de concurrence. De nouvelles recherches sont nécessaires dans ce domaine afin de construire un corps théorique capable de conceptualiser et de valider cette hypothèse de non-symétrie.

4 Nous avons montré que la définition d'un indice de concurrence entre deux marques met en jeu deux niveaux différents d'analyse:

1- Le choix d'un concept servant de base à la mesure (transferts entre marques, intervalles inter-achats, parts de choix, achats simultanés) que nous avons appelé " input de base au calcul". A ce niveau, nous avons souligné que de nouvelles recherches devraient permettre:

- d'exploiter les données d'achats simultanés, qui ne sont que très peu exploitées à l'heure actuelle
- de caractériser des concepts permettant de distinguer, sous l'appellation de concurrence, différents types de relations dues à la "substituabilité perçue", la "recherche de variété", la diversité des "situations d'usage du produit", "l'hétérogénéité des goûts des consommateurs au sein du ménage".

2- Le choix d'une formule de calcul permettant, à partir d'une matrice d'intrants particulière, de calculer une matrice d'indices de concurrence entre marques. De nouvelles recherches devraient viser à mettre au point des formes de calcul ayant les propriétés suivantes:

- propriété d'une distance: nous avons vu l'avantage de telles mesures en terme de modélisation spatiale de la concurrence (analyse des similarités), avec les indices Attaque et Défense, qui possèdent de telles propriétés
- prise en compte de l'assymétrie des relations de concurrence: les indices Attaque et Défense permettent de prendre en compte cette assymétrie.
- robustesse: les mesures ne doivent pas être trop sensibles aux caractéristiques des consommateurs, des points de vente, à la période d'observation et enfin à la diminution de la taille de l'échantillon des panélistes. A ce niveau, nous avons vu que c'est l'indice de Lehmann qui présente les meilleures garanties de robustesse.
- les formules de calcul doivent permettre de prendre en compte des différences dans les volumes de vente des marques, c'est à dire dans les parts de marché de ces marques

- enfin, les formules de calcul doivent être capables de prendre en compte les différences des niveaux de fidélité aux marques. A ce niveau encore, les indices Attaque et Défense sont satisfaisant d'un point de vue conceptuel. Il serait par contre nécessaire de mener des études spécifiques à l'étude de ce problème en manipulant (contrôlant) la variable niveau de fidélité des marques.

Nous avons souligné, qu'il peut être intéressant d'effectuer des mesures de concurrence en utilisant la même formule de calcul appliquée à différentes matrices d'inputs (par exemple l'indice Attaque et Défense calculé sur des matrices de transferts entre marques ou d'achats simultanés). Une telle procédure comporte en effet deux avantages:

- il est alors possible d'effectuer des analyses comparatives des cartes du marché qui sont obtenues à partir de la mesure de différentes connotations du concepts de concurrence (par exemple achats simultanés et transferts entre marques).
- les résultats obtenus sont comparables et ne sont pas biaisés par la simple utilisation de formules de calcul.différentes. En effet, nous avons montré que la variable qui explique le mieux les différences entre deux mesures effectuées sur des matrices d'input différentes n'est pas la variable "input de base" mais la variable "formule de calcul". Ceci signifie que les résultats (sous forme de cartes ou de graphes) dépendent plus des modes de calcul utilisés que des données (input de base) elles-mêmes.

La deuxième partie de la recherche empirique, qui est consacrée à l'étude de l'effet des facteurs manipulés sur la mesure de la concurrence, a permis de caractériser les points suivants.

1- Parmi les facteurs que nous avons manipulés (revenu, CSP, taille du foyer, quantité achetée, temps, point de vente), le point de vente s'est avéré être le facteur qui a la plus forte influence sur la mesure de la concurrence et ce, de manière très significative. Nous avons développé au chapitre VI les implications que représente un tel résultat, notamment relativement à la constitution des panels permettant d'obtenir des mesures de la concurrence *"représentatives"* au niveau du marché dans son ensemble. L'obtention de tels indices représentatifs passe donc par un échantillonnage soigné des points de vente. En particulier, les panels de consommateurs qui ne sont représentatifs que d'une zone de chalandise particulière (zone fermée) paraissent plus aptes, d'après ces résultats, à la mise en oeuvre d'études expérimentales qu'à fournir des mesures représentatives de la concurrence.

Nous avons d'autre part caractérisé l'importance de l'effet de la manipulation du facteur "quantité consommée". Cependant, l'influence de la manipulation de ce facteur n'est pas importante tant qu'il s'agit des "gros" ou des "moyens" consommateurs, il n'y a pas de différences notables entre les structures de concurrence obtenues à partir de ces deux segments. Par contre, les petits consommateurs ne permettent pas d'obtenir des représentations comparables à celles que l'on obtient avec l'échantillon total. Il est donc conseillé de veiller à la représentativité de l'échantillon des panélistes relativement à la variable "quantité consommée" et surtout de ne pas sur-représenter les petits consommateurs. Ceci est d'autant plus important que l'on utilise un indice estimé au plan individuel, comme c'est le cas avec l'indice proposé par Fraser et Bradford. De tels indices, pour être estimés de manière fiable, nécessitent de disposer,

pour chaque ménage, d'historiques suffisamment longs, ce qui signifie qu'à période d'observation identique, il vaut mieux étudier des consommateurs présentant un niveau de fréquence d'achat suffisant.

2 - Nous avons pu observer, d'autre part, que les facteurs manipulés ont un niveau d'influence variable selon les indices. La majeure partie des indices étudiés se sont révélés fortement sensibles à la manipulation des facteurs ou au partage de l'échantillon lié à cette manipulation (nous citerons à ce propos les indices de Rao et Sabavala, de Fraser et Bradford et de Merunka et Bourgeat). A l'inverse, les indices de "transferts bruts" (IND et ACT) ont fait la preuve d'une très bonne stabilité lors de la manipulation de tous les facteurs, excepté le facteur relatif au point de vente. Les transferts entre marques sont ainsi une mesure très stable de la concurrence, peu sensible à la manipulation des facteurs. Ce résultat signifie que même calculés sur un échantillon non-représentatif des panélistes (échantillon non-représentatif relativement aux variables que nous avons manipulées, excepté le point de vente), ces indices permettent d'obtenir une mesure de la concurrence qui reste proche de celle qui est obtenue avec l'échantillon tout entier (échantillon représentatif). Les indices de transferts bruts entre marques ont montré par la même occasion qu'ils sont très stables à la diminution de la taille de l'échantillon des panélistes, diminution qui résultait de la manipulation des facteurs étudiés. Ils peuvent ainsi être qualifiés de *"mesure robuste de la concurrence"*.
L'indice de Lehmann est le seul, mis à part les indices de transferts bruts, qui présente un niveau de stabilité satisfaisant. Tous les autres indices s'avèrent très sensible à la manipulation des facteurs ou (et) à la diminution de la taille de l'échantillon. Ainsi, les deux versions de l'indice de Rao et Sabavala, qui sont estimées sur la base de données de transfert bruts, doivent leur instabilité à leur mode de calcul et non pas aux données de transfert elles-mêmes.

3 - Tous les indices ne sont pas influencés dans les mêmes proportions par les même facteurs. Nous citerons en particulier le cas de la variable "temps" qui a peu d'influence sur la majorité des indices mais qui est le facteur le plus influent (parmi ceux que nous avons étudiés) sur l'indice proposé par Merunka et Bourgeat.
A l'inverse, la variable "point de vente" compte parmi les variables qui ont le moins d'influence sur cet indice (nous avons proposé une explication à ce phénomène dans la conclusion du chapitre VI).

La mesure de la concurrence est donc une opération complexe qui nécessite, si le chercheur ou le praticien ne souhaite pas prendre le risque d'obtenir des *"mesures arbitraires"*, de développer la connaissance théorique du concept qui lui est sous-jacent.
Les données d'historiques d'achats ont connu et connaîtront encore des progrès considérables grâce aux progrès réalisés par les technologies relatives au recueil et à la transmission de l'information. Ces progrès portent à la fois sur la richesse, la finesse et la précision des informations collectées, mais aussi sur la quantité et la représentativité de ces données. Ce développement de l'information disponible s'est accompagné d'une multiplication des recherches proposant de nouvelles méthodes de mesure et de diagnostic de la concurrence entre les marques. Face à cette multiplication, il nous semble opportun de développer le contenu théorique du concept de concurrence ou, plus généralement, de *"relation entre deux marques"*. Ce concept nous est apparu comme comprenant plusieurs dimensions et comme étant non-

symétrique. Il est nécessaire, par de nouvelles recherches de construire un corps théorique permettant d'expliquer ces phénomènes. Si la concurrence comprend plusieurs dimensions, il est opportun d'en préciser la nature afin de savoir comment les mesurer et quand utiliser une dimension plutôt qu'une autre pour effectuer l'analyse diagnostic des relations entre marques.

De la valeur du champ théorique relatif au concept de "relation entre deux marques" dépendent la qualité des mesures que les chercheurs et les praticiens du marketing peuvent en faire et la pertinence des analyses qui reposent sur de telles mesures.

*
* *

BIBLIOGRAPHIE

G.M. ALLENBY, 1988, "A Unified Approach to Identifying, Estimating and Testing Demand Structures with Aggregate Scanner Data", Working Paper, Chicago Illinois, January.

T. AMEMIYA, 1981, "Qualitative Response Models : a SURVEY", Journal of Economic Litterature, vol XIX december (1483-1536).

J.C ANDERSON, 1985, "A Measurement Model to Assess Measure Specific Factors in Multiple-Informant Research", Journal of Marketing Research, vol XXII February (86-92).

J.V. ANDERSON, D.W. GERBING, J.E. HUNTER, 1987, "On the Assesment of Unidimensional Measurement: Internal and External Consistency, an Overall Consistency Criteria", vol XXIV november (432-437).

APPLIED PSYCHOLOGICAL MEASUREMENT, "Multidimensional Scaling and its Applications", M.L. DAVISON & L.E Johnes, Special Issue, Volume 7 1983 n° 4.

P. ARABIE, D. CARROLL, W. DE SARBO, J. WIND, 1981 "Overlapping Clustering : a New Method for Product Positioning" Journal of Marketing Research, Vol XVIII, August (310-317)

H. ASSAEL, 1971, "Perceptual Mapping to Reposition Brands", Journal of Advertising Research, vol 11, n°1 (39-42).

P.AURIER, 1989, "Segmentation: une Approche Méthodologique", Recherche et Applications en Marketing, vol 3.

E. BABAKUS, C.E. FERGUSSON, K.G. JORESKOG, 1987, "The Sensitivity of Confirmatory Maximum Likelihood Factor Analysis to Violations of Measurement Scale and Distributional Assumptions" Journal of Marketing Research, vol XXIV may (22-228).

R. BAGOZZI, 1977, "Structural Equations Models in Experimental Research", Journal of Marketing Research, vol XIV may (209-226).

R. BAGOZZI, "Causal Models in Marketing", John Wiley & Sons,1980.

A. BARNETT, 1976, "More on Market Share Theorem", Journal of Consumer Research, vol XIII february (104-109).

F.M. BASS, W.L. Wilkie, 1973, "A Comparative Analysis of Attitudinal Prediction of Brand Preference", Journal of Marketing Research, vol X august (262-269).

F.M. BASS, 1974, "The Theory of Stochastic Preference and Brand Switching", Journal of Marketing Research, vol XI february (1-20).

F.M. BASS, A. JEULAND, G.P. WRIGHT, 1976,"Equilibrium Stochastic Choice and Market Penetration Theories: Derivations and Comparisons", Management Science, vol 22 n° 10 june (1051, 1063).

F.M. BASS, T.L. PILON, 1980, "A Stochastic Brand Choice Framework for Econometric Modeling of Time Series Market Share Behavior", Journal of Marketing Research, vol XVII november (486-497).

F.M. BASS, M. GIVON, M.U. KALWANI, D. REBSTEIN, G.P. WRIGHT, 1984, "An Investigation into the Order of the Brand Choice Process", Marketing Science, vol 3 n° 4 (267-287)

R.R. BATSELL, L.M. LODISH, 1981, "A Model and Measurement Methodology for Predicting Individual Consumer Choice", Journal of Marketing Research, vol XVIII february (1-12).

R.R. BATSELL,.J.C. POLKING, 1985, "A new Class of Market Share Models", Marketing Science, vol 4 n° 3 (177-198).

J BAUDRILLARD, 1968, "Le Système des Objets", Mediations, Gallimard.

R.W. BELK, 1974, "An Exploratory Assesment of Situational Effects in Buyer Behavior", Journal of Marketing Research, vol 11 may (156-163)

R.W. BELK, 1975, "Situational Variables and Consumer Behavior Effects in Buyer Behavior", Journal of Consumer Research,2 december (157-164)

R.W. BELK, 1979, "A Free Response Approach to Devoloping Product-Specific Consumption Situation Taxonomies", in Analytic Approaches to Product and Marketing Planning, A.D. SHOCKER, ed. Cambridge, MA: Marketing Science Institute (177-196)

P. BENTLER, 1980, "Multivariate Analysis with Latent Variables: Causal Modeling", Annual Review of Psychology, vol 31 (419-456).

J.R. BETTMAN, 1971, "The Structure of Consumer Choice Processes", Journal of Marketing Research, vol VIII november (461-471).

J.R. BETTMAN, 1976, "Modélisation du Traitement de l'Information par l'Acheteur", Encyclopédie du Marketing, vol 4.

J.R. BETTMAN, 1979, "An Information Processing Theory of Consumer Choice", Reading, Madison-Wesley.

R.C. BLATTBERG, 1981, "Evaluation of Stochastic Brand Choice Models", in "Marketing Decision Models", Schultz & Zoltners, North Holland, chap 9 (183-206).

R.C. BLATTBERG and R.J. DOLLAN, 1981, "An assesment of the Contribution of log-Linear Models to Marketing Research", Journal of Marketing Research, Vol45 (89-97)

G.W. BOHRNSTEDT, E.F. BORGATTA, "Social Measurement Current Issues", Sage Publications, Beverly Hills, London, 1980.

J.F. BOSS, 1973, "La Segmentation du Marché : Le Point des Définitions et des Techniques", Revue Française du Marketing, n° 48 (3-24)

J.F. BOSS, 1974, "Typologie et Segmentation : un Etat des Techniques", Revue Française du Marketing, n° 51 (69-93)

J.F. BOSS, 1974, "Le Positionnement: Problèmes et Premiers Résultats", Revue Française du Marketing, n°52 (57-75)

J.D. BOURGEOIS, G.H. HAINES, M.S. SOMMERS, 1979, "Defining an Industry", TIMS/ORSA , Stanford, CA, march 26.

J.M. BOUROCHE, 1977, "Analyse des Données en Marketing", Masson, Monographie de l'AFCET..

J.M. BOUROCHE, 1978, "Analyse des Proximités et des Préférences", note de travail Coref n° 30, novembre.

BRICKMAN, Philip, et D'AMATO, Barbara, (1975), "Exposure Effects in a Free-Choice Situation", *Journal of Personality and Social Psychologie*, 32, (415-420)

R.D. BUZZELL, 1981, "Are there "Natural" Market Structures?", Journal of Marketing, vol 45, (42-51)

R.E. BUCKLIN, V. SRINIVASAN, 1989, "Determining Cross-Price Elasticities and Product-Market Structure Through the Measurement of Consumer Preference Structures", Working Paper n° 193, J.E Anderson Graduate School of Management, UCLA, september.

D.T. CAMPBELL, D.W FISKE, 1959, "Convergent and Discriminant Validity by the Multitrait Multimethod Matrix", Psychological Bulletin, n° 56 (81-10).

J.M. CARMAN, 1966, "Brand Switching and Linear Learning Models", Journal of Advertising Research, vol 6 n° 2 (23-31).

G.S. CARPENTER, D.R. LEHMANN, 1985, "A Model of Marketing Mix, Brand Switching, and Competition", Journal of Marketing Research, vol XXII august (318-329).

J.D. CARROLL, 1968, "Generalization of Canonical Correlations Analysis to Three or More Sets of Variables", Proceedongs 76th annual convention, APA, 3, (227-228).

W.A. CHANCE, N.D. FRENCH, 1972, "An Exploratory Investigation of Brand Switching", Journal of Marketing, vol IX may (226-229).

J.L. CHANDON, S. PINSON, "Analyse Typologique, Théories et Applications", MASSON 1981.

J.L. CHANDON, A. STRAZZIERI, 1986, "Une Analyse de Structure de Marché sur la Base de la Mesure de l'Emsemble Evoqué", Recherche et Applications en Marketing, n°1, avril (17-40)

R.G. CHAPMAN, R. STAELIN, 1982, "Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model", Journal of Marketing Research, august (288-301).

G.A. CHURCHILL, 1979, "A Paradigm for Developing Better Measures of Marketing Constructs", Journal of Marketing Research, vol 16 (64-73).

G.A CHURCHILL, A. PECOTICH, 1982, "A Strctural Equation Investigation of the Pay Satisfaction - Valence Relationship Among Salespeople", Journal of Marketing, (114-124).

G.A. CHURCHILL, J.P. PETER, 1984, "Research Design Effects on the Reliability of Rating Scales: Meta-Analyses", Journal of Marketing Research, vol XVI february (360-375).

C.C. CLOGG, 1979, "Some Latent Structure Models for the Analysis of Likert-Type data", Social Science Research, vol 8 december (287-301).

D.L. COCKS, J.R. VIRTS, 1975, "Market Definition and Concentration in the Ethical Pharmaceutical Industry", International Publication of Eli Lilly and co, Indianapolis.

T.D. COOK, D.T. CAMPBELL,, "Quasi Experimentation: Design & Issues for Field Settings", Rand McNally College Compagny, Chicago 1979.

C.H. COOMBS, 1964, "A Theorie of Data", (John Wiley & Sons, inc, New York.

L.G. COOPER, M. NAKANISHI, 1983, "Two Logit Models for External Analysis of Preferences", Psychometrika, vol 48 n° 4 december (607-620).

L.G. COOPER, 1988, "Competitive Maps: The Structure Underlying Asymmetric Cross Elasticities", Management Science, vol 34 n° 6 june (707-723).

D.L. COURSEY, 1988, "Preference Trees, Preference Hierarchies and Consumer Behavior", Journal of Consumer Resarch, Vol 15 december (407-409)

E.P. COX, 1980, "The Optimal Number of Response Alternatives for a Scale: a Review, Journal of Marketing Research vol XVI november (407-422).

L.J.CRONBACH, 1951, "Coefficient Alpha and the Internal Structure of Tests", Psychometrika, vol 16 september (297-334).

L.J. CRONBACH, 1955, "Construct Validity in Psychological Test", Psychological Bulletin, vol 52 july (281-302).

L.J. CRONBACH, G.C. GLEZER, H. NANDA, N. RAJARATNAM, "Dependability of Behavioral Measurements: Theory Generalizability for Scores and Profiles", New York, Wiley & Sons, Inc, 1972.

I.S. CURRIM, 1981, "Using Segmentation Approaches for Better Prediction and Understanding from Consumer Mode Choice Models", Journal of Marketing Research, Vol XVIII (301, 309)

I.S.CURRIM, 1982, "Predictive Testing of Consumer Choice Models no Subject to Independance of Irrelevant Alternatives, Journal of Marketing Research, vol XIX may (208-222).

P.M. DAVIES, A.P.M. COXON, 1981,"Key Texts in Multidimentional Scaling" Heinemann Educational Books

G.S. DAY, A.D. SHOCKER, R. SRIVASTAVA, 1979, "Customer Oriented Approaches to Identifying Products-Markets", Journal of Marketing, vol 43 (8-19)

G.S. DAY, R WENSLEY, 1983, "Marketing Theorie with a Strategic Orientation", Journal of Marketing, vol 47 (79-89).

G.S. DAY, R WENSLEY, 1988, "Assessing Advantage: A Framework for Diagnosing Competitive Superiority", Journal of Marketing, vol 52 april (1-20).

W.S. De SARBO and D.K. HILDEBRAND, 1980, "A. Marketer's Guide to Log-Linear Models for Qualitative Data Analysis", Journal of Marketing, Vol (40-51)

W.S. DeSARBO, 1982, "GENNCLUS: New Models for General Nonhierarchical Clustering Analysis", Psychometrika vol 47 n° 4 december (449-475).

W.S. DeSARBO, DE SOETE, 1984, "On the Use of Hierarchical Clustering for the Analysis of Nonsymmetric Proximities", Journal of Consumer Research, vol 11 june (601, 610).

X.S. DeSARBO, D.L. HOFFMAN, 1987, "Constructing MDS Joint Spaces from Binary Choice Data: a Multidimensional Unfolding Threshold Model for Marketing Research, Journal of Marketing Research, vol XXIV february (40-54).

W.S. DeSARBO, A.K. MANRAI, 1988, "A New Multidimensional Scaling Methodology For The Analysis of Asymmetric Proximities Data Incorporating the Notion of Distance-Density", Working Paper, July

W.S. DeSARBO, J. CHO, 1989, "A Stochastic Multidimensional Scaling Vector Threshold Model for the Spatial Representation of "Pick Any/N" Data", Psychometrika, vol 54 n° 1 march (105-129).

G. De SOETE, 1983, "A Least Square Algorithm for Fitting Additive Trees to Proximity Data", Psychometrika vol 48 n° 4 december (621-626).

P.R. DICKSON, 1982, "Person-Situation: Segmentation's Missing Link", Journal of Marketing, vol 46 (56-64)

P.R. DICKSON, J.L. GINTER, 1988, "Segmentation de Marché, Différenciation de Produit et Stratégie Marketing", Recherche et Applications en Marketing, vol III n° 1 (35-51).

W.R. DILLON, D.G. FREDERICK, V. TANGPANICHDEE, 1985, "Decision Issues in Building Perceptual Product Spaces with Multi-Attribute Rating Data", Journal of Consumer Research, vol 12 june (47-63).

W.R. DILLON, T. DOMZAL, T.J. MADDEN, 1986, "Evaluating Alternative Product Positioning Strategies", Journal of Advertising Research, vol 26, august (29-35).

J. DOLLARD, 1949, "Under What Conditions Do Opinions Predict Behavior", Public Opinion Quarterly, vol 12 (623-632).

P. DOYLE 1972 " Market Segmentation by Factor Analysis", European Journal of Marketing, n° 6/1 (1-6)

P. DOYLE, 1977, "The Application of Probit, Logit and, Tobit in Marketing: a Review", Journal of Business Research, 5, september (235-248).

DRIVER, Michael, et STREUFERT, Siegfried, (1964), "The 'general Incongruity Adaptation Level' Hypothesis: An Analysis and Integration of Cognitive Approach to Motivation", *paper* n°114, Institute for Research in the Behavioral, Economic and Management Sciences, Krannert Graduate School of Management, Perdue University.

C. DRODGE, R.Y. DARMON, 1987, "Associative Positioning Strategies Through Comparative Advertising: Attribute Versus Overall Similarity Approaches", Journal of Marketing Research, vol XXIV november (377-388).

B. DUBOIS, Y EVRARD, 1975, "Some Problems Associated with the Use of Multidimensional Scaling", Proceedings 28 Esomar Wapor Congres, Special Group.

P.L. DUBOIS, 1977, "Le Positionnement du Produit, Strategie de Marketing: Fondements, Efficacité", Thèse pour le Doctorat d'Etat es Sciences de Gestion, juillet, Université de Lille 1.

P.L DUBOIS, 1979, "Pour une Pratique du Positionnement des produits", Revue Française du Marketing, cahier 77 (5-31)

C. DUSSART, 1983, "Comportement du Consommateur et Stratégie de Marketing", Mc Graw-Hill Québec.

C. DUSSART, 1985, "Le Marketing de la Troisième Vague", Gaetan Morin Editeur, ESKA, Montréal.

A.S.C. EHRENBERG, 1972, "Repeat-Bying", Amsterdam : North-Holland, chap 3 "The Repeat Buying Structure of Market" (33-54), chap 4 "Basic Theorie" (55-82).

H.J. EINHORN, 1970, "Use of Nonlinear Noncompensatory Models in Decision Making", Psychological Bulletin, vol 73 (221-230).

T. ELROD and R.S. WINER : 1982 "An Empirical Evaluation of Aggregation Approaches for Developing Market Segments", Journal of Marketing Research, n°46 (65-74)

T. ELROD, 1988, "Choice Map: Inferring a Product Market Map From Panel Data", Marketing Science, vol 7 n° 1 (21-40).

T. ELROD, R.S. WINER, 1988, "An Empirical Comparison of Mapping Methods Based on Panel Data", Working Paper, Vanderbilt University december, Nashville.

Y EVRARD, G. LAURENT, 1982, "Les Modèles Log-Linéaires: Applications en Marketing", Cahier du Troisième Séminaire de Méthodologie de la Recherche en Marketing, FNEGE, (335-358).

Y. EVRARD, 1985 "Validité des Mesures et Causalité en Marketing", Revue Française de Marketing, n° 101 janvier (17-32).

P.H. FARQHAR, V.R. RAO, 1976, "A Balance Model for Evaluating Subsets of Multiattributed Items", Management Science, n° 5 january (528-539).

G FENNELL, 1978, "Consumers' Perceptions of the Product-Use Situation", Journal of Marketing, 42 april (38-47)

M. FISHBEIN, 1966, "The Relationships Between Beliefs, Attitudes, and Behavior", S. Feldman ed, Cognitive Consistency, New York.

C. FORNELL 1982, "A Second Génération of Multivariate Analysis". Volume 1 Methods, Praeger.

C. FORNELL, D.F. LARCKER, 1981, "Structural Equation Model With Unobservable Variables and Measurement Error: Algebra and Statistics", Journal of Marketing Research, vol XVIII august (382-388).

C. FRASER, J.W. BRADFORD, 1983, "Competitive Market Structure Analysis: Principal Partitioning of Revealed Substituabilities, Journal of Consumer Research, vol 10 june (15-30).

C. FRASER, J.W. BRADFORD, 1984, "Competitive Market Structure Analysis: A Reply", Journal of Consumer Research, vol 11 december (842-847).

W.A.H FROST, 1977, "La visualisation des Marchés: La Méthode du Tableau "Produit x Usage", Revue Française du Marketing, sept-oct 1977 (53-64)

D.H. GENSH, W.W. RECKER, 1979, "The Multinomial, Multiattribute Logit Choice Model", Journal of Marketing Research, vol XVI february (24-32).

D.H. GENSCH, 1985, "Empirically Testing a Disaggregate Choice Model for Segments", Journal of Marketing Research, vol XXII november (462-466).

D.H. GENSH, 1987, "A two-Stage Disagregate Attribute Choice Model", Marketing Science, vol 6 n° 3 summer (223-239).

D.W. GERBING, J.C. ANDERSON, 1982, "Some Methods for Respecifying Measurment Models to Obtain Unidimentional Construct Measurement", Journal of Marketing Research, vol XIX november (186-192).

D.W. GERBING, J.C. ANDERSON, 1988, "An Updated Paradigm for Scale Development Incorporating Unidimensionality and its Assesment", Journal of Marketing Research, vol XXV may (186-192).

H.S. GITLOW, 1979, "Discrimination Procedures for the Analysis of Nominally Scalled Data Sets" , Journal of Marketing Research, Vol 16 (387-393)

M. GIVON, 1978, D. HORSKY, 1978, "Market Share Models as Approximators of aggregated Heterogeneous Brand Choice Behavior", Management Science, 26 september (1404-1416).

M. GIVON, D. HORSKY, 1979, "Application of a composite Stochastic Model of Brand Choice", Journal of Marketing Research, vol XVI may (258-267).

M. GIVON, 1984, "Variety Seeking through Brand Switching", Marketing Science, vol 3 n°1 winter (1-23).

M. GIVON, 1985, "Variety Seeking, Market Partitioning and Segmentation", International Journal of Research in Marketing, N° 2 (117-127).

R. GLAZER, 1984, "Multiattribute Perceptual Bias as Revealing of Preference Structure", Journal of Consumer Research, vol 11 june (510-520).

C. GOURIEROUX, 1984, "Econometrie des Variables Qualitatives", Economie et Statistiques Avancées, Economica, Paris.

J.C. GOWER, 1978, "Unfolding: Some Technical Problems and Novel Uses", Presented at European Meeting on Psychometric and Mathematical Psychology, Uppsala.

P.E. GREEN, A MAHESHWARI, V.R. RAO, 1969, "Multidimensional Interpretation and Configuration Invariance of Multidimensional Scaling: an Empirical Study", Multivariate Behavioral Research, 4 avril (159-180).

P.E. GREEN, A MAHESHWARI, 1970, "A Note on the Multimensional Scaling of Conditional Proximity Data", Journal of Marketing Research, vol VII february (106-110).

P.E. GREEN, V. RAO, 1971, "Multidimensional Scaling and Individual Differences", Journal of Marketing Research, vol VIII february (71-77).

P.E. GREEN, V. RAO, 1972, "Configuration Synthesis in Multidimensional Scaling", Journal of Marketing Research, vol IX february (65-68).

P.E. GREEN, Y. WIND, A.K. JAIN, 1973, "Analysing Free Response Data in Marketing Research", Journal of Marketing Research, vol X february (45-52).

P.E. GREEN, F.J. CARMONE, D.P. WACHSPRESS, 1977, "On the Analysis of Qualitative Data in Marketing Research", Journal of Marketing Research, vol XIV february (52-59).

P.E. GREEN, W.S. DeSARBO, 1979, "Componential Segmentation in the Analysis of Consumer Trade-Offs", Journal of Marketing, vol 43 (83-91).

P. GREEN, W.DeSARBO, 1981, "Two Models for Representing Unrestricted Choice data", Advances in Consumer Research, 7 (309-312).

S.T. GROSS, C.M. NIMAN, 1975, "Attitude Behavior Consistency: a Review", Public Opinion Quarterly, (358-368).

R. GROVER, W.R. DILLON, 1985, "A Probabilistic Model for Testing Hypothesized Hierarchical Structures", Marketing Science, vol 4 n° 4 (312-335).

R. GROVER and V. SRINIVASAN, 1987, "A Simultaneous Approach to Market Segmentation and Market Structuring", Journal of Marketing Research, Vol XXIV (139-153)

R. GROVER, W.R. DILLON, 1988, "Understanding Market Caracteristics from Aggregated Brand Switching Data by the method of Spectral Decomposition", Working paper, University of Pittsburh, june.

P.M. GUADAGNI, J.D.C. LITTLE, 1983, "A Logit Model of Brand Choice Calibrated on Scanner Data", Marketing Science, vol 2 n° 3 summer (203-238).

J.P. GUILFORD, "Psychometric Methods", McGraw Hill, 1954.

D. GUPTA, 1984, "Some Practical Guidelines for Selection of Appropriate Perceptual Mapping Techniques", EMAC/ESOMAR, october (107-135).

E. GYSBRECHTS, P.V. ABEELE, 1988, "A General Class of Asymmetric Market Share Models: Specification and Estimation Issues", Working Paper, Belgium.

M.R. HAGERTY, J.M. CARMAN, G.J.RUSSELL, 1988, "Estimating Elasticities with PIMS Data: Methodological Issues and Substantive Implications", Journal of Marketing Research, vol XXV february (1-9).R.I. HALEY, 1968, "Benefit Segmentation: A Decision-Oriented Research Tool", Journal of Marketing, vol 32 july (30-35).

H.H. HARMAN, 1976, "Modern Factor Analysis", Third Edition, University of Chicago Press, Chicago.

R.A. HARSHMAN, P.E. GREEN, Y. WIND, M.E.LUNDY, 1982, "A Model for the Analysis of Asymmetric Data in Marketing Research", Marketing Science, vol 1 n° 2 spring (205-242).

J.R.HAUSER, F.S. KOPPELMAN, 1979, "Altenative Perceptual Mapping Techniques: Relative Accuracy and Usfulness, Journal of Marketing Research, vol XVII november (495-506).

J.R. HAUSER, P. SIMMIE, 1981, "Profit Maximisation Perceptual Positions: an Integrated Theory for the Selection of Products Features and Price, Management Science, vol 27 n° 1 january (33-56).

J.R. HAUSER, M. SHUGAN, 1983, "Defensive Marketing Strategies", Marketing Science, n° 2, (319-360).

J.R. HAUSER, S.P. GASKIN, 1986, "Application du Modèle de Comportement Defender", Recherche et Applications en Marketing, n°1 avril (59-92)

J.R HAUSER, G.L. URBAN, 1986, "The Value Priority Hypotheses for Consumer Budget Plans", Journal of Consumer Research, vol 12 march (446-462).

J. HAUSMAN, 1980, "Les Modèles Probit de Choix Qualitatifs", Annales de l'INSEE.

R.M. HEELER, M.L. RAY, 1972, "Measure Validation in Marketing", Journal of Marketing Research, vol IX november (361-370).

R.M. HEELER, T.W. WIPPLE, T.P. HUSTAD, 1977, "Maximum Factor Analysis of Attitude Data", Journal of Marketing Research, vol XIV february (42-51).

W.A. HENRY, R.V. STUMPF, 1975, "Time Accuracy Measures for Alternative Multidimensional Scaling Data Collection Methods", Journal of Marketing Research, vol XII may (165-170).

J.R. HICKS, 1946, "Value and Capital", second Edition. Oxford, England: Clarendon Press.

M.B. HOLBROOK, 1981, "Integrating Compositional and Decompositional Analyses to Represent the Intervening Role of Perceptions in Evaluative Judgments", Journal of Marketing Research, vol XVIII february (13-28).

J.H. HOLMES, 1973, "Profitable Product Positioning", MSU Business Topics, Spring (27-32).

HUBER,Joel, et REIBSTEIN, David (1978), "Attitude Measures and Choice Frequency: Some Pitfalls to be Avoided", *Attitute research plays for high stakes*, ed J.C Maloney, Chicago AMA.

J. HUBER, M.B. et HOLBROOK, 1979, "Using Attribute Ratings for Product Positioning: Some Distinctions Among Compositional Approaches", Journal of Marketing Research, vol XVI, november (507-516).

J. HUBER et C. PUTO, 1983, "Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects", Journal of Consumer Research, Vol 10 June (31-44)

J.W. HUCHINSON, 1990, "Netscal: A Network Scaling Algorithm for Nonsymmetric Proximity Data", in Press Psychometrika.

H. HRUSCHKA, 1986, "Market Definition and Segmentation Using Fuzzy Clustering Methods", International Journal of Research in Marketing, n° 3 (117-134)

D.N. JACKSON, 1969, "Multimethod Factor Analysis in the Evaluation of Convergent and Discriminant Validity", Psychological Bulletin, (259-271).

J. JACOBY, D.B. KYNER, 1973, "Brand Loyalty Vs Repeat Purchasing Behavior", Journal of Marketing Research, vol 10 february (1-9).

J. JACOBY, 1984, "Perspective on Information Overload", Journal of Cosumer Research, vol 10, march.

D. JAIN, F.M. BASS, Y.M. CHEN, 1990, "Estimation of Latent Class Models With Heterogeneous Choice Probabilities: An Application to Market Structuring, Journal of Marketing Research, vol XXVII february (94-101).

K. JEDIDI, W.S. DeSARBO, 1988, "A Stochastic Multidimensional Scaling Methodology for the Spatial Representation of Three-Mode Three-Way Binary Data, Working Paper, Graduate School of Business University of Michigan, Ann Arbor, MI, october.

A.P. JEULAND, 1978, "Brand Preference Over Time: A Partialy Deterministic Operationalization of the Notion of Variety Seeking" in AMA Research Frontiers in Marketing, ed S. Jain, Educators' Procedings, serie 43, 33-37.

A.P. JEULAND, 1979, "The Interaction Effect of Preference and Availability on Brand Switching and Market Share", Management Science, vol 25 n° 10 october (953-965).

A.P. JEULAND, F.M. BASS, G.P. WRIGHT, 1980, "A Multibrand Stochastic Model Compounding Heterogeneous Erlang Timing and Multinomial Choice Process", Operation Research, 28 march april (255-277).

S. JOHNSON, 1967, "Hierarchical Clustering Schemes", Psychometrika vol 32 n° 3.

M.D. JOHNSON, 1984, "Consumer Choice Strategies for Comparing Noncomparable Alternatives", Journal of Consumer Research, Vol 11 December, (741-753)

M.D. JOHNSON, 1988, "Comparability and Hierarchical Processing in Multialternative Choice", Journal of Consumer Research, Vol 15 december (303-314)

K.J. JORESKOG, 1969, "A General Approach to Confirmatory Maximum Likelihood Factor Analysis", Psychometrika, 1969, vol 34 (183-202).

K.J. JORESKOG, 1978, "Structural Analysis of Covariance and Correlation Matrices", Psychometrika, vol 43 (443-477).

K.J. JORESKOG, D. SORBOM, 1981, "Lisrel VI: Analysis of Linear Relationships by the Method of Maximum Likelihood. Chicago: National Education Ressource, Inc.

K.J. JORESKOG, D. SORBOM, 1982, "Recent Developpement in Structural Equation Modeling", Journal of Marketing Research, vol XIX november (404-416).

B.E. KAHN, M.U. KALWANI, D.G. MORRISSON, 1986, "Measuring Variety Seeking and Reinforcement Behaviors Using Panel Data", Journal of Marketing Research, Vol XXIII may (86-100)

B.E. KAHN, D.G. MORRISSON, G.P. WRIGHT, 1986,"Aggregating Individual Purchases to the Household Level", Marketing Science vol 5 n° 3 (260-268)

B.E. KAHN, M.U. KALWANI, D.G. MORRISSON, 1987, "Using Purchase Frequencies and Penetration Rates to Infer Brand Positioning: Niching Versus Change-of-Pace", Working Paper series N°162, J.E. Anderson Graduate School of Management, UCLA

M.U. KALWANI, D.G. MORRISSON, 1977, "A Parsimonious Description of the Hendry System", Management Science, vol 22 n° 5 january (467-477).

W.A.KAMATURA, R.K. SRIVASTAVA, 1984, "Predicting Choice Shares Under Conditions of Brand Interdependance", Journal of Marketing Research, november (420-433).

W.A.KAMATURA, R.K. SRIVASTAVA, 1986, "An Ideal Point Probabilistic Choice Model for Heterogeneous Preferences", Marketing Science, vol 5 n° 3 summer (199-218).

W.A. KAMAKURA, G.J. RUSSELL, 1989, "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure", Journal of Marketing Research, vol XXVI november (379-390)

J.W. KEON, 1983, "Product Positioning: TRINODAL Mapping Brand Images, Ad Images, and Consumer Preference", Journal of Marketing Research, vol XX november (380-392).

H. KONIG, M. NERLOVE, G. OUDIZ, 1979, "Modèles Log-Linéaires pour l'Analyse de Données Qualitatives", Annales de l'INSEE, n° 36, (30-81).

P. KOTLER et B. DUBOIS, 1986, "Marketing Management" PARIS, PUBLI-UNION

L. KRISHNAMURTHI, S.P. RAJ, 1988, "A Model of Brand Choice and Purchase Quantity Price Sensitivities", Marketing Science vol 7 n°1 winter (1-20).

P.M. KROONENBERG, J.D. LEEW, 1980, "Principal Component Analysis of Three-Mode Data by Means of Alternating Least Squares Algorithm", Psychometrika, vol 45 n° 1 march (69-97).

C.L. KRUMHANSL, 1978, "Concerning the Applicability of Geometric Models to Similarity Data: The Interrelationship Between Similarity and Spatial Density", Psychological Review, vol 85 (445-463).

A. KUMAR and W. DILLON, 1987, "Constrained Discrimination via MDI Estimation : the Use of Additional Information in Segmentation Analysis", Journal of Marketing Research, Vol XXIV november (396-403).

A. KUMAR, C.M. SASHI, 1989, "Confirmatory Analysis of Aggregate Hierarchical Market Structures: Inferences from Brand-Switching Behavior", Journal of Marketing Research, vol XXVI november (444-453).

L'Age de la Science, "Lectures Philosophiques: 2, Epistémologie", Edition Odile Jacob, 1989, Premier Chapitre.

K. LANCASTER, 1966, "A New Approach to Consumer Theory", Journal of Political Economy, vol 74 (132-157).

J.M. LATTIN and L. McALISTER, 1985, "Using a Variety-Seeking Model to Identify Substitute and Complementary Relationships Among Competing Products" Journal of Marketing Research, Vol XXII august (330-339).

G. LAURENT, 1978, "A Study of Multiple Variant Consumption for Frequently Purchased Consumer Products", Unpuplished PhD Dissertation, Sloan School of Management, MIT, Cambridge, MA.

L. LEBART, A. MORINEAU, K.M. WARWICK, "Multivariate Descriptive Analysis : Correspondance Analysis and Related Techniques for Large Matrices", New York: Wiley-Interscience, 1984.

J.P. LECOUTRE, P TASSI, 1987, "Statistique Non Paramétrique et Robustesse", Economie et Statistiques Avancées, Economica, Paris.

J.H. LEIGH, C MARTIN, 1981, "A Review of Situational Influence Paradigms and Research", Review of Marketing, AMA: Chicago (57-74).

D.R. LEHMANN, 1972, "Juged Similarity and Brand-Switching Data as Similarity Measures", Journal of Marketing Research, vol IX august (331-334).

J.H. LEVINE, 1979, "Joint-Space Analysis of "Pick-Any" Data: Analysis of Choices From an Unconstrained Set of Alternatives", Psychometrika, vol 44 n° 1, march (85-92).

R.D. LUCE, 1959, "Individual Choice Behavior: A Theoritical Analysis", New York: Wiley.

T. LUNN, 1972, "Segmentating and Constructing Markets", in Consumer Market Research Handbook, R.M. Worcester, Maidenhead, McGraw-Hill.

R.J. LUTZ, P. KAKKAR, 1975, "The Psychological Situation as a Determinant of Consumer Behavior", Advances in Consumer Research, vol 2, Ann Arbor, MI: Association for Consumer Research, (439-453).

J.G. LYNCH, 1985, "Uniqueness Issues in the Decompositional Modeling of Multiattribute Overall Evaluations: An Information Integration Perspective", Journal of Marketing Research, vol XXII february (1-17).

D.B. MacKAY, J.L. ZINNES, 1982, "Proscal: A Program for Probabilistic Scaling", Discussion Paper 218, Idiana University, December, Bloomington.

D.B. MacKey, J.L. ZINNES, 1986, "A Probabilistic Model for the Multidimensional Scaling of Proximity Data", Maketing Science, vol 5 n° 4 (325-344)

G.S. MADDALA, "Limited-Dependent and Qualitative Variables in Econometrics", Cambridge University Press, 1983.

N. MALHOTRA, 1984, "Reflexion on Information Overload: Paradig in Consumer Decision Making", Journal of Consumer Research, vol 10, march (436-).

N. MALHOTRA, 1984, "The Use of Linear Logits Models in Marketing Research", Journal of Marketing Research, vol XXI february (20-31).

N. MALHOTRA, 1987, "Validity and Structural Reliability of Multidimensional Scaling", Journal of Marketing Research, vol XXIV may (164-173).

N. MALHOTRA, A.K. JAIN, C PINSON, 1988, "The Robustness of MDS Configuration in the Case of Incomplete Data", Journal of Marketing Research, vol XXV february (95-102).C.F. MANSKI, D. McFADDEN, "Structural Analysis of Discrete Data with Econometrics Applications", The MIT Press, 1981.

A.A.J. MARLEY, 1981, "Multivariate Stochastic Processes Compatible with "Aspects" Models of Similarity and Choice", Psychometrika, vol 46 n° 4 december, (421-428).

A. MASLOW, "Motivation and Personnality, New York, Haroer & Row, 1952.

W.F. MASSY, B. MONTGOMERY, D.G. MORRISSON, 1970, "Stochastic Models of Buying Behavior", Cambridge: MIT Press.

L. McALISTER, 1979, "Choosing Multiple Items From a Product Class" Journal of Consumer Research, vol 6 december (213-224)

L. McALISTER, 1982, "A Dynamic Attribute Satiation Model of Variety-Seeking Behavior", Journal of Consumer Research, vol 9 september (141-150).

L. McALISTER, E. PESSEMIER, 1982, "Variety Seeking Behavior: an Interaction Disciplinary Review", Journal of Consumer Research, vol 9 december (311-322).

D. McFADDEN, 1980, "Econometric Models for Probabilistic Choice among Products", Journal of Business, vol 53 n° 3 (s13-s29).

D. McFADDEN, 1986, "The Choice Theory Approach to Market Research" Marketing Science, vol 3 n° 4 (275-297).

D.R. MERUNKA, P. BOURGEAT, 1988, "Une Methode de Mesure et de Representation de la Concurrence entre les Marques", Recherche et Applications en Marketing, vol III n° 2 (1-27).

A. MICALLEF, 1982, "Gestion Commerciale des Entreprises", Dalloz, Paris.

K.E. MILLER et J.L. GINTER, 1979, "An Investigation of Situational Variation in Brand Choice Behavior and Attitude", Journal of Marketing Research, vol 16 february (111-123).

W.L. MOORE et D.R. LEHMANN, 1982, "Effects of Usage and Name on Perceptions of New Products" vol 1 No 4,(351-370)

W.L. MOORE, R.S. WINER, 1987, "A Panel-DATA based Method for Merging Joint Space and Market Response Function Estimation", Marketing Science, vol 6 n° 1 winter (25-42).

P.E. MURPHY and B.M. ENIS, July 1986, "Classifying Products Strategically", Journal of Marketing, No 3 (24,42)

J.H. MYERS, E. TAUBER, 1977, "Market Structure Analysis", American Marketing Association, USA.

L.A. NEIDELL, 1972, "Procedures for Obtaining Similarities Data", Journal of Marketing Research, vol IX august (335-337).

NIELSEN, Recherches Marketing, "Single Source, La Proposition Nielsen", mars 1990.

T.P. NOVACK, C. STANGOR, 1987, "Testing Competitive Market Structures: an Application of Weighted Least Squares Methodology to Brand Switching Data", Marketing Science, vol 6 n° 1 (82-97).

J.C. NUNNALLY, "Psychometric Theory", McGraw Hill, second edition, 1978.

K. OGAWA, I. FURUKAWA, S. YAGI, 1989, "Inferring the Market Structure and Predicting the Choice Share of the Non-Alcoholic Drinks", ORSA/TIMS Conference, october, New York.

R. PARAMESWARAN, B.A. GREENBERG, D.N. BELLENGER, D.H. ROBERTSON, 1979, "Measuring Reliability: A Comparison of Alternative Techniques", Journal of Marketing Research, vol XVI february (18-25).

M. PARRY, C. GENTLER, 1988, "Market Partitions and Latent Structure Analysis: Accounting for Household Purchase Frequency", Working Paper, University of Texas at Dallas, march.

J.L. PAYNE, 1976, "Task Complexity and Contingent Processing in Decision Making, an Information Search and Protocol Analysis", Organizational Behavior and Human Performance, 16 august (366-387).

E. PESSEMIER, 1971, P. BURGER, R. TEACH, D. TIGERT, 1971, "Using Laboratory Brand Preference Scales to Predict Consumer Brand Purchases", Management Science, vol 17 february (371-385).

E. PESSEMIER, 1978, "Stochastic Properties of Changing Preferences", American Marketing Association, vol 68 n° 2 may (380-385).

E. PESSEMIER, M. HANDELSMAN, 1984, "Temporal Variety in Consumer Behavior", Journal of Marketing Research, vol XXI november (435-443).

J.P. PETER, 1979, "Reliability: a Review of Psychometric Basics and Recent Marketing Practices", Journal of Marketing Research, vol XVI february (6-17).

J.P. PETER, 1981, "Construct Validity: A Review of Basic Issues and Marketing Practices", Journal of Marketing Research, vol XVIII may (133-145).

J.P. PETER, M.L. RAY, "Measurement Readings for Marketing Research", American Marketing Association, Chicago 1984.

J.P. PETER, G.A. CHURCHILL, 1986, "Relationships Among Research Design Choices and Psychometric Properties of Rating Scales: a Meta Analysis", Journal of Marketing Research vol XXIII february (1-10).

A PIOCHE, 1974, "Images de Marques et Positionnements Relatifs: Quelques Techniques et Exemples d'Etude de la Structure d'un Marché", Revue Française de Marketing, n°51 (17-35)

K.R. POPPER, 1963, "Conjectures et Réfutations", Routlege & Kegan Paul, Payot, Trad Franç 1985, Paris.

M. PORTER, 1986, "L'Avantage Concurrentiel", InterEditions, Paris.

B. PRAS, J. SUMMERS, 1975, "A Comparison of Linear and non-Linear Evaluation Process Models", Journal of Marketing Research, vol XII august (276-281).

B. PRAS, 1977, "Comment les Consommateurs Opèrent-ils leurs Choix? Les Modèles Multi-attributs", Encyclopédie du Marketing, Volume 1 (1-42).

G.N. PUNJ, D.W. STEWART, 1983, "An Interaction Framework of Consumer Decision Making", Journal of Consumer Research, vol 10 september (181-196).

J.O. RAMSEY, 1977, "Maximum Likelihood Estimation in Multidimensional Scaling", Psychometrika, vol 42 (241-266).

V.R. RAO, 1972, "Changes in Explicit Information and Brand Perceptions", Journal of Marketing Research, vol IX may (209-213).

V.R. RAO, D.J. SABAVALA, 1981, "Inference of Hierarchical Choice Process from Panel Data", Journal of Consumer Research, vol 8 june (85-96).

S. RATNESHWAR, A.D. SHOCKER et D.W. STEWART, 1987, "Toward Understanding the Attraction Effect: The Implications of Product Stimulus Meaningfulness and Familiarity", Journal of Consumer Research, Vol 13 March (520-533)

S. RATNESHWAR, A.D. SHOCKER, 1988, "Substitution in Use and The Cognitive Structure of Product Categories", Working Paper, University of Minesota, November

D. REIBSTEIN , (1975), "An experimental Study of brand Choice and Switching Behavior", unpublished doctoral dissertation, Perdue University, West Lafayette, Indiana

A. RIES, J.TROUT, "Le Positionnement: la Conquête de l'Esprit", manager, McGraw-Hill Paris 1987.

J.R.B. RITCHIE, 1974, "An exploratory Analysis of the Nature and Extent of Individual Differences in Perceptions", Journal of Marketing Research, vol XI february (41-49).

J.R. RUBINSON, W.R. VANHONACKER, F.M. BASS, 1980, "On "A Parsimonious Description of the Hendry System", Management Science, vol 26 n° 2 (215-226).

G.J.RUSSELL, R.N. BOLTON, "Implications of Market Structure for Elasticity Structure", Journal of Marketing Research, vol XXV august (229-241).

SAS INSTITUTE, 1981, "The Adclus Procedure", in SAS Supplemental Library User's Guide, Cary, NC: Statistical Analysis Institute

SAS INSTITUTE, 1981, "The Alscal Procedure", in SAS Supplemental Library User's Guide, Cary, NC: Statistical Analysis Institute

S. SATTATH, A TVERSKY, 1977, "Additive Similarity Trees", Psychometrika vol 42 n° 3 (319-345).

S. SCHIFFMAN, I. REYNOLDS, F YOUNG, 1981, "Introduction to Multidimensional Scaling", Academic Press, INC, LONDON.

N. SCHMITT, B.W. COYLE, B. SAARI, 1977, "A Review and Critique of Analyses of Multitrait-Multimethod Matrices", Multivariate Behavioral Research, vol 12 october (447-478).

D.C. SCHMITTLEIN, A BEMMAOR, D.G. MORRISSON, 1985, "Why does the NBD Model Work? Robustness in Representing Product Purchases, Brand Purchases and Imperfectly Recorded Purchases", Marketing Science, vol 4 n° 3 (255-266).

R.N. SHEPARD, A.K. ROMMEY, S.B. NERLOVE, "Multidimensional Scaling Theory and Applications in the Behavioral Sciences", Volume 1 theorie, Seminar Press, 1972.

R.N. SHEPARD et P. ARABIE (1979), "Additive Clustering: Representation of Similarities as Combination of Discrete Overlapping Properties", Psychological Review, 86 (87-123).

R.N. SHEPARD, 1980, "Multidimensional Scaling, Tree-Fitting and Clustering", Science 210, (390-398)

A.D. SHOCKER, R.K. SRINIVASAN, 1974, "A Consumer-Based Methodology for the Location of New Product Ideas", Management Science, vol 20 n° 6 february (921-937).

A.D. SHOCKER, R.K. SRINIVASAN, 1979, "Multi-Attribute Application for Product-Concept Evaluation and Generation: A Critical Review" Journal of Marketing Research, 15, may (159-180)

A.D. SHOCKER, R.K. SRIVASTAVA, 1979, "Customer-Oriented Approaches to Identifying Product-Markets", Journal of Marketing, Vol 43 (8-19)

A.D. SHOCKER, A.J. ZAHORICK, D.W. STEWART, 1984, "Competitive Market Structure Analysis: A Comment on Problems", Journal of Consumer Research, vol 11 december (836-841).

A.D. SHOCKER, D.W. STEWART and A.J. ZAHORICK, 1987, "Market Structure Analysis : Practice, Problems, and Promise", Working Paper, September, University of Southern California, Los Angeles.

S.M. SHUGAN, 1987, "Estimating Brand Positioning Maps Using Supermarket Scanning Data", Journal of Marketing Research, vol XXIV february (1-18).

I. SIMONSON, 1989, "Choice Based on Reasons : The case of Attraction and Compromise Effects", Journal of Consumer Research, Vol 16 september, 1989 (158-174)

I. SIMONSON, 1990, The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior", Journal of Marketing Research, vol XXVII may (150-162).

J.Z. SISSORS, 1966, "What is a Market?", Journal of Marketing, vol 30 (17-21).

A. SOMIA, 1974, "Fondement Théorique de l'Analyse Multidimensionnelle des Concepts Nouveaux. Expériences Internationales", Revue Française du Marketing, n° 51 (55-68).

V. SRINIVASAN et SHOCKER, Allan (1973), "Linear Programming Techniques for Multi-dimensional Analysis of Preferences" *Psychometrika*, vol 38, (sept),.337-369

R.K. SRIVASTAVA, A.D. SHOCKER, G.S. DAY, 1978, "An Exploratory Study of the Influences of Usage Situation on Perceptions of Product-Markets", in Advances in Consumer Research, vol V, Ann Arbor: Association for Consumer Research (32-38)

R.K. SRIVASTAVA, 1980, "Usage-Situational Influences On Perceptions of Products-Markets: Response Homogeneity and its Implications for Consumer Research, in Advances in Consumer Research, vol 7, Chicago: Association for Consumer Research (644-649)

R.K. SRIVASTAVA, 1981, "Usage-Situational Influences On Perceptions of Products-Markets: Theoritical and Empirical Issues, in Advances in Consumer Research, vol 8, Chicago: Association for Consumer Research (106-111)

R.K. SRIVASTAVA, R.P. LEONE and A.D. SHOCKER, 1981, "Market Structure Analysis : Hierarchical Clustering of Products Based on Substitution-in-Use", Journal of Marketing, Vol 45 (38-47)

R.K. SRIVASTAVA, M. ALPERT and A.D. SHOCKER, 1984, "A Customer Oriented Approach for Determining Market Structures", Journal of Marketing, Vol 48 spring (32-45)

J.C. STANLEY, 1961, "Analysis of Unreplicated Three Way Classification with Applications to Rater Bias and Trait Independance" Psychometrika, vol 26(2105-219).V.J. STEFFLRE, 1971, "New Products and New Entreprises: A Report of an Experiment in Applied Social Science", Irvine, CA: University of California.

V.J. STEFFLRE, 1979, "New Products: Organizational and Technical Approaches to Product and Marketing Planning, A.D. SHOCKER, ed Cambridge, MA: Marketing Science Institute", (415-480).

Y. TAKANE, 1981, "Multidimensional Successive Categories Scaling: A Maximum Likelihood Method, Psychometrika, vol 46, n° 1 march (9-27).

Y. TAKANE, 1981, "Nonmetric Maximum Likelihood Multidimensional Scaling from Directional Ranking of Similarities", Psychometrika, vol 46 n° 4 december (389-405).

Y. TAKANE, 1983, "Choice Model Analysis of the "pick-any/n" type of binary data". Handout for the presentation at the European Meatings of the Psychometric and Classification Societies, July Jouy en Josas, France.

M. TENENHAUS, 1977, "Analyse en Composantes Principales d'un ensemble de variables Nominales ou Numériques", Revue de Statistique Appliquée, vol 25 n° 2 (39-56).

M. TENENHAUS, 1979, "La Regression Qualitative", Revue de Statistique Appliquée, vol 27 n° 2 (5-21).

M. TENENHAUS, 1983, "L'Analyse Canonique Généralisée de variables Numériques, Ordinales ou Nominales par des Méthodes de Codage Optimal", in Diday et Al "Data Analysis and Informatics", North Holland Publishing (71-84).

M. TENENHAUS, F.W. YOUNG, 1985, "An Analysis and Synthesis of Multiple Correspondance Analysis, Optimal Scaling, Dual Scaling, Homogeneity Analysis and other Methods for Quantifying Categorical Multivariate Data", Psychometrika, vol 50 n° 1 (91-120).

M. THENENHAUS, 1987, "Analyse des Données Historiques de la Qualité de l'Eau en France", Rapport de Recherche, Centre HEC-ISA, juillet.

L. THURSTONE, 1927, "A Law of Comparative Jugement", Psychological Review, vol 34 n° 4 july (272-286).

J.C. TOTTEN, T. ELROD, 1988, "A Complete System for Inferring Market Definition and Market Structure from Panel Data", Working Paper, Vanderbilt University, march, Nashville.

A. TVERSKY, 1972, "Elimination by Aspects: a Theory of Choice", Psychological Review, vol 34 n° 4 july (272-286).

A. TVERSKY, 1977, "Feature of Similarity", Psychological Review, vol 84 (327-352).
A. TVERSKY, S. SATTATH, 1979, "Preference Trees", Psychological Review, vol 86 (542-573).

G.L. URBAN, 1977, "Perceptor: un Modèle de Positionnement de Produits", Revue Française de Marketing, sept-oct (35-53).

G. URBAN , J.R. HAUSER, 1980, in Design and Marketing of New Products and Services, Englewood Cliffs, NJ: Prentice Hall.

G. URBAN, P.L. JOHNSON, R.H. BRUNWICK, 1981, "Market Entry Strategy Formulation", Working Paper 1103-80, Cambridge Ma, Sloan School of Management,MIT.

G. URBAN, P.L. JOHNSON, J.R. HAUSER, 1984,"Testing Competitive Market Structure", Marketing Science, vol 3 n°2, spring (83-112)

W.R. VANHONACKER, 1984, "Structuring and Analysing Brand Competition Using Scanner Data", Working Paper, september, Columbia University.

H. WAINER, S MESSICK, 1983, "Principals of Modern Psychological Measurement", Lawrence Erlbaum Associates, Publishers.

D.G. WEEKS, P.M. BENTLER, 1982, "Restricted Multidimensional Scaling Models for Asymmetric Proximities", Psychometrika, vol 47 n° 2 june, (201-208).

T.W. WHIPPLE, 1976, "Variation Among Multidimensional Scaling Solutions: An Examination of the Effect of Data Collection Differences", Journal of Marketing Research, vol XIII february (98-103).

B. WIERENGA, 1983, "Model and Measurement Methodology for the Analysis of Consumer Choice of Food Products", Journal of Food Quality vol 6 (119-137).

B. WIERENGA, 1984, "Empirical Test of the Lancaster Characteristics Model", International Journal of Research in Marketing, 1 (263-293).

R.E. WILKES, 1977, "Product Positioning by Multidimensional Scaling", Journal of Advertising Research, n°4 (15-19)

W.L. WILKIE, E.A. PESSEMIER (1973), "Issues in Marketing's Use of Multiattribute Attitude Models", Journal of Marketing Research, vol X november (428-441).

Y. WIND, P.J. ROBINSON, 1972, "Product Positioning: An Application of Multidimentional Scaling", in I. Haley (Ed), Attitude Research in Transition, Chicago: American Marketing Association.

J.K. WONG, J.N.SHETH, 1985, "Explaining Intention-Behavior Discrepancie: A Paradigm", ACR (378-384).

F.W. YOUNG, 1975, "An Asymmetric Euclidian Model for Multi processes Asymmetric data", Presented at US Japon Seminar on Multidimensional Scaling, University of California, San Diego, La Jolla.

G. ZALTMAN, C. PINSON, R. ANGELMAR, 1973, "Metatheory and Consumer Research", New York: Holt, Rinehart and Wintson.

J. ZINNES, D.B. MacKEY, 1983, "Probabilistic Multidimensional Scaling: Complete and Incomplete Data", Psychometrika, vol 48 n° 1 march (27-48).

A.A. ZOLTNERS et J.A. DODSON, 1983, "A Market Selection Model for Multiple End-Use Products", Journal of Marketing, vol 47 spring (76-88)

M. ZUCKERMAN, (1979), *"Sensation Seeking Scale: Beyond the Optimal Level of Aroudsal"*, Hillsdale, NJ: Lawrence Erlbaum

ANNEXES

PLAN DES ANNEXES

ANNEXES A LA PREMIERE PARTIE

ANNEXE 1: MESURE DE LA RECHERCHE DE VARIETE
ANNEXE 1A: INDEX OF TEMPORAL VARIETY (PESSEMIER ET HANDELSMAN 1984) 1
ANNEXE 1B: LE MODELE DAS (DYNAMIC ATTRIBUTE SATIATION McALISTER 1982) 3

ANNEXE 2: PANORAMA DES INDICES D'ASSOCIATION ENTRE DEUX OBJETS 6

ANNEXE 3: CACLUL DES TRANSFERTS THEORIQUES DANS LE MODELE DE HENDRY.

ANNEXES A LA DEUXIEME PARTIE

ANNEXE 1: DESCRIPTION DES MARCHES 9

ANNEXE 2: RESULTATS EMPIRIQUES RELATIFS AUX INDICES DE FRASER ET BRADFORD 17

ANNEXE 3:PROGRAMMES DE CALCULS DES INDICES ETUDIES 21

ANNEXE 4: DETAILS DES INDICES CALCULES SUR LES DIFFERENTS MARCHES ET LES DIFFERENTS UNIVERS EXTRAIT D'UN FICHIER DE CORRELATIONS 37

ANNEXE 5: TABLEAU DES MOYENNES DES CORRELATIONS PAR CATEGORIES DE CORRELATIONS 42

ANNEXE 6: ETUDE DE LA COHERENCE INTERNE DES INDICES: ANALYSES FACTORIELLES (coefficient de corrélation de Pearson). 44

ANNEXE 7: ETUDE DE LA COHERENCE INTERNE DES INDICES: ANALYSES FACTORIELLES (coefficient de rang de Spearman).

ANNEXE 8: MATRICE MULTITRAITS-MULTIMETHODES (coefficient de corélation de Pearson)

ANNEXE 9: ANALYSE DE LA VALIDITE DE TRAITS PAR L'ANALYSE FACTORIELLE (coefficient de corrélation de Pearson) 67

ANNEXE 10: ANALYSE DE LA VALIDITE DE TRAITS PAR L'ANALYSE FACTORIELLE (coerfficient de rang de Spearman) 69

ANNEXE 11: ANALYSE DE LA VALIDITE DE TRAITS: UTILISATION DE LA TYPOLOGIE HIERARCHIQUE ET DE L'ANALYSE DES SIMILARITES 71

ANNEXE 12 ANALYSE DE LA VALIDITE DE TRAITS PAR L'ANALYSE FACTORIELLE CONFIRMATOIRE 79

ANNEXE 14: ANALYSE DE LA VALIDITE DE TRAITS: UTILISATION DU MODELE LISREL 86

ANNEXE 15: ETUDE DE LA VALIDITE FACIALE DES INDICES 90

ANNEXE 16 ETUDE DES FAMILLES D'INDICES: MATRICES DE CORRELATIONS ETUDIEES 151

ANNEXE 17: ETUDE DES FAMILLES D'INDICES: ANALYSE DES SIMILARITES 155

ANNEXE 18: ANALYSE DES FAMILLES D'INDICES: ANALYSE TYPOLOGIQUE 167

ANNEXE 19: ANALYSE DES FAMILLES D'INDICES: ANALYSE FACTORIELLE 179

ANNEXE 20: EFFET DE LA MANIPULATION DES FACTEURS SUR LES INDICES 191

ANNEXE 21: EFFET DES FACTEURS SUR CHAQUE INDICE 193

ANNEXE 22: VALEURS DES CORRELATIONS INTER-FACTEURS POUR CHAQUE INDICE 197

ANNEXES A LA PREMIERE PARTIE

ANNEXE 1 MESURE DE LA RECHERCHE DE VARIETE

ANNEXE 1A INDEX DE VARIETE TEMPORELLE

Pessemier et Handelsman (1984)

Suivant les auteurs, l'indice de la variété temporelle est construit comme la somme pondérée de trois éléments:

$$ITV = w_1 PRD + w_2 PRE + w_3 RNB$$

-w_1, w_2 et w_3 sont les poids assignés à chacun des éléments de la mesure, ils peuvent caractériser :

- différents individus
- différentes situations
- différents sous-marchés ou différentes marques

- PRD est le pourcentage de dissimilarité réalisée
- PRE est le pourcentage d'entropie réalisée
- RNB est le non-groupage relatif dans la séquence d'achat étudiée

Ces trois éléments se déterminent de la manière suivante:

Soit:

- un ensemble de M marques disponibles sur un marché
- K attributs caractérisant ces marques, k=1 à K
- X_{kj}, le niveau perçu de la marque i selon l'attribut k (donnée de jugement collectée par questionnaire)
- d_{ij} la distance perceptuelle entre les marques i et j

$$d_{ij} = \left[\sum_{k=1}^{K} (X_{ki} - X_{kj})^2\right]^{1/2}$$

- D_0 la somme des distances entre les m marques achetées sur la période (m < M)
- D_{max} la somme des distances entre les m marques de M les plus dissimilaires

$PRD = D_0/D_{max}$, PRD appartient à l'intervalle [0,1]

Soit:

- p_i la proportion observée d'achat de la marque i (donnée de comportement)
- N le nombre d'occasions d'achats
- H_0 l'entropie observée

$$H_0 = -\sum_{i=1}^{m} p_i \ln(p_i)$$ (ln signifie le logarithme népérien)

-H_{max}, l'entropie maximum, $H_{max} = \sum_{i=1}^{m} (1/N^*) \ln(1/N^*)$ avec

$N^* = \min(M,N)$

$PRE = H_0/H_{max}$, PRE appartient à l'intervalle [0,1]

Soit:

O_N le nombre de changements contigus

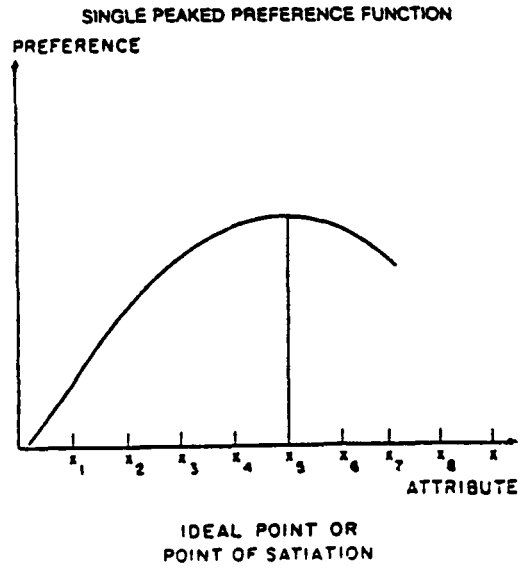
$RNB = O_N/(N-1)$, RNB appartient à l'intervalle [0,1]

Par exemple,

pour l'historique d'achat 01100111:

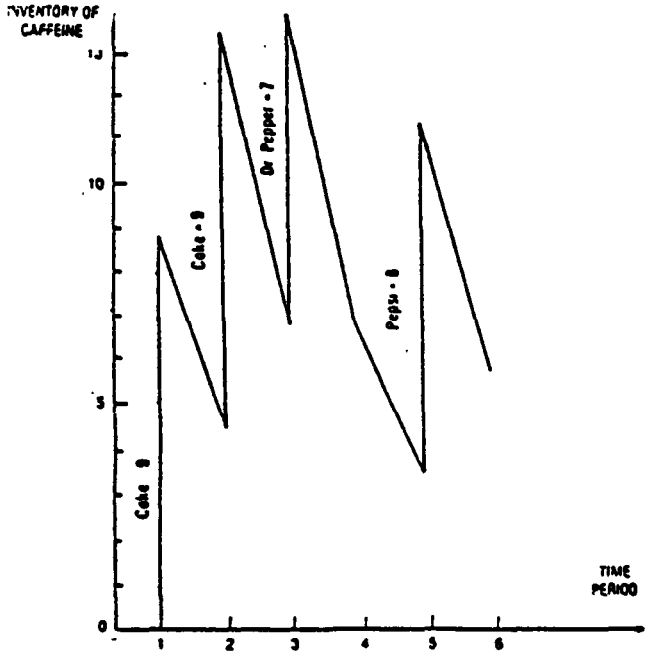
$O_N=3$, N=8 et RNB = 3/7 = 0,43

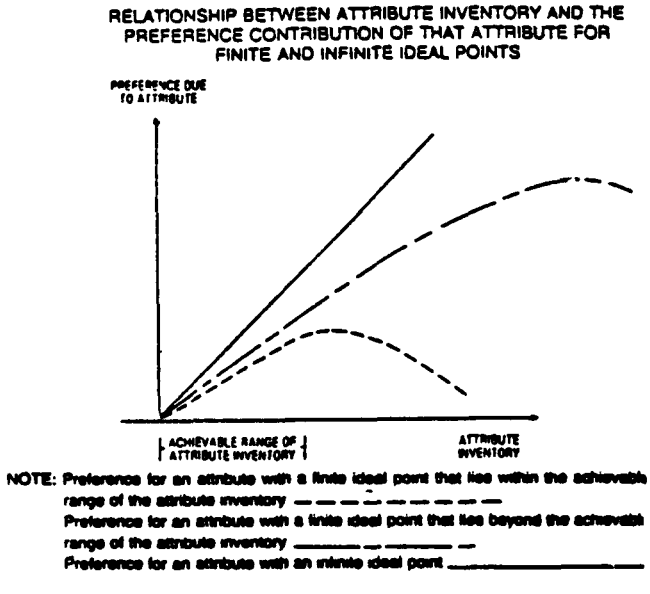
FONCTION DE PREFERENCE ET INVENTAIRE CUMULE D'UN ATTRIBUT



"Extrait de McAlister (1982) p 143"

- Xj (X5 sur la figure précédente), le niveau idéal cumulé (à la différence des modèles de préférence classiques), attendu sur l'attribut j, en tenant compte des quantités de cet atttribut consommées dans les périodes précédentes (à travers la consommation de k ou de tout autre produit possédant cet attribut)





"Extrait de McAlister et Pessemier (1982) p143"

ANNEXE 1B MODELE DAS (Dynamic Attribute Satiation) McAlistair 1982

Soit :

- K produits appartenant à un marché, k=1 à K
- J , le nombre d'attributs caractérisant chaque produit, j=1 à J
- t une occasion d'achat particulière, t=1 à T
- DAS_{kt}, la préférence pour le produit k selon le modèle DAS, à l'occasion t

$$DAS_{kt} = \sum_{j=1}^{J} P_{kjt}$$

P_{kjt} est la contribution de l'attribut j à la satisfaction résultant de la consommation du produit k à l'occasion t

P_{kjt} est fonction de l'inventaire cumulé de l'attribut j à l'occasion t: I_{jt} , c'est la somme, pondérée par un taux de rémanence, des quantités consommées de cet attribut j, à l'instant t .

$P_{kjt} = w_j [(I_{jt} + X_{kj}) - X_j)^2] (-1)$

- w_j est l'importance de l'attribut j
- I_{jt} est l'inventaire de l'attribut j en t
- X_{kj} est la quantité perçue de l'attribut j dans le produit k
- X_j est le niveau idéal cumulé (à la différence des modèles de préférences classiques), attendu sur l'attribut j, en tenant compte des quantités de cet attribut qui ont été consommées lors des périodes précédentes , à travers la consommation du produit k ou de tout autre produit possédant cet attribut.j

I_{jt} peut donc s'écrire:

$$I_{jt} = \sum_{t=1}^{T} m_j^{T-t} X_{kjt}$$

m_j est le facteur de rétention de l'attribut j

m_j appartient à l'intervalle (0,1) pour tout attribut k

X_{kjt} est la quantité de l'attribut j dans l'item k choisi à l'instant t

Collecte de l'information et estimation

- les X_{kj} sont collectés par questionnaire
- on observe l'historique d'achat sur les t = 1 à T occasions d'achat
- il est nécessaire de fournir une valeur pour m_j, taux de rémanence de la consommation passée d'une caractéristique k en quantité X_{kjt}

L'estimation est effectuée en utilisant le modèle LINMAP (Srinivasan et Shocker 1973). L'intérêt conceptuel majeur de cette approche est de considérer que pour la prévision des préférences, l'importance des attributs ne suffit pas. Il y a donc deux éléments qui participent à la formation de la préférence pour un produit, à partir de sa perception selon les attributs déterminants: le poids des attributs (approche classique) et l'inventaire de ces attributs à l'instant t. McAlistair a ainsi montré que le pourcentage de variance expliquée des préférences est significativement supérieur lorsque l'on inclut dans l'analyse ce deuxième élément. On notera cependant que la procédure d'estimation est lourde et que les valeurs de m_j doivent être fournies arbitrairement.

ANNEXE 2 PANORAMA DES INDICES D'ASSOCIATION ENTRE DEUX OBJETS

- Indices de proximité définis sur des matrices logiques.

DENOMINATEUR	NUMERATEUR	
1) Avec co-absences	Sans co-absences	Avec co-absences
a- Egale pondération des coïncidences et des non-coïncidences	RUSSEL et RAO (1940) (1) $\frac{P}{T}$	SOKAL et MICHENER (1958) (2) $\frac{P+A}{T}$
b- Pondération double des coïncidences	(3) $\frac{2P}{2(P+A)+N}$	(4) $\frac{2(P+A)}{2(P+A)+N}$
c- Pondération double des non-coïncidences	(5) $\frac{P}{P+A+2N}$	ROGER et TANIMOTO (1960) (6) $\frac{P+A}{P+A+2N}$
2) Sans co-absences	JACCARD (1908)	
a- Egale pondération des coïncidences et des non-coïncidences	(7) $\frac{P}{P+N}$	non sens
b- Pondération double des coïncidences	DICE (1945) (8) $\frac{2P}{2P+N}$	non sens
c-Pondération double des non-coïncidences	SOKAL et SNEATH (1963) (9) $\frac{P}{P+2N}$	non sens
d- Pondération nulle des coïncidences	KULCZYNSKI (1927) (10) $\frac{P}{N}$	(11) $\frac{P+A}{N}$
e- Pondération des coïncidences par le nombre d'attributs possédés par les 2 objets	KULCZYNSKI 1927) (12) $\frac{1}{2}\left(\frac{P}{\|X\|}+\frac{P}{\|Y\|}\right)$ OCHIAI (1957) (14) $\frac{P}{\sqrt{\|X\|\,\|Y\|}}$	(13) $\frac{1}{4}\left(\frac{P+A}{\|X\|}+\frac{P+A}{\|Y\|}\right)$

P = co-présences
A = co-absences
N = non coïncidences

|X|= nombre d'attributs possédés par l'objet X
|Y|= nombre d'attributs possédés par l'objet Y
T = nombre total d'attributs

Les proximités utilisables sur des matrices de fréquence ou d'occurrence.

Distance du Chi Deux

$$F1 : d_{ik}^2 = \sum_{j=1}^{p} \frac{1}{p.j}\left(\frac{p_{ij}}{p_{i.}} - \frac{p_{kj}}{p_{k.}}\right)^2$$

Distance de Jambu

$$F2 : d_{ik}^2 = \sum_{j=1}^{p} \frac{1}{(p_{ij}+p_{kj})}\left(\frac{p_{ij}}{p_{i.}} - \frac{p_{kj}}{p_{k.}}\right)^2$$

Information mutuelle

$$F3 : S_{ik} = \sum_{j=1}^{p}\sum_{i=1}^{n} p_{ji} \log_2\left[\frac{p_{ji}}{p_{j.}.p_{.i}}\right]$$

Distance de RAJSKI

$$F4 : d_{ik} = \frac{\sum_{j=1}^{p}\sum_{i=1}^{n} p_{ji}\,(\log_2 p_{j.} + \log_2 p_{.i} - 2\log_2 p_{ji})}{-\sum_{j=1}^{p}\sum_{i=1}^{n} p_{ji}\log_2 p_{ji}}$$

Coefficient de cohérence de RAJSKI (similarité)

$$F5 : R_{ik} = \left[1 - d_{ik}^2\right]^{1/2}$$

"Extrait de Pinson et Chandon 1980, p 67 et 74"

ANNEXE 3 CALCUL DU TRANSFERT THEORIQUE DANS LE MODELE DE HENDRY

Le modèle de Hendry fournit la valeur théorique de TS, le changement total au sein d'un marché.

$$TStheo = \sum_{i=1}^{g} \frac{-pi.\ln(pi)}{1 - pi.\ln(pi)} \quad (3)$$

Cette valeur est obtenue (voir Bass et al 1980 en annexe) en exprimant le comportement de N consommateurs aléatoires sans préférences particulières concernant g marques. Le nombre de possibilités de répartir les choix sur ces g marques selon le vecteur de choix (n1, n2, ...ng) est alors :

$$N! / n1! \,.\, n2! \,...ng!$$

La probabilité d'une répartition particulière des choix est donnée par :

$$Prob(n1,n2,..ng) = 1 / (N! / (n1! \,.\, n2! \,...ng!)) \quad (4)$$

qui est la probabilité qu'une distribution particulière de (N , g , {ni}) se réalise.
L'incertitude relative à cette distribution est directement liée à N! / (n1! . n2! ...ng!) , soit monotoniquement à :

Ln (N! / (n1! . n2! ...ng!)), facteur qui est appelé "Entropie" en thermodynamique .

L'espérance de cette distribution (pour un consommateur moyen), est donnée par :

$$[\ Prob(n1,...,ng)\ ]^{1/N}$$

Lorsque N tend vers l'infini, ce qui est généralement le cas:

$$\lim [\ Prob(n1,...,ng)\ ] \rightarrow \prod_{i=1}^{g} p(\pi i) = \prod_{i=1}^{g} \pi i$$

P(πi) probabilité qu'un consommateur moyen choisisse i
p(πi) = exp(πi.ln(πi))

En posant πi = ui + vi

où (v1....vg) est un vecteur de macro-préférences (habitudes d'achats)
(u1....ug) est un vecteur d'incertitude

On aboutit à un système de choix avec préférences :

$$vi = \pi i/(1 - \pi i.\ln(\pi i))$$

Or, comme $\sum_{i=1}^{g} ui = - \sum_{i=1}^{g} vi.\pi i.\ln(\pi i)$ (6)

$$ui = - vi.\ln(\pi i) \qquad (7)$$

(7) dans (6) donne finalement (3)

ANNEXES A LA DEUXIEME PARTIE

ANNEXE 1 DESCRIPTION DES MARCHES

I MARCHE DES HUILES ALIMENTAIRES

α SELECTION DES MARQUES

Les dix marques retenues sur ce marché ont été sélectionnées parmi les items disponibles dans la base de données SCAN 5000 qui, pour les huiles, en comportait 185. Le choix des marques a été opéré à partir de l'avis d'experts de ce marché. Seuls les attributs "marque" et "parfum" ont été conservés. Par contre, l'attribut contenance n'a pas été pris en compte, les différents contenants d'une même marque sont donc agrégés. Nous sommes conscients du rique de réduction du phénomène concurrentiel que ceci peut présenter. L'item 'autres marques' est constitué (dans la base SCAN 5000) de l'agrégat de plusieurs marques aux parts de marché faibles.

β TABLEAU DES PARTS DE MARCHE EN ACTES D'ACHATS

CODE	MARQUE	NOMBRE D'ACTES	PART DE MARCHE EN ACTES
1	LESSIEUR TOURNESOL	1989	14.7
2	LESSIEUR ARRACHIDE	638	4.7
3	LESSIEUR MAIS	313	2.3
4	LESSIEUR OLIVE	155	1.1
5	FRUIT D'OR	1976	14.6
6	ECLAT D'OR	52	.4
7	EPI D'OR	77	.6
8	PUJET	44	.4
9	AUTRES MARQUES	2228	16.5
10	MARQUES DISTRIBUTEURS	6178	45.7

Γ FICHE TECHNIQUE

ACTES D'ACHATS	13507
PANELISTES	2295
NOMBRE MOYEN D'ACHATS	30.8 (28.6)
QUANTITE MOYENNE ACHETEE	26.9 (19.7)

π TRIS SUR LES VARIABLES

QUANTITE CONSOMMEE

	QUANTITE	% ACTES
PETIT	< 4.7	33.6
MOYEN	] 4.7,13]	33.8
GROS	> 15	32.6

FREQUENCE D'ACHAT

	FREQUENCE	% ACTES
PETIT	< 16	33.5
MOYEN	] 16,30]	33.1
GROS	> 30	33.4

TAILLE DU FOYER

	TAILLE	% ACTES
PETIT	1 et 2	34.1
MOYEN	3	32.6
GROS	4 et plus	34.3

REVENU

	REVENU	% ACTES
PETIT	< 6000	30.1
MOYEN	]6000,9000]	32.2
GROS	> 9000	37.7

MAGASIN

	CATEGORIE	% ACTES
PETIT	1	31.1
MOYEN	2	51.2
GRAND	3	17.7

CSP

	CATEGORIE	% ACTES
CSPS	4 et 5	28.3
CSP6	6	35
CSP7	7	23

II MARCHE DES BOISSONS GAZEUSES (CARBONEES)

α SELECTION DES MARQUES

Les quatorze marques retenues sur ce marché ont été sélectionnées parmi les 378 items de la base de données SCAN 5000 relative à ce segment de marché (qui n'est qu'un sous marché du marché des Softs Drinks qui lui est beaucoup plus large).

Le choix des marques a été opéré à partir de l'avis d'experts de ce marché. Par rapport au marché des carbonnés tel qu'il est disponible dans la base SCAN 5000, nous n'avons pas opéré d'agrégations selon un facteur particulier.

Par contre, certaines marques aux parts de marché trop faibles ont été abandonnées (SPRITE par exemple).

β TABLEAU DES PARTS DE MARCHE EN ACTES D'ACHAT

CODE	MARQUE	NOMBRE D'ACTES	PART DE MARCHE EN ACTES
1	ORANGINA STANDARD	1298	10.8
2	INDIAN TONIC STANDARD	858	7.2
3	CANADA DRY	876	7.3
4	GINI	353	2.9
5	RICQLES	218	1.8
6	COCA COLA STANDARD	2468	20.6
7	COCA COLA SANS CAFEINE	997	8.3
8	PEPSI COLA	258	2.1
9	SEVEN UP	466	3.9
10	LIMONADES VERRE	1470	12.3
11	LIMONADES PET	1300	10.9
12	PERRIER ZESTE	350	3
13	BRUT DE POMME	367	3
14	AUTRES BAF ET SODAS	1860	15.5

Γ FICHE TECHNIQUE

ACTES D'ACHATS	12000	
PANELISTES	1939	
NOMBRE MOYEN D'ACHATS	13.5	(26.2)
QUANTITE MOYENNE ACHETEE	13.5	

π TRIS SELON LES PRINCIPALES VARIABLES

QUANTITE CONSOMMEE

	QUANTITE	% ACTES
PETIT	< 4.6	33.8
MOYEN	]4.6,14.4]	32.2
GROS	> 14.4	34

TAILLE DU FOYER

	TAILLE	% ACTES
PETIT	1 et 2	35.1
MOYEN	3	32.6
GROS	4 et plus	32.5

REVENU

	REVENU	% ACTES
PETIT	< 6000	34.2
MOYEN	]6000,9000]	29.9
GROS	> 9000	35.9

REVENU MOYEN 8557 (4384)

CSP

	CATEGORIE	% ACTES
CSPS	4,5	27.5
CSP6	6	32.8
CSP7	7	16.4

MAGASIN

	CATEGORIE	% ACTES
PETIT	1,2,7,8,9	43.2
MOYEN	3,6,0	39.6
GRAND	4,5	17.2

III MARCHE DES BOISSONS AUX FRUITS GAZEUSES ET SODAS (PARFUM)

α LE CHOIX DES MARQUES

Les quinze items retenus sur ce marché ont été sélectionnés parmi les 378 (voir Annexe 2) que compte la base de données SCAN 5000 relative à ce segment de marché des Boissons aux Fruits Gazeuses et sodas (qui n'est qu'un sous marché du marché des Softs Drinks, qui lui est plus large).

Le choix des marques a été opéré à partir de l'avis d'experts de ce marché. Par rapport au marché des BAF et SODAS tel qu'il est disponible dans la base SCAN 5000, nous n'avons pas opéré de simplifications particulières, exeptée la suppréssion de certains items aux parts de marché trop faibles et qui correspondent à des parfums peu vendus comme Framboise & Raisin, Pomme ...

Par rapport au marché précédent où on ne dispose que de la marque, ce marché permet de prendre en compte l'aspect Recherche de Variété entre les parfums qui est certainement un aspect non négligeable de la concurrence entre les marques .

β TABLEAU DES MARQUES ET DE LEURS PARTS DE MARCHE EN ACTES D'ACHAT

CODE	MARQUE	NOMBRE D'ACTES	PART DE MARCHE EN ACTES
1	DRY ORANGE	364	3.1
2	DRY CITRON	250	2.1
3	DRY EXOTIQUE	230	2
4	DRY PAMPLEMOUSSE	47	0.4
5	ORANGINA	4993	42.5
6	FANTA CITRON	407	3.5
7	FANTA ORANGE	885	7.5
8	FANTA EXOTIQUE	207	1.8
9	FRUITE ORANGE	307	2.6
10	FRUITE POMME & CASSIS	98	0.8
11	FRUITE ANANAS PAMPL	41	0.3
12	DISTRIBUTEUR ORANGE	819	7
13	DISTRIBUTEUR CITRON	398	3.4
14	AUTRES MARQUES ORANGE	1650	14
15	AUTRES MARQUES CITRON	1036	8.8

Γ FICHE TECHNIQUE

ACTES D'ACHATS	11109	
PANELISTES	1939	
NOMBRE MOYEN D'ACHATS	8.25	(16.5)
QUANTITE MOYENNE ACHETEE	14.2	(26.3)

π TRIS SUR LES VARIABLES PRINCIPALES

QUANTITE CONSOMMEE

	QUANTITE	% ACTES
PETIT	< 3.2	34.5
MOYEN	]3.2,10]	32.5
GROS	> 10	33

TAILLE DU FOYER

	TAILLE	% ACTES
PETIT	1 et 2	23.6
MOYEN	3	21
GROS	4 et plus	55.4

REVENU

	REVENU	% ACTES
PETIT	< 6000	36.6
MOYEN	]6000,9000]	32.8
GROS	> 9000	30.6

MAGASIN

	CATEGORIE	% ACTES
PETIT	1,2,7,8,9	37
MOYEN	3,6,0	46.3
GROS	4,5	16.6

CSP

	CATEGORIE	% ACTES
CSPS	4 et 5	15.5
CSP6	6	40.7
CSP7	7	15.5

IV MARCHE DES SHAMPOOINGS (SHAMP)

α LE CHOIX DES MARQUES

Les vingt quatre items retenus sur ce marché ont été sélectionnés parmi les 449 (voir annexe 2) que compte la base de données SCAN 5000 relative au marché des Shampooings .

Le choix des marques a été opéré à partir de l'avis d'experts de ce marché . Par rapport au marché des Shampooings tel qu'il est disponible dans la base SCAN 5000, nous n'avons pas conservé le segment des Shampooings Douche qui contient un trop grand nombre de références, assez peu diffusées . D'autre part, les seuls attributs conservés sont la marque et le fait que le produit ait ou non une action anti-pelliculaire . Les attributs Contenance, Type de Cheveux (gras, normaux , secs) et Parfum du produit n'ont pas été retenus, ce qui, encore une fois, présente le risque de gommer une partie de la concurrence entre les marques . D'autre part, les marques aux parts de marché trop faibles ont été agrégées dans autres .

β TABLEAU DES MARQUES ET DE LEURS PARTS DE MARCHE EN ACTES D'ACHAT

CODE	MARQUE	NOMBRE D'ACTES	PART DE MARCHE EN ACTES
1	MIXA	746	6.3
2	BABYDOP	319	2.7
3	BABIVEA	174	1.5
4	ULTRADOUX	1475	12.6
5	FLOREAL	336	2.9
6	TIMOTEI	851	7.2
7	NIVEA	345	2.9
8	DIMENSION	306	2.6
9	ELSEVE	855	7.9
10	LONGUEURS ET PONTES	713	6.1
11	SOYANCE HAP *	82	0.7
12	ENERGANCE	337	2.9
13	PETROL HAHN	881	7.5
14	MOELLE ET GARNIER HAP	111	0.9
15	DOP HAP	277	2.4
16	PALMOLIVE HAP	283	2.4
17	EQUILIBRE	175	1.5
18	HEAD & SHOULDERS	351	3
19	CLEAR	371	3.2
20	TOTAL DISTRIBUTEURS	549	4.7
21	SOYANCE AP **	346	2.9
22	MOELLE ET GARNIER AP	217	1.8
23	DOP AP	1260	10.7
24	PALMOLIVE AP	1101	9.4

* HAP : HORS ANTIPELLUCULAIRE
** AP : ANTIPELLICULAIRE

β FICHE TECHNIQUE

ACTES D'ACHATS	11740
PANELISTES	2679
NOMBRE MOYEN D'ACHATS	9.62 (8.3)
QUANTITE MOYENNE ACHETEE	10.45 (8.7)

Γ TRIS SELON LES VARIABLES PRINCIPALES

QUANTITE CONSOMMEE

	QUANTITE	% ACTES	% INDIVIDUS
PETIT	< 5	33.1	66.7
MOYEN	]5 ,11]	34.6	17.2
GROS	> 11	32.3	16.1

FREQUENCE D'ACHAT

	QUANTITE	% ACTES
PETIT	< 8.2	36.5
MOYEN	]8.2 10]	30.7
GROS	> 10	32.8

TAILLE DU FOYER

	TAILLE	% INDIVIDUS	% ACTES
PETIT	1 et 2	45.5	34.2
MOYEN	3	28.8	32.8
GROS	4 et plus	8.1	33

MAGASIN

	CATEGORIE	% ACTES	% INDIVIDUS
PETIT	1,2,7,8,9	37.5	40
MOYEN	3,6,0	32.2	33.9
GRAND	4,5	30.3	26.1

ANNEXE 2 RESULTATS EMPIRIQUES RELATIFS AUX INDICES DE FRASER ET BRADFORD

Nous présenterons dans cette annexe les résultats concernant deux points particuliers à cet indice :

- l'hypothèse $E(Xii) \geq E(Xij)$
- l'estimation des coefficients de forme de la loi gamma des Xij, les ri

Concernant ces deux points, nous présenterons les résultats pour les deux versions que nous avons calculé :

- BRAA qui ne tient pas compte pour l'estimation des Xij des autres marques. Il n'y a pas de définition d'un univers de concurrence pour le calcul de cet indice. C'est la définition de l'indice telle qu'elle figure dans l'article originel de Fraser et Bradford.

- FRASA qui tient compte d'un univers de concurrence préalablement défini afin d'estimer les Xij.

1 ETUDE DE L'HYPOTHESE $E(Xii) \geq E(Xij)$

Le modèle de Fraser et Bradford étant bâti autour de cette hypothèse, il était important de la valider empiriquement. Pour cela nous avons calculé, sur l'ensemble des couples de marques, le pourcentage des couples vérifiant l'hypothèse posée par les auteurs. Les résultats figurent, pour les quatre marchés, dans le tableau suivant (en pourcentage de relations vérifiées).

	HUILE	BOISSON	PARFUM	SHAMP
BRAA	88	80	68	-*
FRASA	83	77	91.5	93

* L'indice BRAA n'a pas été calculé sur le marché SHAMP .

Empiriquement, on peut donc constater que la relation posée par Fraser et Bradford n'est pas toujours vérifiée, même si elle l'est dans la plus part des cas.
D'autre part, on n'observe pas de différence entre les deux versions de l'indice quant au degré de vérification de cette hypothèse.

2 ETUDE DE LA VALEUR DES COEFFICIENTS DE FORME ri

La valeur du coefficient de forme est importante car c'est de la qualité de son estimation que dépent en grande partie la qualité de l'estimation de l'indice de concurrence. Nous présenterons pour chaque marché les valeurs obtenues pour ce coefficient sur chacun des quatre marchés.

α MARCHE HUILE

Mi	PdMi	Ri1	Ri2
1	14.7	1.72	2.23
2	4.7	1.71	2.07
3	2.3	1.76	2.33
4	1.1	2.53	4.26
5	14.6	1.64	2.57
6	0.4	3.3	8.55
7	0.6	3.2	3.8
8	0.5	-	-
9	16.5	0.95	1.9
10	45.7	1.22	1.5

Mi Indice de la Marque
PdMi Part de marché de la marque
Ri1 Valeur du coefficient de forme pour l'indice BRAA
Ri2 Valeur du coefficient de forme pour l'indice FRASA

β MARCHE BOISSON

Mi	PdMi	Ri1	Ri2
1	10.8	1.51	2.23
2	7.2	1.71	1.86
3	7.3	2.01	2.97
4	2.9	2.08	2.1
5	1.8	2.52	2.86
6	20.6	1.38	1.88
7	8.3	1.26	2.13
8	2.1	1.16	1.6
9	3.9	1.76	2.49
10	12.3	1.27	1.65
11	10.9	1.1	1.68
12	3.0	1.42	1.3
13	3.0	1.59	2.49
14	15.5	1.32	1.71

Γ MARCHE PARFUM

Mi	PdMi	Ri1	Ri2
1	3.1	1.37	1.3
2	2.1	1.21	2
3	2.0	1.70	1.8
4	0.4	1.53	3
5	42.5	1.03	1.5
6	3.5	1.27	1.3
7	7.5	0.81	1.8
8	1.8	1.1	3.2
9	2.6	2.57	1.8
10	0.8	1.33	1.4
11	0.3	1.4	1.6
12	7.0	1.45	1.15
13	3.4	-	1.48
14	14.0	-	1.24

π MARCHE SHAMP

Mi	PdMi	Ri2
1	6.3	2.7
2	2.7	4.0
3	1.5	3.5
4	12.5	3.0
5	2.9	3.1
6	7.2	2.8
7	2.9	3.6
8	2.6	3.9
9	7.9	3.4
10	6.1	3.3
11	0.7	31.7
12	2.9	2.4
13	7.5	2.7
14	0.9	3.2
15	2.4	3.5
16	2.4	8.0
17	1.5	3.3
18	3.0	6.0
19	3.2	3.3
20	4.7	4.2
21	2.9	4.4
22	1.8	4.0
23	10.7	3.1
24	9.4	3.0

Après analyse de ces tableaux, on retiendra les éléments suivants :

- plus les parts de marché sont faibles et plus les les valeurs de ri sont élevées. Ceci s'explique par l'augmentation du nombre d'individus pour lesquels les intervalles Xij prennent une valeur égale à zéro, ce qui augmente alors la stabilité apparente du comportement agrégé à l'égard de ces deux marques. Ceci constitue une limite à l'utilisation du modèle lorsque la part de marché d'un produit est faible, ce qui est le cas pour un produit nouveau. En effet, plus la part de marché devient faible et plus la valeur de l'indice de concurrence entre deux marques dépend du coefficient de normalisation (ri -1)/2*ri . On a pu dailleurs constater que l'utilisation de ce coeficient de normalisation n'aboutit pas aux effets escomptés à savoir l'obtention d'un indice compris entre 0 et 1.
C'est pour cette raison que nous avons calculé les versions de BRAA et FRASA sans utiliser ce coefficient de normalisation (respectivement les indices BRAB et FRASB).

- on peut constater que les coefficients de forme ri prennent des valeurs plus élevées quand on définit l'univers de concurrence (c'est à dire pour le calcul de l'indice FRASA). Ceci est logique puisque ce faisant on introduit un nombre plus important de marques dans l'historique d'achat du panéliste ce qui diminue d'autant le nombre de transferts entre chaque couple de marques i et j. Le nombre d'individus pour lesquels la valeur Xij est nulle s'acroit, ce qui augmente la stabilité apparente des comportements agrégés relativement aux deux marques. Ainsi on observe que Ri1 est inférieur à Ri2.

ANNEXE 3 PROGRAMMES DE CALCUL DES INDICES ETUDIES (*)

1) PROGRAMME DE CALCUL DE L'INDICE DE FRASER ET BRADFORD, DANS SA FORME ORIGINENELLE (INDICES BRAA ET BRAB).

```
********** CALCUL DE L'INDICE DE FRASER ET BRADFORD
******************;
*****************************************************************

************ CE FICHIER EST TRIE PAR PANEL DATUA ;
******* SASIN1 EST LA BASE TRIEE SELON PANEL PUIS DATUA ;

DATA BIDON2 ;
  INFILE SASIN  ;
     INPUT CODEM 1-2   datua 4-9     PANEL 12-17
        +1 (QR1 -QR14)(5.3 );
     INPUT;
      ld =lag(datua) ;
      IF QR1  NE 0 OR QR6   NE 0 ;
DATA BIDON1; SET BIDON2; BY PANEL;
        ARRAY QR( 14) QR1-QR14 ;
        ARRAY SQR(14 ); RETAIN SQR1-SQR14 0 ;
            do i=1 to 14 ;
            if first.PANEL THEN SQR(I)=QR(I);
            ELSE  SQR(I)=SUM(SQR(I),QR(I));
            end;
            if last.panel then output bidon1;
            KEEP  PANEL sqr1-SQR14  ;

DATA bidon3; merge bidon1 bidon2 ; by panel ; quant=sqr1+sqr6
            ;
            IF  quant   gt  5 then output bidon3 ;
DATA BIDON1 ; SET BIDON3 ; BY PANEL ;
        ARRAY LQ  (14 );
        ARRAY QR  (14 )QR1-QR14 ;
        ARRAY LQTR(14 ) LQR1 -LQR14 ;
             ld=lag(datua);
             LPANEL=LAG(PANEL); RETAIN  LPANEL 0;
             RETAIN  LQR1 -LQR14  0;
             LQR1 =LAG(QR1);LQR2 =LAG(QR2 )  ;
             LQR3 =LAG(QR3 );LQR4 =LAG(QR4 );LQR5 =LAG(QR5 );
             LQR6 =LAG(QR6 )  ;LQR7 =LAG(QR7 )  ;LQR8
             =LAG(QR8 ); LQR9 =LAG(QR9 )  ;LQR10=LAG(QR10)
             ;LQR11=LAG(QR11)  ; LQR12=LAG(QR12)
             ;LQR13=LAG(QR13)  ;LQR14=LAG(QR14)  ; RETAIN
             LQ1-LQ14 0;
                IF PANEL NE LPANEL THEN  DO; DO k=1 TO 14; LQ
                (k)=0; END;
                       END;
                IF PANEL=   LPANEL  THEN DO;
                DO k=1 TO 14;   LQ (k)=LQTR(k) ; END;
```

```
                    END;
          KEEP LQ1-LQ14  QR1 -QR14 PANEL DATUA ld ;

DATA BIDON2;SET BIDON1 ; BY PANEL ;
        ARRAY QR(14 ) QR1-QR14 ;
        ARRAY LQ(14) LQ1-LQ14 ; RETAIN LQ1-LQ14 0 ;
        ARRAY  X(4) ; RETAIN X1 X2 X3 X4 0 ;
        ARRAY XB(4); RETAIN XB1 XB2 XB3 XB4 0 ;
        ARRAY moy(4); RETAIN moy1-moy4 0 ;
        ARRAY  V(4); RETAIN V1 V2 V3 V4  0 ;
        ARRAY SC(4); RETAIN SC1 SC2 SC3 SC4 0 ;
        ARRAY M(4); RETAIN M1 M2 M3 M4 0 ;
        RETAIN    P r1 r3 RI1 RI3  FRAS 0 ;
        I=1 ; J=6 ;
        IF QR(I) NE 0 AND LQ(I) NE 0 THEN DO ;   X1=(DATUA-
        LD)/lq(i); XB1  = sum(XB1,X1)   ;
        M1=M1+1 ;
        SC1=sum(SC1,X1**2) ;
        END ;
        IF QR(J) NE 0 AND LQ(I) NE 0 THEN DO ;
        X2=(DATUA-LD)/lq(i);
        XB2  =sum( XB2, X2)   ;
        M2  =M2+1 ;
        SC2=sum(SC2,X2**2) ;
        END ;
        IF QR(J) NE 0 AND LQ(J) NE 0 THEN DO ;
        X3=(DATUA-LD)/lq(j); XB3  =SUM( XB3, X3)   ;
        M3  =M3+1 ;
        SC3=SUM(SC3,X3**2) ;
        end;

        IF QR(I) NE 0 AND LQ(J) NE 0 THEN DO ; X4=(DATUA-
        LD)/lq(j); XB4 =sum(xb4, X4)  ;
        M4  =M4+1 ;
        SC4=sum(SC4,X4**2) ;
        END ;
        IF LAST.PANEL THEN DO ;
        DO K=1 TO 4 ;
          if m(k) lt 2 then do ;
             MOY(K)=0 ;V(K)=0 ;  END ;
             ELSE DO ;
             moy(K)=XB(K)/M(K) ;
             V(K)=(SC(K)/M(K))  -(MOY(K)**2)  ;
                end ;
                end ;
         do k=1 to 4 ;m(k)=0;xb(k)=0;sc(k)=0;x(k)=0; end ; if
         v1=0 then r1=0 ;
         else  r1= v1/(moy1**2)  ;
         IF V3=0 THEN R3=0 ;
         ELSE
          r3=v3/(moy3**2)  ;
          P=P+1 ;
          output bidon2 ;
          end ;
```

```
          drop qr2-qr5 lq2-lq5 qr7-qr14 lq7-lq14 ;

data bidon1 ; set bidon2 end=last ;
      array xbb(4 ) ;retain xbb1-xbb4 0 ;
      array moy(4 ) moy1-moy4 ;
      array pop(4);  retain pop1-pop4 0 ;
      array rot(2) r1 r3 ; retain rii1 rii3 0 ;
      array parm(2) rii1  rii3 ;
          do k=1 to 2 ;
     parm(k)=sum(parm(k),rot(k));
     end ;
          do k=1 to 4 ; if moy(k) ne 0 then pop(k)=pop(k)+1 ;
          xbb(k)=sum(xbb(k),moy(k)) ; end ;
      if last then do ;
          DO K=1 TO 4 ;
          XBB(K)=XBB(K)/Pop(k)    ; END ;
          RIi1  =( RIi1/Pop1)**-1 ;
          RIi3  =( RIi3/Pop3 )**-1 ;
          IF rii1 =0 AND rii3 =0 THEN FRAS=0 ;
          IF rii1 =0 AND rii3  NE 0 THEN
     FRAS  =1 - (((XBB3-XBB4)/XBB3)
          /((RIi3-1)/2*RIi3)) ;
          IF rii1  NE 0 AND rii3  EQ 0 THEN
     FRAS  =1 -(((XBB1-XBB2)/XBB1 )
          /((RIi1-1)/2*RIi1)) ;
          IF rii1  NE 0 AND rii3  NE 0  THEN
     FRAS  =1-
               ((.5*(xbb1-xbb2)/xbb1 +.5*(xbb3-xbb4)/xbb3 )
             / (.5*(rii1-1)/(2*rii1)+ .5*(rii3-
1)/(2*rii3)));

      FILE SASOUT ;
      PUT (XBB1-XBB4) (5.2 +1) +1  RIi1 5.3  +1 RIi3 5.3 +1
      FRAS 6.4  ; end ;
          /*  ;
```

2) PROGRAMME DE CALCUL DE L'INDICE DE FRASER ET BRADFORD, AVEC PRISE EN COMPTE DE L'UNIVERS CONCURRENTIEL (INDICES FRASA ET FRASB).

```
*************************************************************
********** ********; ******* INDICE  DE FRASER ET BRADFORD
VERSION FRASA FRASB*******************;
*************************************************************;
DATA BAS1  ;
     INFILE SASIN1 ;
     INPUT CODEM 1-2  DATUA 4-9       PANEL 12-17
     +1         ( QR1-QR14 ) (5.3 ) +1    ;
     INPUT  (VAR1- VAR14) (1.) +1 ( QA1- QA14  ) (2.)    ;
******* SASIN1 EST LA BASE TRIEE SELON PANEL PUIS DATUA ;

DATA RETARD ; SET BAS1; BY PANEL DATUA;
           LPANEL=LAG(PANEL); RETAIN  LPANEL 0;
           RETAIN L1 -L14 LQR1 -LQR14  0;
           L1 =LAG(VAR1 ); L2 =LAG(VAR2 );
           L3 =LAG(VAR3 );L4 =LAG(VAR4 ); L5 =LAG(VAR5 );
           L6 =LAG(VAR6 );L7 =LAG(VAR7 ); L8 =LAG(VAR8 );
           L9 =LAG(VAR9 );L10=LAG(VAR10); L11=LAG(VAR11);
           L12=LAG(VAR12);L13=LAG(VAR13); L14=LAG(VAR14);
           LQR1 =LAG(QR1 )  ;LQR2 =LAG(QR2 )  ;
           LQR3 =LAG(QR3 )  ;LQR4 =LAG(QR4 )  ;LQR5 =LAG(QR5
           )  ; LQR6 =LAG(QR6 )  ;LQR7 =LAG(QR7 )  ;LQR8
           =LAG(QR8 )  ; LQR9 =LAG(QR9 )  ;LQR10=LAG(QR10)
           ;LQR11=LAG(QR11)  ; LQR12=LAG(QR12)
           ;LQR13=LAG(QR13)  ;LQR14=LAG(QR14)  ;
      RETAIN LI1-LI14     LQ1-LQ14 0;
      ARRAY LI(14) ;
      ARRAY LQ  (14 );
      ARRAY L(14 )  L1 -L14  ;
      ARRAY LQTR(14 ) LQR1 -LQR14 ;
           IF PANEL NE LPANEL THEN  DO; DO I=1 TO 14;
                  LQ (I)=0;      LI(I)=0 ; END;
                                 END;
           IF PANEL=   LPANEL  THEN DO;
           DO I=1 TO 14; LI(I)=L(I);  LQ (I)=LQTR(I) ; END;
           END;
           KEEP LQ1-LQ14 LI1-LI14 QR1 -QR14 PANEL DATUA VAR1
           -VAR14 LPANEL;

DATA RETARD2; SET RETARD;
      RETAIN  LDATUA   0;
      LDATUA=LAG(DATUA);
      ARRAY QR(14) QR1 -QR14   ;
      ARRAY ACTUEL(14) VAR1 -VAR14 ;
      ARRAY LQ   (14 ) LQ1-LQ14 ;
      ARRAY RETARD(14) LI1-LI14  ;
      ARRAY PER (196 ) ;
      ARRAY NPER (196 ) ;
           DO I=1 TO 14  ;DO J=1 TO 14 ;  K=(I-1)*14 +J  ; IF
           ACTUEL(J)=RETARD(I) AND ACTUEL(J)=1  THEN DO;
```

```
            PER(K) =(DATUA -LDATUA)     ;
            NPER(K)=PER(K)/LQ(I)    ;
            END;
            ELSE DO;
            PER(K)=0    ;NPER(K)=0 ; END;
            END;  END;
            KEEP NPER1-NPER196 PANEL DATUA PER1-PER196 LPANEL
            ;

DATA CUMUL; SET RETARD2;BY PANEL;
       ARRAY SNPER(196);RETAIN SNPER1-SNPER196 0;
       ARRAY  NPER(196) NPER1-NPER196 ;
       ARRAY EN(196  ) ;
       RETAIN  EN1-EN196   SN1-SN196  0;
       ARRAY  SN(196) ;
       ARRAY SCPER(196); RETAIN SCPER1-SCPER196  0;
            IF PANEL NE LPANEL THEN  DO;
            DO I=1 TO 196 ;
            SNPER(I)=NPER(I);
            SCPER(I)=   NPER(I)**2;
            IF NPER(I)=0 THEN DO; EN(I)=0 ; SN(I)=EN(I) ;
            END; ELSE  DO ; EN(I)=1; SN(I)=EN(I); END;   END;
            END; ELSE DO;  DO I=1 TO 196  ;
            SNPER(I)=SUM(SNPER(I),NPER(I));
            SCPER(I)=SUM(SCPER(I), NPER(I)**2);
            IF NPER(I)=0 THEN SN(I)=SN(I);
            ELSE SN(I)=SN(I)+1;  END;   END;
             IF LAST.PANEL   THEN OUTPUT CUMU   ;
         KEEP SNPER1-SNPER196   SCPER1-SCPER196  SN1-SN196
PANEL;

DATA  CUMU2; SET CUMU;
       ARRAY SNPER(196 ) SNPER1-SNPER196   ;
       ARRAY MOYE(196 ); RETAIN MOYE1-MOYE196    ;
       ARRAY SCPER(196)  SCPER1-SCPER196  ;
       ARRAY VARI(196); RETAIN VARI1-VARI196  0;
       ARRAY SN(196 ) SN1-SN196  ;
            LPANEL =LAG(PANEL);
            DO I=1 TO 196  ;  IF SN(I)=0 OR SN(I)=1 THEN DO;
            MOYE(I)=0;  VARI(I)=0 ; END;
            ELSE DO;
            MOYE (I)= SNPER(I)/SN(I)    ;
            VARI(I) = SCPER(I)/(SN(I)-1) - (MOYE(I)**2);
               END;       END;

DATA CUMU3 ; SET CUMU2  END=LAST  ;
        ARRAY MOYE(196)     MOYE1-MOYE196 ;
        ARRAY VARI(196)     VARI1-VARI196;
        ARRAY SMOY (196) ; RETAIN SMOY1-SMOY196 0;
        ARRAY SVAR(196); RETAIN SVAR1-SVAR196  0;
        ARRAY MOYP(196);RETAIN MOYP1-MOYP196  0;
        ARRAY  SM(196 ) ; RETAIN SM1-SM196    0;
            DO I=1  TO  196 ;
             iF MOYE(I)=0 THEN DO  ; MOYP(I)=0   ;
```

```
            SMOY(I) = SUM(SMOY(I), MOYE(I));
            SVAR(I) = SUM(SVAR(I) , MOYP(I))  ;
            SM(I)=SM(I);    END;
              ELSE DO  ;
            SMOY(I) = SUM(SMOY(I), MOYE(I));
            MOYP(I) = VARI(I)/( MOYE(I)**2)  ;
            SVAR(I) = SUM(SVAR(I) , MOYP(I) ) ; SM(I)=SM(I)+1
            ;
            END;  END;
             IF LAST THEN OUTPUT CUMU3;

DATA CUMU4;  SET CUMU3 ;
              ARRAY SMOY (196)     SMOY1-SMOY196  ;
              ARRAY SVAR(196) SVAR1-SVAR196   ;
         ARRAY   SM(196 )  SM1-SM196   ;
         ARRAY   RII(196 ) ; RETAIN RII1-RII196  0;
         ARRAY MOYT(196); RETAIN MOYT1- MOYT196 0;
               DO I=1 TO 14 ; DO J=1 TO 14 ; K=(I-1)*14 +J ;
               IF  SM(K)=0 THEN DO; MOYT(K)=0;
               END ;
               ELSE DO;     MOYT(K)= SMOY(K)/SM(K);
               IF I=J THEN DO;
               RII(L)=(SVAR(K)/SM(K) )**(-1);   END;    END;
               END;   END;
               KEEP MOYT1-MOYT196  RII1-RII196  ;
      proc print ;

      DATA FRASA  ; SET CUMU4  ;
   ARRAY AURIE  (196); RETAIN AURIE1-AURIE196  0;
   ARRAY MOYT(196)   MOYT1-MOYT196 ;
 **ARRAY RII (196)  RII1-  RII196  ;
         I=1 ;  J=14 ; L=1 ; M=1 ;
      OK: DO K=I TO J ; IF MOYt(L)=0 THEN DO ;
           **FRASB(K)=0; AURIE(K)=0 ; END ;
          ELSE DO ;
      AURIE(K)= 1- ((MOYt(L)-MOYt(K))/MOYt(L));
   ** FRASB(K)= 1-(((MOYT(L)-MOYT(K))/MOYT(L))/((RII(M)-
1)/2*RII(M))) ;
            END ;
            IF MOYt(L) LT MOYt(K) THEN DO ;
            AURIE (K)=1 ; END;
            END ;
            I=I+14; J=J+14; L=L+15 ; M=M+1 ;
            IF J LE 196 THEN GOTO OK ;ELSE GOTO FIN ;
   FIN:  FILE SASOUT2;
            DO K=1 TO 196;
            PUT                    AURIE(K) 1-10  .7 ; END ;
            proc print  ;
    /*
```

3) PROGRAMME DE CALCUL DE L'INDICE DE LEHMAN (LEM)

```
**************************************************************
******* CALCUL DE L'INDICE DE LEHMAN**************************
**************************************************************
  ************ ce fichier est trie par panel datua ;
DATA BIDON1 ;
     INFILE SASIN  ;
     INPUT CODEM 1-2  DATUA 4-9      PANEL 12-17
           +1         ( QR1 -QR14  ) (3.1 )    ;
     INPUT( VAR1 -VAR14) (1.) +1 (QA1- QA14  ) (2.);
     FILE SASOUT ;
     PUT CODEM 1-2  DATUA 4-9      PANEL 12-17
       +1         ( QR1 -QR14  ) (3.1 )  +1 (QR1-QR14) (3.1) ;
     PUT( VAR1 -VAR14) (1.) +1 (VAR1 -VAR14) (1.) + 1 (QA1-
     QA14) (2.) (QA1 -QA14) (2.)  ;

 DATA bidon2; set bidon1;by panel;
     array variable (28 ) var1-var28    ;
     LPANEL=LAG(PANEL); LDATUA=LAG(DATUA);
     array li(28 ); retain li1-li28 0;
         retain l1-l28 ;
         L1=lag(var1);l2=lag(var2);l3=lag(var3);l4=lag(var4);
         l5=lag(var5);l6=lag(var6);l7=lag(var7);l8=lag(var8);
         L9
         =LAG(VAR9);L10=LAG(VAR10);L11=LAG(VAR11);L12=LAG(VAR
         12); L13=LAG(VAR13);L14=LAG(VAR14);
         L15=LAG(VAR15);L16=LAG(VAR16);L17=LAG(VAR17);L18=LAG
         (VAR18);
         L19=LAG(VAR19);L20=LAG(VAR20);L21=LAG(VAR21);L22=LAG
         (VAR22);
         L23=LAG(VAR23);L24=LAG(VAR24);L25=LAG(VAR25);L26=LAG
         (VAR26); L27=LAG(VAR27);L28=LAG(VAR28);
      array l(28 ) l1-l28 ;
         if panel ne lpanel then  do; do i=1 to 28 ;
         li(i)=0 ; end;  end;
         if panel=   lpanel  then do;
         do i=1 to 28 ; li(i)=l(i); end;
          END;

DATA bidon1; SET bidon2 ; BY PANEL;
         retain somme1-somme784 no1-no28 0;
      array somme(784 );
      array actuel(28 ) var1-var28    ;
      array retard(28 ) li1-li28 ;
      array no(28 ) ;
         IF PANEL NE LPANEL THEN DO;
         DO I=1 TO 28 ; DO J=1 TO 28 ; K=(I-1)*28 + J;
         SOMME(K)=0; END; END;
         DO J=1 TO 28 ; NO(J)=ACTUEL(J);
                END; END;
         if panel eq lpanel then do;
         DO I=1 TO 28 ; DO J=1 TO 28 ; K=(I-1)*28 + J;
         if retard(i)=actuel(j) and actuel(j)=1 then
```

```
          somme(k)=somme(k) +1; END;END;
          DO I=1 TO 28 ;
         no(i)=no(i)+actuel(i); end;
            end;
         keep somme1-somme784 no1-no28  panel ;
         if last.PANEL THEN OUTPUT bidon1 ;

 DATA BIDON2; SET BIDON1   ;
       ARRAY SOMME (784 ) SOMME1-SOMME784 ;
       RETAIN TOT1-TOT784 MO1-MO28 0;
       ARRAY TOT (784 );
       ARRAY NO(28 ) NO1-NO28 ; ARRAY MO(28 );
                  DO I=1 TO 784 ;
                  IF SOMME(I) GE 1 THEN SOMME(I)=1;
                  TOT(I)= SUM(TOT(I),SOMME(I)); END;
                  DO I=1 TO 28 ;
                  IF  NO(i) ge 1 then NO(i)=1;
                  MO(I)=SUM(MO(I),NO(I)) ; END ;
                  KEEP TOT1-TOT784 MO1-MO28       ;
DATA bidon1; SET bidon2  END=LAST  ;
      IF LAST THEN OUTPUT bidon1 ;
DATA INDICE  ; SET BIDON1 ;
           ARRAY SNO(28) MO1-MO28 ;
           ARRAY SMO(28) MO1-MO28 ;
           RETAIN   TOTAL1  -TOTAL784 0 ;
           ARRAY TOTAL(784) ;
           ARRAY TOT (784) TOT1 -TOT784 ;
               DO I=1 TO 28 ; DO J=1 TO 28 ;
               K=(I-1)*28 + J  ;
               IF  SNO(J) +SMO(I) =0 THEN TOTAL(K) =0   ; ELSE
               TOTAL(K)=TOT(K)/(SNO(J)+SMO(I))  ;
                 END;   END;
  /*
```

4) PROGRAMME DE CALCUL DE L'INDICE DE MERUNKA, SELON DEUX VERSIONS (MERA et MERR)

```
*************************************************************
********** ********; ******* CALCUL DES INDICES MERA ET MERR
*************************************************************
********** ********;
DATA BAS1  ;
     INFILE SASIN1 ;
     INPUT CODEM 1-2  DATUA 4-9      PANEL 12-17
     +1        (     QR1-QR14  ) (5.3 )  +1              ;
     INPUT  (   VAR1- VAR14) (1.) +1 ( QA1- QA14  ) (2.)   ;
******* SASIN1 EST LA BASE TRIEE SELON PANEL PUIS DATUA ;
DATA TOTAL1; SET BAS1; BY PANEL;
        array qa( 14) qa1-qa14 ;
        array qr(14 ) qr1-qr14 ;
        array sqa(14 ); retain sqa1-sqa14 0 ;
        array sqr(14 ); retain sqr1-sqr14  0;
          do i=1 to 14 ;
          if first.PANEL THEN DO;
          sqa(i)=qa(i);  sqr(i)=qr(i);  end;
          else do; sqa(i)=sum(sqa(i),qa(i));
          sqr(i)=sum(sqr(i), qr(i));  end;
           end;
          if last.panel then output total1;
          keep sqr1-sqr14 sqa1-sqa14         panel datua;

DATA INDICE; set total1;
       array sqa (14  ) sqa1-sqa14 ;
       array sqr (14 )  sqr1-sqr14 ;
       array quota( 196 ); retain quota1-quota196 0;
       array quotr(196 ); retain quotr1-quotr196 0;
       array soma (196 ); retain soma1-soma196 0 ;
       array somr(196 ) ; retain somr1-somr196  0;
         do i=1 to 14; do j=1 to 14;
         k=(i-1)*14  +J   ;
         if sqa(i)*sqa(j)=0 then do ;quota(k)=0;
         soma(k)=0  ;  end;
         else do; if sqa(i) lt sqa(j) then do;
         quota(k)=sqa(i)/sqa(j);
         soma(k)=sum(of sqa1 -sqa14 )  ;
              end ;
         else do ; quota(k) = sqa(j)/sqa(i)  ; soma(k)=sum(of
         sqa1 -sqa14 )  ; end ;
              end;
          end; end;
         do i=1 to 14; do j=1 to 14;
         k=(i-1)*14  +J   ;
         if sqr(i)*sqr(j)=0 then do ;quotr(k)=0;
          somr(k)=0  ;  end;
         else do;     if sqr(i) lt sqr(j) then do;
          quotr(k)=sqr(i)/sqr(j);
          somr(k)=sum(of sqr1 -sqr14 )  ;
           end ;
```

```
          else do ; quotr(k) = sqr(j)/sqr(i)   ; somr(k)=sum(of
          sqr1 -sqr14 )   ; end ;
            end;
           end; end;
          keep sqr1-sqr14 sqa1-sqa14 quota1-quota196 quotr1-
          quotr196 soma1 -soma196 somr1 -somr196    ;

DATA  AGREGA; set indice  end=last  ;
        array soma(196 ) soma1 -soma196 ;
        array somr(196 ) somr1 -somr196 ;
        array quota(196 )  quota1-quota196 ;
        array quotr(196 ) quotr1-quotr196 ;
        array pondr(196 ); retain pondr1 -pondr196   0;
        array ponda(196 ); retain ponda1 -ponda196   0;
        array ssoma (196 ); retain ssoma1 -ssoma196 0;
        array ssomr (196 ); retain ssomr1 -ssomr196 0;
                do k=1 to 196 ;
               ssoma(k)=sum(ssoma(k), soma(k));
               ssomr(k)=sum(ssomr(k), somr(k));
               pondr(k)=sum(pondr(k ),quotr(k)*somr(k));
               ponda(k)=sum(ponda(k ),quota(k)*soma(k) );
                  end;
         keep  pondr1-pondr196 ponda1-ponda196 ssomr1 -
ssomr196
                                                ssoma1 -
ssoma196  ;
               if last then output agrega ;
       PROC PRINT; title 'tableau cumule';

DATA MERU ; set agrega    ;
      array ponda(196 ) ponda1 -ponda196 ;
      array pondr (196 ) pondr1 -pondr196 ;
      array merua (196 ); retain merua1 -merua196 0;
      array merur (196 ); retain merur1-merur196 0;
      array ssomr (196) ssomr1 -ssomr196  ;
      array ssoma (196) ssoma1 -ssoma196  ;
         do k=1  to 196  ;
         if ponda(k)*ssoma(k)=0 then  merua(k)=0 ;
         else  merua(k)= ponda(k )/SSOMA(k)  ;
            end ;
         do k=1 to 196  ;
         if pondr(k)*ssomr(k)=0 then  merur(k)=0 ;
         else  merur(k)= pondr(k )/SSOMR(k)  ;
          end;
         keep merua1-merua196   merur1-merur196  ;
         proc print ; title  'indices' ;

       file sasout1   ;
              do i=1 to 196  ;
              put merur(i) ;
              end ;
       file sasout2  ;
              do i=1 to 196  ;
              put merua(i)   ;
```

```
    end  ;
PROC PRINT; title 'indices de MERUNKA';
```

5) PROGRAMME DE CALCUL DE L'INDICE DE RAO ET SABALAVA (RAO et SABA)

```
**************************************************************
********** * ******* INDICE DE RAO ET SABALAVA
**************************************************************
********** ********;

DATA        BAS1   ;
     INFILE SASIN1 ;
******* SASIN1 EST LA BASE TRIEE SELON PANEL PUIS DATUA ;

     INPUT  DATUA 4-9       PANEL 12-17   ;
     INPUT  (VAR1 - VAR14) (1.)    ;
        IF DATUA EQ 10091 OR DATUA EQ 10133 ;
        LPANEL=LAG(PANEL); LDATUA=LAG(DATUA);
      ARRAY LI(14 ); RETAIN LI1-LI14 0;
      RETAIN L1-L9      LD 0;
         L1=LAG(VAR1);L2=LAG(VAR2);L3=LAG(VAR3);L4=LAG(VAR4);
         L5=LAG(VAR5);L6=LAG(VAR6);L7=LAG(VAR7);L8=LAG(VAR8);
         L9=LAG(VAR9);L10=LAG(VAR10);L11=LAG(VAR11);L12=LAG(V
         AR12); L13=LAG(VAR13);L14=LAG(VAR14);

      ARRAY L( 14) L1-L14 ;
         IF PANEL NE LPANEL THEN  DO; DO I=1 TO 14 ;
          LI(I)=0 ; END;
          LD=0;  END;
         IF PANEL=   LPANEL  THEN DO;
          LD=LDATUA     ;
         DO I=1 TO 14; LI(I)=L(I);  END;
         END;
         DROP LDATUA LPANEL L1-L14    ;

DATA PARDEM;  INFILE SASIN2 ;
************ CE FICHIER EST  TRIE SELON DATUA PUIS PANEL ;
     INPUT             DATUA 4-9       PANEL 12-17   ;
     INPUT  (  VAR1 - VAR14) (1.)    ;
DATA PARDEM2; SET PARDEM;  BY DATUA;
           LPANEL=LAG(PANEL);   LDATUA=LAG(DATUA); LD  =
           DATUA  ; ARRAY VAM(14) VAR1-VAR14    ;
           ARRAY SM (14); RETAIN SM1-SM14 0;
           ARRAY RM (14) ;RETAIN RM1-RM14 0 ;
 *********** RM(I) EST LA PDM JOUR 'DEPART' DE LA MARQUE ;
 *********** SN(I) EST LA PDM JOUR 'ARRIVEE' DE LA MARQUE ;
           IF DATUA NE LDATUA THEN DO; DO I=1 TO 14;
           SM(I)=VAM(I); END; END;
           ELSE DO; DO I=1 TO 14; SM(I)=SUM(SM(I), VAM(I));
            END;END;
           DO I=1 TO 14 ;
           RM(I)= SM(I) ;  END ;
        IF LAST.DATUA   THEN OUTPUT PARDEM2;
        KEEP LD DATUA SM1-SM14 RM1-RM14 ;
       PROC PRINT; TITLE 'TABLEAU AVEC LES PDM';
```

```
DATA FUSION1; SET BAS1  ; IF LD NE 0 ; PROC SORT; BY LD
    ; DATA FUSION2 ; SET PARDEM2 ;
       KEEP LD RM1-RM14     ;
DATA FUSION3;  MERGE    FUSION1   FUSION2 ;   BY LD      ;
       PROC SORT ;  BY DATUA       ;

DATA FUSION4 ; SET PARDEM2 ;
      KEEP DATUA SM1-SM14 ;
DATA       COMPLE ; MERGE FUSION3 FUSION4 ; BY DATUA;
DATA CUMUL ; SET       COMPLE ; BY DATUA ;
      ARRAY SOMME (196);
       RETAIN SOMME1-SOMME196 0;
       ARRAY CUMUL (196);
       RETAIN CUMUL1-CUMUL196 0;
       ARRAY ACTUEL(14) VAR1-VAR14    ;
       ARRAY RETARD(14) LI1-LI14   ;
       ARRAY  RM (14 );RETAIN  RM1- RM14 0;
       ARRAY SRM(196);RETAIN SRM1-SRM196  0;
       RETAIN LDATUA     0;
       LDATUA=LAG(DATUA)  ;
       IF DATUA NE  LDATUA THEN DO ;
       DO I=1 TO 14 ;DO J=1 TO 14;  K=(I-1)*14+J  ;
       IF ACTUEL(J)=RETARD(I) AND ACTUEL(J)  =1     THEN DO;
          SOMME(K)=1 ;
          SRM(K)= RM(I)   ;
          END ;
       ELSE DO;  SOMME(K)=0; SRM(K)=0 ; END;
       CUMUL(K)=SOMME(K); END ; END ;
       END ;
       ELSE   DO ;
        DO I=1 TO 14 ;DO J=1 TO 14;  K=(I-1)*14+J  ;
        IF ACTUEL(J)=RETARD(I) AND ACTUEL(J)=1  THEN DO ;
        SOMME(K)=1;
         SRM(K) = SUM(SRM(K),RM(I)) ;
            END ;
        ELSE  DO  ;SOMME(K)=0; SRM(K)=SRM(K); END ;
           CUMUL(K)= SUM(CUMUL(K),SOMME(K));
              END; END ;
                END ;
           IF LAST.DATUA THEN OUTPUT       CUMUL ;
           KEEP  CUMUL1-CUMUL196 DATUA SRM1-SRM196  SM1-SM14
           ;
DATA  CUMUL2 ; SET CUMUL ;
       ARRAY CUMUL (196) CUMUL1 -CUMUL196  ;
       ARRAY SRM(196)         SRM1-SRM196  ;
       ARRAY NNJ(196); RETAIN NNJ1-NNJ196 0;
       ARRAY SM(14) SM1-SM14;
          DO I=1 TO 14 ;DO J=1 TO 14;  K=(I-1)*14+J  ;
          IF CUMUL(K)=0 THEN DO ;
          SRM(K)=0 ;
         NNJ(K)=0   ;
         END ;
         ELSE DO ;  SRM(K)=SRM(K)/CUMUL(K) ;
```

```
            NNJ(K)=SM(J)   ;
             END;
           END ; END ;
       KEEP CUMUL1-CUMUL196 SRM1-SRM196 NNJ1-NNJ196 DATUA ;

DATA  SOMM ;  SET      CUMUL2 END=LAST ;
      ARRAY SSOM(196);
      ARRAY SRM(196) SRM1-SRM196 ;
      ARRAY NNJ(196) NNJ1-NNJ196 ;
      ARRAY SSNMI(196) ;
      ARRAY SSNNJ(196) ;
      RETAIN SSOM1-SSOM196 SSNMI1-SSNMI196 SSNNJ1-SSNNJ196 0;
      ARRAY CUMUL(196) CUMUL1-CUMUL196;
      DO I=1 TO 196;  SSOM(I)= SUM(SSOM(I), CUMUL(I)) ;
         SSNMI(I)= SUM(SSNMI(I),SRM(I))   ;
         SSNNJ(I)= SUM(SSNNJ(I),NNJ(I))   ;
         END  ;
      IF LAST THEN OUTPUT SOMM;  KEEP SSOM1-SSOM196
      SSNMI1-SSNMI196 SSNNJ1-SSNNJ196  ;


  DATA SABALAVA ; SET SOMM ;
       ARRAY SSNMI(196) SSNMI1-SSNMI196 ;
       ARRAY SSNNJ(196) SSNNJ1-SSNNJ196 ;
       ARRAY SSOM(196) SSOM1-SSOM196 ;
       ARRAY SABA(196) ; RETAIN SABA1-SABA196  0;
             DO I=1 TO 196 ;
             IF SSNMI(I)*SSNNJ(I)=0  THEN SABA(I) =0 ;
             IF SSNMI(I)*SSNNJ(I) NE 0 THEN
             SABA(I)=SSOM(I)/(SSNMI(I)*SSNNJ(I)) ;
             END  ;
             FILE SASOUT ;
             DO I=1 TO 196 ;
          PUT SABA(I) 1-15  .7    SSOM(I) 17-22   ;
             END  ;
       PROC PRINT ; TITLE 'TABLEAU FINAL' ;
  /*
```

6) PROGRAMME DE CALCUL DES INDICES "ATTAQUE" ET "DEFENSE" (TOTT, TRTO, ELRO et TREL).

```
**********************************************************
********** * ******* ******************; *******
CE PROGRAMME PERMET DE CALCULER DES INDICES DE
CONCURRENCE EN UTILISANT UNE DISTANCE DU CHI DEUXSUR LES
PROFILS DE TRANSFERTS ENTRE MARQUES **********************
**********************************************************;
DATA LECTURE ;
     INFILE SASIN  ;
      ARRAY SWIT ( 14); M=1 ;
      DO I= 1 TO 14  ;
       INPUT SWIT(I)  1-15 ;    END ;
DATA LECT2 ; SET LECTURE  ; RETAIN SN 0 ;
       SN = SUM(SN,N)  ;
       PDM=SUM(OF  switching ) ; PDM EST LA SOMME DES
       TRANSFERTS EXEPTE DE LA MARQUE VERS ELLE MEME
DATA COLONNE ; SET LECTURE   END=LAST  ;
     ARRAY SWIT(14) SWIT1-SWIT14  ;
     ARRAY SCOL(2744) ; RETAIN SCOL1-SCOL196  0;
      DO I=1 TO 14 ;
      DO J=1 TO 14 ;
      DO H=1   TO 14 ;
       K=(J-1)*14+H ;
       L=196*(J-1) +K ;
       IF N NE I AND N NE J THEN SCOL(L)=SUM(SCOL(L),SWIT(I))
       ;
       ELSE SCOL(L)=SCOL(L) ;
       END ; END ;
       END ;
       IF LAST THEN OUTPUT COLONNE ;
       KEEP SCOL1-SCOL2744  ;
   DATA DISTANCE; MERGE  COLONNE LECTURE ;
      ARRAY SCOL(2744) SCOL1-SCOL2144 ;
      ARRAY SWIT(14) SWIT1-SWIT14 ;
      ARRAY DIST(196); RETAIN DIST1-DIST196 0 ;
      ARRAY SDIST(196); RETAIN SDIST1-SDIST196 0 ;
           DO I=1 TO 14 ;
           DO J=1 TO 14 ;
            K= (I-1)*196+ (I-1)*3 +J ;
            L= (J-1)*196 + (J-1)*3  +I ;
            M=(I-1)*14 +J ;
       if N NE I AND SN NE J THEN
        DIST(M)=( SWIT(I)/SCOL(K)-SWIT(J)/SCOL(L))**2/PDM ;
        ELSE DIST(M)=0 ;
        SDIST(M)=SUM(SDIST(M),DIST(M)) ;
            END ; END ;
```

(*) Ces programmes ne sont pas les versions définitives qui ont été utilisées lors du calcul effectif des indices. Celles-ci ne sont malheureusement plus disponibles, elles figuraient dans une bibliothèque de programmes qui a été perdue au centre de calcul de Montpellier (CNUCS) lors de la destruction accidentelle de la "cassette" sur laquelle était sauvegardée cette bibliothèque, au mois de juin dernier (voir lettre ci-jointe).

Ministère de l'Education Nationale
Académie de Montpellier - Chancellerie des Universités

CENTRE NATIONAL UNIVERSITAIRE SUD DE CALCUL

950, rue de St-Priest B.P. 7229 F 34083 MONTPELLIER CEDEX Tél. 67.14.14.14 Télécopie 67.52.37.63 Télex 480 438 F.

Montpellier, le 11 Juillet 1990

Le 20 juin 1990, à la suite d'un incident sur le système de gestion de fichiers du nouveau système d'exploitation MVS/ESA, les fichiers de Monsieur AURIER ont été détruits ainsi que leurs copies de sauvegarde.

Il s'agit des bibliothèques:

- EEC 2501. AURIER. NILSEN2
- EEC 2501. AURIER. CNUSC
- EEC 2501. AURIER. COREL

C'est le premier accident de ce type depuis la création du CNUSC en 1980, les procédures de sauvegarde permettant d'assurer, normalement, l'intégrité des fichiers du CNUSC et de ses utilisateurs.

J.L DELHAYE

ANNEXE 4 DETAIL DES INDICES CALCULES SUR LES DIFFERENTS MARCHES ET LES DIFFERENTS UNIVERS

Voici, sous forme de tableau, le détail des indices calculés, pour chaque marché, et chaque Univers. Chaque croix indique qu'un calcul a été effectué pour l'indice et l'univers correspondants .

1 MARCHE HUILE 10 MARQUES

INDICE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
UNIVERS															
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2			+	+	+	+	+	+	+	+	+	+	+	+	+
3			+	+	+	+	+	+	+	+	+	+	+	+	+
4			+	+	+	+	+	+	+	+	+	+	+	+	+
5			+	+	+	+	+	+	+	+	+	+	+	+	+
6			+	+	+	+	+	+	+	+	+	+	+	+	+
7			+	+	+	+	+	+	+	+	+	+	+	+	+
8			+	+	+	+	+	+	+	+	+	+	+	+	+
9			+	+	+	+	+	+	+	+	+	+	+	+	+
10			+	+	+	+	+	+	+	+	+	+	+	+	+
11			+	+	+	+	+	+	+	+	+	+	+	+	+
12			+	+	+	+	+	+	+	+	+	+	+	+	+
13			+	+	+	+	+	+	+	+	+	+	+	+	+
15			+	+	+	+	+	+	+	+	+	+	+	+	+
16			+	+	+	+	+	+	+	+	+	+	+	+	+
17			+	+	+		+	+	+	+	+	+	+	+	+
18					+		+	+	+	+	+	+	+	+	+

Soit, 234 fichiers de résultats

2 MARCHE BOISSON

INDICE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
UNIVERS															
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
2			+	+	+	+	+	+	+	+	+	+	+	+	+
3			+	+	+	+	+	+	+	+	+	+	+	+	+
4			+	+	+	+	+	+	+	+	+	+	+	+	+
5			+	+	+	+	+	+	+	+	+	+	+	+	+
6			+	+	+	+	+	+	+	+	+	+	+	+	+
7			+	+	+	+	+	+	+	+	+	+	+	+	+
8			+	+	+	+	+	+	+	+	+	+	+	+	+
9			+	+	+	+	+	+	+	+	+	+	+	+	+
10			+	+	+	+	+	+	+	+	+	+	+	+	+
11			+	+	+	+	+	+	+	+	+	+	+	+	+
12			+	+	+	+	+	+	+	+	+	+	+	+	+
13			+	+	+	+	+	+	+	+	+	+	+	+	+
14			+	+	+	+	+	+	+	+	+	+	+	+	+
15			+	+	+	+	+	+	+	+	+	+	+	+	+
16			+	+	+	+	+	+	+	+	+	+	+	+	+
17			+	+	+	+	+	+	+	+	+	+	+	+	+
18			+	+	+	+	+	+	+	+	+	+	+	+	+

Soit, 238 fichiers de résultats

3 MARCHE PARFUM

INDICE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
UNIVERS															
1	+	+	+	+	+		+	+	+	+	+	+	+	+	+
2			+	+	+		+	+	+	+	+	+	+	+	+
3			+	+	+		+	+	+	+	+	+	+	+	+
4			+	+	+		+	+	+	+	+	+	+	+	+
5			+	+	+		+	+	+	+	+	+	+	+	+
6			+	+	+		+	+	+	+	+	+	+	+	+
7			+	+	+		+	+	+	+	+	+	+	+	+
8			+	+	+		+	+	+	+	+	+	+	+	+
9			+	+	+		+	+	+	+	+	+	+	+	+
10			+	+	+		+	+	+	+	+	+	+	+	+
11			+	+	+		+	+	+	+	+	+	+	+	+
12			+	+	+		+	+	+	+	+	+	+	+	+
13			+	+	+		+	+	+	+	+	+	+	+	+
14			+	+	+		+	+	+	+	+	+	+	+	+
15			+	+	+		+	+	+	+	+	+	+	+	+
16			+	+	+		+	+	+	+	+	+	+	+	+
17			+	+	+		+	+	+	+	+	+	+	+	+
18			+	+	+		+	+	+	+	+	+	+	+	+

Soit, 231 fichiers de résultats .

4 MARCHE SHAMPOOING 24 MARQUES

INDICE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
UNIVERS															
1			+	+	+	+	+	+			+			+	+
2															
3															
4															
5															
6															
7															
8			+	+	+	+	+	+			+			+	+
9			+	+	+	+	+	+			+			+	+
10			+	+	+	+	+	+			+			+	+
11			+	+	+	+	+	+			+			+	+
12			+	+	+	+	+	+			+			+	+
13			+	+	+	+	+	+			+			+	+
14			+	+	+	+	+	+			+			+	+
15			+	+	+	+	+	+			+			+	+
16			+	+	+	+	+	+			+			+	+
17			+	+	+	+	+	+			+			+	+
18			+	+	+	+	+	+			+			+	+

Soit, 130 fichiers de résultats .

EXTRAIT D'UN FICHIER DE CORRELATIONS

FRASA	GLOBAL	1444443	FRASA	GLOBAL	1444443	1.00
FRASA	REVG	1144443	FRASA	GLOBAL	1444443	0.00
FRASA	REVG	1144443	FRASA	REVG	1144443	1.00
FRASA	REVM	1244443	FRASA	GLOBAL	1444443	0.11
FRASA	REVM	1244443	FRASA	REVG	1144443	0.02
FRASA	REVM	1244443	FRASA	REVM	1244443	1.00
FRASA	REVP	1344443	FRASA	GLOBAL	1444443	0.44
FRASA	REVP	1344443	FRASA	REVG	1144443	0.07
FRASA	REVP	1344443	FRASA	REVM	1244443	0.15
FRASA	REVP	1344443	FRASA	REVP	1344443	1.00
FRASA	CSPS	1414443	FRASA	GLOBAL	1444443	-.22
FRASA	CSPS	1414443	FRASA	REVG	1144443	-.24
FRASA	CSPS	1414443	FRASA	REVM	1244443	0.13
FRASA	CSPS	1414443	FRASA	REVP	1344443	0.14
FRASA	CSPS	1414443	FRASA	CSPS	1414443	1.00
FRASA	CSP5	1424443	FRASA	GLOBAL	1444443	0.51
FRASA	CSP5	1424443	FRASA	REVG	1144443	-.13

ANNEXE 5 TABLEAU DES MOYENNES

Dans case du tableau suivant figurent la moyenne, l'écart-type et l'effectif de la strate sur laquelle ont été calculées les corrélations.

α MARCHE HUILE

VARIABLE	COL1	COL2	COL3	COL4	COL5
CORREL	.062/.24 19107	.261/.33 1648	.270/.28 173	.072/.23 1913	.110/.29 969
CORRN	-.160/.13 7733	-.168/.13 389	-.121/.12 30	-.151/.11 686	-.152/.12 360
CORRP	.213/.19 11374	.394/.25 1259	.353/.23 143	.197/.17 1227	.265/.25 609
ACORR	.192/.17 19107	.340/.25 1648	.312/.23 173	.181/.15 1913	.223/.22 969

COL1 : Calculs effectués sur la totalité des corrélations
COL2 : Calculs effectués sur les corrélations "intra-indices" (I1=I2) COL3 : Calculs effectués sur les corrélations "intra-indices" et "intra-facteurs" (I1=I2 et F1=F2).
COL4 : Calculs effectués sur les corrélations "intra-facteurs" (F1=F2)
COL5 : Calculs effectués sur les corrélations "intra-univers" (U1=U2)

β MARCHE BOISSON

VARIABLE	COL1	COL2	COL3	COL4	COL5
CORREL	.04/.21 19497	.28/.26 1682	.22/.23 176	.033/.19 1936	.059/.29 985
CORRN	-.150/.12 8277	-.13/.12 202	-.13/.11 25	-.14/.11 802	-.19/.16 417
CORRP	.181/.16 11220	.335/.22 1480	.26/.19 151	.16/.13 1134	.24/.23 568
ACORR	.168/.14 19497	.31/.22 1682	.26/.19 176	.15/.12 1636	.22/.20 985

Γ MARCHE PARFUM

VARIABLE	COL1	COL2	COL3	COL4	COL5
CORREL	.027/33 16087	.29/.33 1526	.25/.28 160	.032/.24 1600	.087/.35 793
CORRN	-.17/.13 8298	-.18/.14 286	-.14/.12 26	-.15/.11 787	-.18/.13 375
CORRP	.24/.24 7791	.40/.26 1240	.33/.24 134	.21/.21 813	.33/.32 418
ACORR	.020/.19 16089	.36/.26 1526	.30/.23 160	.18/.17 1600	.26/.26 793

π MARCHE SHAMPOOING

VARIABLE	COL1	COL2	COL3	COL4	COL5
CORREL	.063/.21 3479	.28/.24 459	.20/.20 70	.038/.17 490	.11/.31 249
CORRN	-.12/.09 1369	-.08/.06 43	-.05/.04 7	-.10/.09 196	-.15/.10 110
CORRP	.18/.19 2210	.322/.22 416	.23/.19 63	.13/.14 294	.26/.30 139
ACORR	.15/.16 3479	.30/.22 459	.21/.19 70	.12/.12 490	.22/.25 249

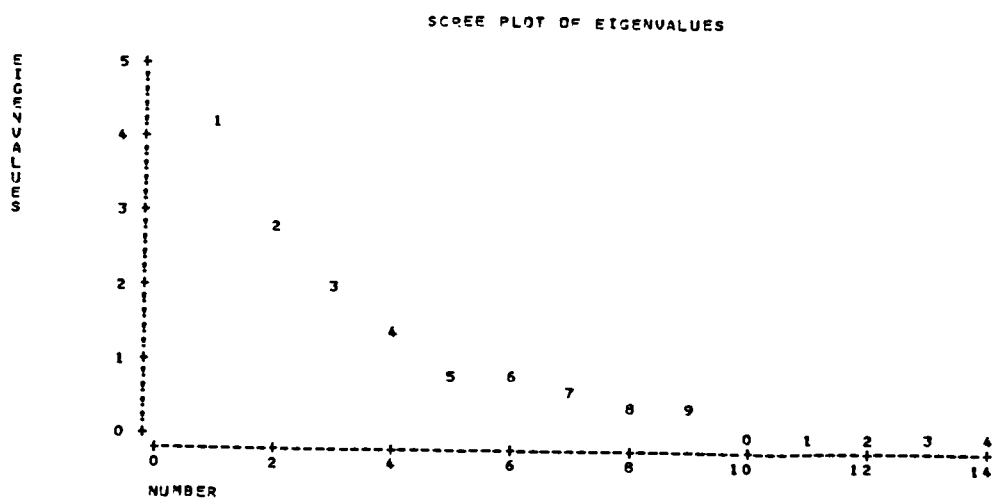
Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

ANNEXE 6 ETUDE DE LA COHERENCE INTERNE DES INDICES ANALYSES FACTORIELLES (COEFFICIENT DE CORRELATION LINEAIRE DE PEARSON)

ANNEXE 6.α MARCHE DE L'HUILE

1) VALEURS PROPRES ET POURCENTAGE D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	30.6	30.6
2	20.3	50.9
3	13.9	64.7
4	10.0	74.8
5	6.4	81.1
6	6.0	87.1
7	4.6	91.8
8	3.6	95.3
9	2.7	98.0
10	.07	98.7
11	.06	99.2
12	.05	99.9
13	.02	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES QUATRE PREMIERS FACTEURS (74.8% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4
BRAA	11	10	-4	50 *
FRASA	20	2	4	88 *
FRASB	27	9	10	92 *
LEM	94 *	29	3	13
MERA	8	3	89 *	3
MERR	2	0	91 *	8
SABA	9	12	58 *	19
TOTT	9	89 *	9	9
TRTO	14	79 *	30	4
RAO	20	18	68 *	10
ELRO	23	82 *	28	8
TREL	24	67 *	22	14
IND	97 *	18	-4	9
ACT	97 *	16	-3	-6

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace 4 facteurs) TOTAL = 10.5

BRAA	FRASA	FRASB	LEM	MERA	MERR	SABA	TOTT	TRTO	RAO
0.28	0.82	.94	.98	.79	.84	.39	.82	.73	.54

ELRO	TREL	IND	ACT
.81	.57	.98	.96

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES QUATRE PREMIERS FACTEURS (74.8% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4
BRAA	-8	12	-8	-47 *
FRASA	29	7	10	90 *
FRASB	38	15	16	96 *
LEM	99 *	48 *	4	30
MERA	-7	3	88 *	12
MERR	-2	1	92 *	16
SABA	13	12	58 *	26
TOTT	29	89 *	6	15
TRTO	31	79 *	28	13
RAO	25	23	68 *	8
ELRO	39	86 *	32	3
TREL	36	70 *	26	9
IND	99 *	38	2	27
ACT	98 *	35	2	24

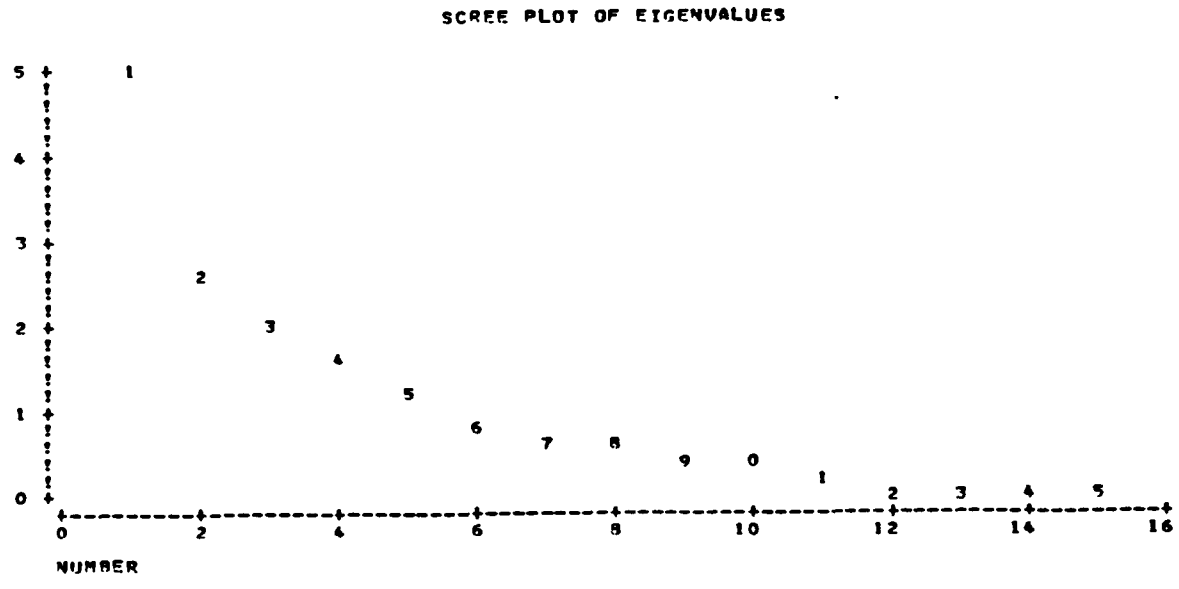
5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4
FACTEUR1	100 *			
FACTEUR2	41	100 *		
FACTEUR3	2	6	100 *	
FACTEUR4	28	9	16	100 *

ANNEXE 6.β MARCHE BOISSON

1) VALEURS PROPRES ET POURCENTAGES D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	33.5	33.5
2	16.8	50.3
3	13.7	64.0
4	10.2	74.0
5	7.7	81.8
6	5.1	86.9
7	3.9	90.8
8	3.5	94.3
9	.22	96.6
10	.21	98.7
11	.09	99.6
12	.04	100
13	.00	100
14	.00	100
15	.00	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES CINQ PREMIERS FACTEURS (81.8% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4	FACTEUR5
BRAA	-20	6	62 *	9	53 *
BRAB	-1	9	78 *	2	14
FRASA	49 *	-4	76 *	28	-14
FRASB	52 *	-6	73 *	18	-5
LEM	94 *	16	9	1	24
MERA	1	11	1	96 *	9
MERR	0	10	12	95 *	2
SABA	40 *	1	7	10	53 *
TOTT	2	86 *	9	10	4
TRTO	10	94 *	-3	9	9
RAO	62 *	52 *	25	6	8
ELRO	32	6	4	7	81 *
TREL	14	92 *	-3	5	-1
IND	94 *	3	4	1	23
ACT	93 *	11	5	2	17

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace à 5 facteurs) TOTAL = 12.27

BRAA	BRAB	FRASA	FRASB	LEM	MERA	MERR	SABA	TOTT
0.71	.64	.91	.84	.97	.94	.93	.45	.76

TRTO	RAO	ELRO	TREL	IND	ACT
.91	.73	.77	.87		

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES CINQ PREMIERS FACTEURS (81.8% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4	FACTEUR5
BRAA	-9	8	58 *	-5	54 *
BRAB	9	11	77 *	15	14
FRASA	57 *	2	84 *	38	-18
FRASB	60 *	0	81 *	28	-9
LEM	96 *	24	25	8	20
MERA	8	15	11	96 *	5
MERR	8	14	21	96 *	2
SABA	44 *	7	16	14	50 *
TOTT	12	87 *	13	16	7
TRTO	20	95 *	2	15	13
RAO	68 *	56 *	35	4	-8
ELRO	38	13	12	10	79 *
TREL	23	93 *	2	12	3
IND	95 *	11	20	7	19
ACT	94 *	19	21	11	14

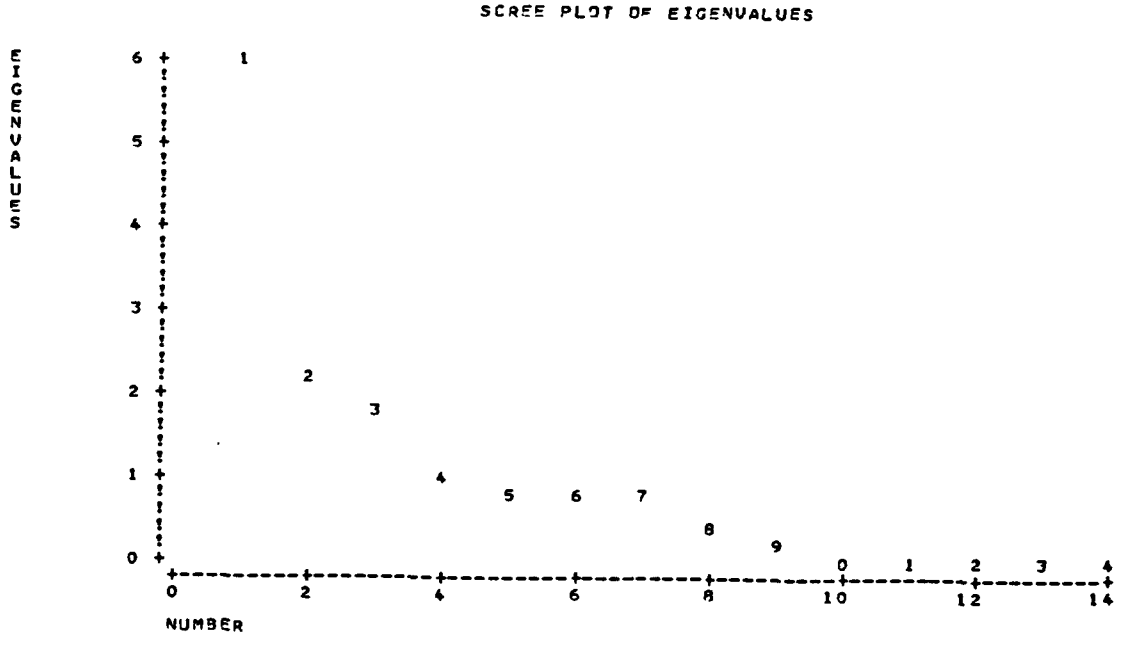
5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4	FACTEUR5
FACTEUR1	100 *				
FACTEUR2	18	100 *			
FACTEUR3	28	6	100 *		
FACTEUR4	15	11	19	100 *	
FACTEUR5	2	9	0	-5	100 *

ANNEXE 6.Γ MARCHE PARFUM

1) VALEURS PROPRES ET POURCENTAGES D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	42.1	42.1
2	16.0	58.2
3	12.4	70.6
4	7.0	77.5
5	6.4	83.8
6	6.0	89.9
7	5.5	95.3
8	2.5	97.9
9	1.6	99.5
10	.29	99.7
11	.15	99.9
12	.07	100
13	.03	100
14	.01	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (70.6% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
BRAA	2	5	90 *
BRAB	3	5	88 *
FRASA	25	65 *	1
FRASB	7	82 *	15
LEM	34	85 *	0
MERR	31	-2	33
SABA	-2	26	31
TOTT	95 *	25	9
TRTO	96 *	23	9
RAO	37	60 *	3
ELRO	95 *	26	0
TREL	95 *	20	-1
IND	10	83 *	9
ACT	12	90 *	13

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace à 3 facteurs) TOTAL = 9.97

BRAA	BRAB	FRASA	FRASB	LEM	MERR	SABA	TOTT
0.81	.78	.48	.71	.84	.20	.17	.97
TRTO	RAO	ELRO	TREL	IND	ACT		
.97	.50	.96	.93	.70	.85		

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (70.6% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
BRAA	5	12	90 *
BRAB	7	12	88 *
FRASA	38	68 *	2
FRASB	26	83 *	18
LEM	52	90 *	2
MERR	30	6	32
SABA	4	28	32
TOTT	98 *	43	8
TRTO	98	42	9
RAO	49	65 *	-2
ELRO	98 *	43	0
TREL	97 *	38	-2
INDL	28	84 *	12
ACT	32	92 *	16

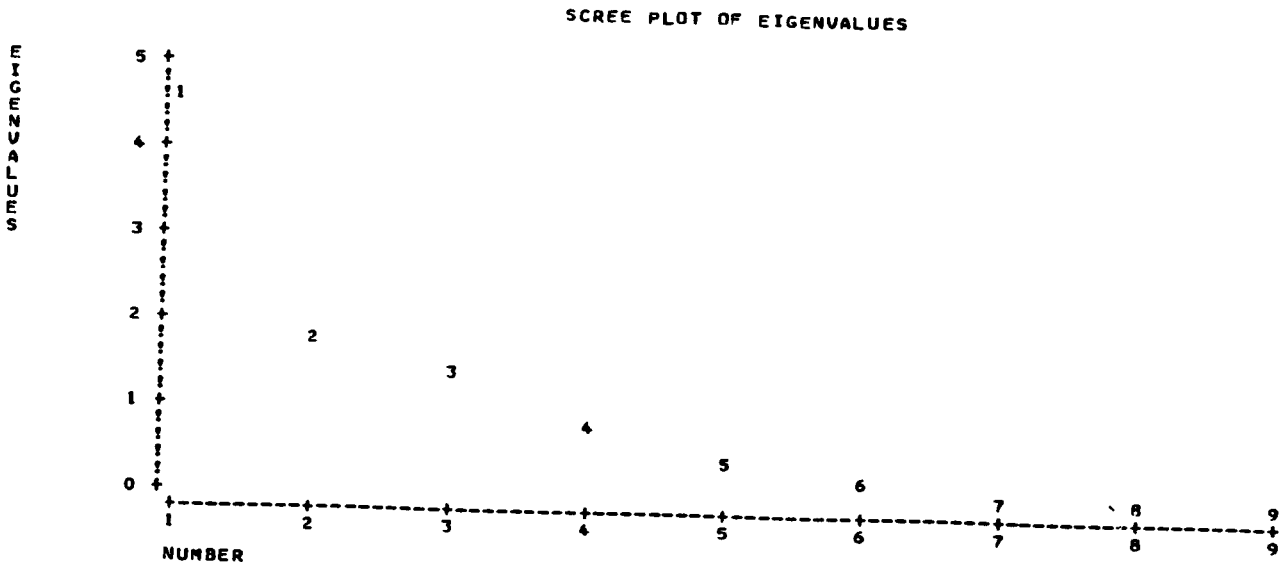
5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3
FACTEUR1	100 *	40	2
FACTEUR2	40	100 *	10
FACTEUR3	2	10	100 *

ANNEXE 6.π MARCHE SHAMP

1) VALEURS PROPRES ET POURCENTAGES D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	50.2	50.2
2	20.3	70.5
3	15.0	85.6
4	8.1	93.7
5	4.8	98.4
6	1.3	99.7
7	0.2	99.9
8	0.1	100
9	0.0	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (85.6% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
FRASA	20	95 *	5
FRASB	23	96 *	3
LEM	97 *	24	5
MERA	5	7	97 *
MERR	7	2	97 *
SABA	35	50	4
RAO	78 *	24	9
IND	95 *	21	3
ACT	94 *	23	3

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace à 3 facteurs) TOTAL = 7.70

FRASA	FRASB	LEM	MERR	SABA	RAO	IND	ACT
.94	.97	.99	.94	.37	.67	.93	.93

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (85.6% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
FRASA	41	97 *	10
FRASB	44	99 *	9
LEM	100 *	45	12
MERA	11	12	97 *
MERR	12	8	97 *
SABA	45	56	8
RAO	82 *	41	14
IND	97 *	41	9
ACT	96 *	44	10

5) CORRELATIONS ENTRE LES FACTEURS

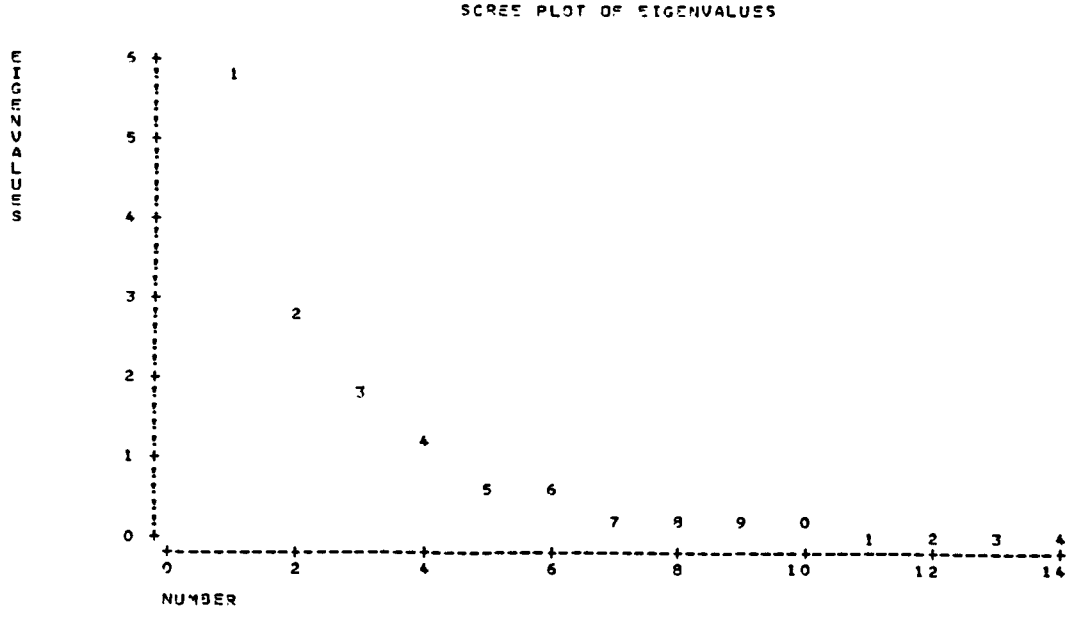
	FACTEUR1	FACTEUR2	FACTEUR3
FACTEUR1	100 *		
FACTEUR2	43	100 *	
FACTEUR3	12	10	100 *

ANNEXE 7 ETUDE DE LA COHERENCE INTERNE DES INDICES ANALYSES FACTORIELLES (COEFFICIENT DE RANG DE SPEARMAN)

ANNEXE 7.α MARCHE DE L'HUILE

1) VALEURS PROPRES ET POURCENTAGE D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	42.0	42.0
2	20.0	62.1
3	13.2	75.4
4	9.1	84.5
5	4.5	89.0
6	4.2	93.2
7	1.9	95.1
8	1.8	97.0
9	1.2	98.1
10	.08	99.0
11	.04	99.4
12	.01	99.9
13	.00	100
14	.00	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES QUATRE PREMIERS FACTEURS (84.5% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4
BRAA	12	19	3	83 *
FRASA	11	3	97 *	2
FRASB	15	10	95 *	12
LEM	82 *	24	22	40
MERA	16	84 *	1	22
MERR	1	90 *	5	16
SABA	10	80 *	12	8
TOTT	86 *	3	9	19
TRTO	88 *	25	11	12
RAO	39	66 *	12	41
ELRO	91 *	3	3	10
TREL	92 *	7	2	3
IND	79 *	15	26	44
ACT	79 *	15	27	43

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace 4 facteurs) TOTAL = 10.14

BRAA	FRASA	FRASB	LEM	MERA	MERR	SABA	TOTT	TRTO	RAO
0.74	0.96	.95	.93	.77	.84	.67	.79	.87	.77

ELRO	TREL	IND	ACT
.84	.86	.91	.90

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES QUATRE PREMIERS FACTEURS (84.5% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4
BRAA	5	-14	-11	76 *
FRASA	22	2	97 *	17
FRASB	27	7	97 *	31
LEM	89 *	35	40	61 *
MERA	13	80 *	8	16
MERR	2	88 *	9	7
SABA	15	81 *	10	16
TOTT	84 *	4	22	2
TRTO	86 *	17	25	7
RAO	48	73 *	22	56 *
ELRO	89 *	11	12	9
TREL	91 *	16	14	17
IND	86 *	26	44	65 *
ACT	86 *	26	45	64 *

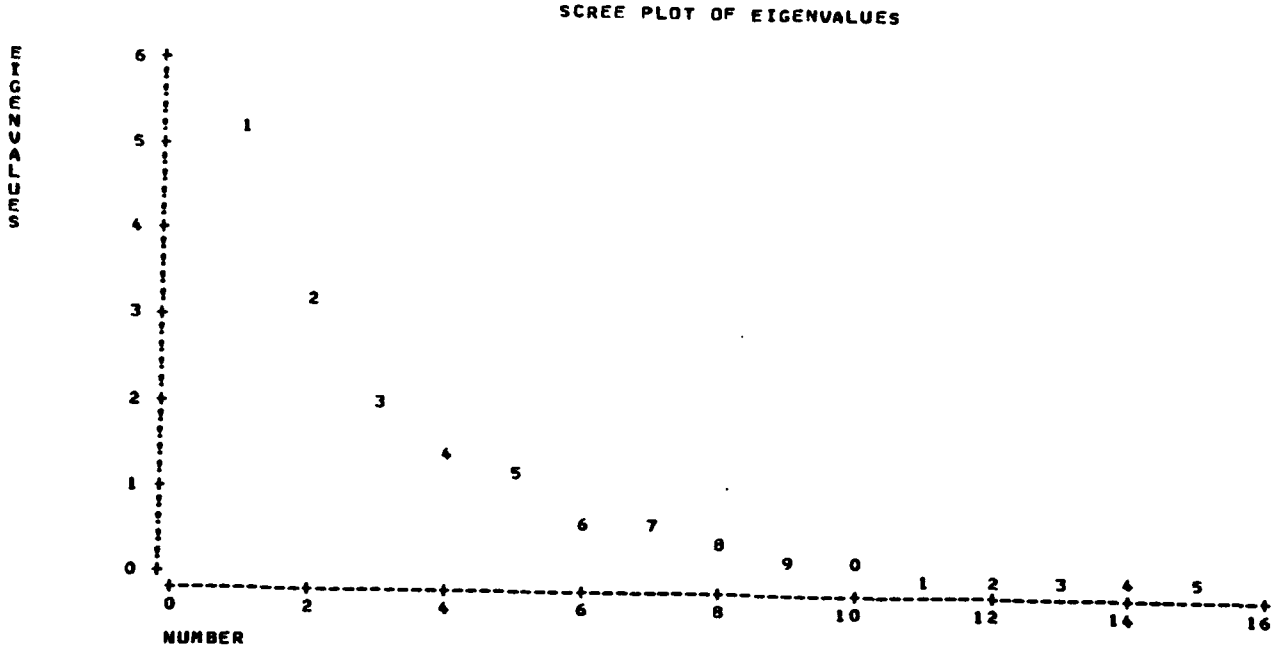
5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4
FACTEUR1	100 *			
FACTEUR2	17	100 *		
FACTEUR3	30	-1	100 *	
FACTEUR4	32	20	-32	100 *

ANNEXE 7.β MARCHE BOISSON

1) VALEURS PROPRES ET POURCENTAGES D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	34.9	34.9
2	20.8	55.7
3	13.0	68.7
4	9.6	78.3
5	8.2	86.4
6	4.0	90.4
7	3.7	94.0
8	2.6	96.6
9	1.5	98.1
10	.75	98.9
11	.42	99.3
12	.30	99.6
13	.24	99.8
14	.13	99.9
15	.00	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX) DANS L'ESPACE DES CINQ PREMIERS FACTEURS (86.4% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4	FACTEUR5
BRAA	15	7	53 *	22	64 *
BRAB	0	-7	19	3	88 *
FRASA	32	4	89 *	16	-10
FRASB	40	-6	86 *	11	-4
LEM	94 *	17	22	5	3
MERA	12	-6	3	94 *	5
MERR	-2	13	13	95 *	1
SABA	94 *	13	14	12	8
TOTT	10	89 *	10	11	7
TRTO	2	95 *	-3	10	8
RAO	46 *	61 *	26	5	10
ELRO	57	36	-1	12	38
TREL	-9	91 *	-1	4	2
IND	98 *	1	16	2	1
ACT	94 *	7	21	6	9

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace à 5 facteurs) TOTAL = 12.96

BRAA	BRAB	FRASA	FRASB	LEM	MERA	MERR	SABA	TOTT
0.76	.81	.94	.92	.96	.89	.93	.93	.83

TRTO	RAO	ELRO	TREL	IND	ACT
.92	.67	.61	.84	.98	.96

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE DANS L'ESPACE DES CINQ PREMIERS FACTEURS (86.4% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4	FACTEUR5
BRAA	-8	8	47 *	22	65 *
BRAB	5	-4	19	1	87 *
FRASA	45 *	9	95 *	-16	-14
FRASB	51 *	0	93 *	10	-9
LEM	97 *	23	44	6	2
MERA	13	12	4	94 *	8
MERR	1	20	16	95 *	3
SABA	94 *	-8	35	16	15
TOTT	6	88 *	9	20	2
TRTO	8	96 *	7	18	12
RAO	53 *	64 *	38	0	10
ELRO	57 *	42	14	12	36
TREL	14	92 *	4	12	6
IND	99 *	6	38	5	7
ACT	97 *	12	43	8	14

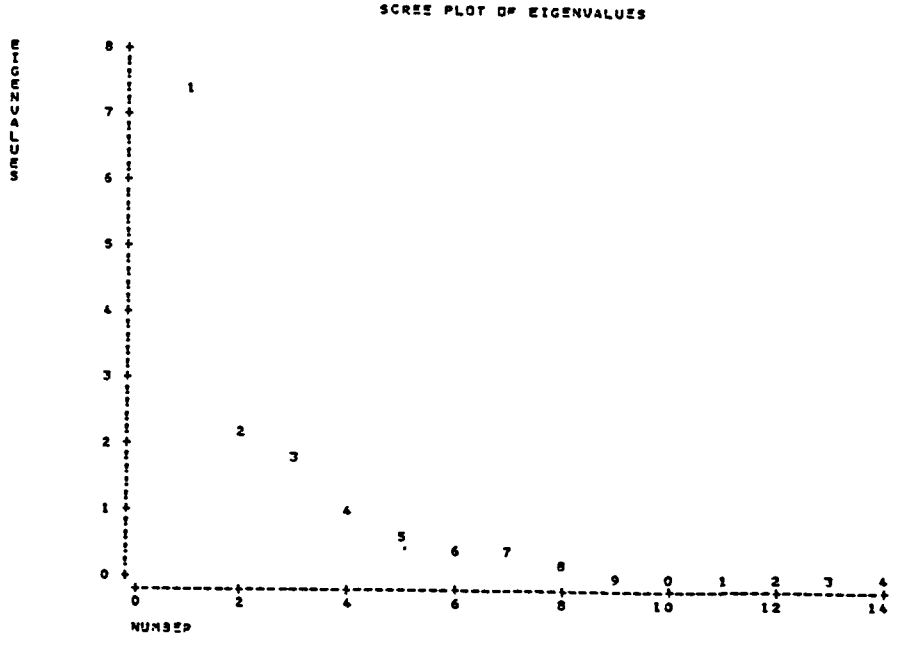
5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3	FACTEUR4	FACTEUR5
FACTEUR1	100 *				
FACTEUR2	13	100 *			
FACTEUR3	35	8	100 *		
FACTEUR4	7	16	4	100 *	
FACTEUR5	9	8	3	7	100 *

ANNEXE 7.Γ MARCHE PARFUM

1) VALEURS PROPRES ET POURCENTAGES D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	52.7	52.7
2	15.8	68.2
3	12.8	81.0
4	6.9	87.9
5	4.8	92.7
6	2.8	95.5
7	2.5	98.0
8	1.2	99.3
9	0.5	99.8
10	.35	100
11	.11	100
12	.03	100
13	.01	100
14	.00	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (81% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
BRAA	3	2	92 *
BRAB	8	2	92 *
FRASA	84 *	5	8
FRASB	88 *	14	19
LEM	91 *	30	4
MERR	20	62 *	7
SABA	11	60 *	25
TOTT	57 *	75 *	14
TRTO	60 *	73 *	7
RAO	83 *	37	11
ELRO	55 *	76 *	20
TREL	59 *	74 *	16
IND	94 *	11	0
ACT	96 *	11	5

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace à 3 facteurs) TOTAL = 11.34

BRAA	BRAB	FRASA	FRASB	LEM	MERR	SABA	TOTT
0.85	.85	.72	.84	.91	.42	.42	.97

TRTO	RAO	ELRO	TREL	IND	ACT
.90	.83	.92	.92	.90	.93

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (81.% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
BRAA	3	0	91 *
BRAB	14	3	92 *
FRASA	83 *	18	18
FRASB	86 *	10	30
LEM	94 *	53	16
MERR	10	53	8
SABA	3	54	24
TOTT	70 *	88 *	23
TRTO	72 *	87 *	16
RAO	87 *	58 *	0
ELRO	68 *	88 *	28
TREL	72 *	87 *	25
INDL	94 *	36	-11
ACT	96 *	36	-7

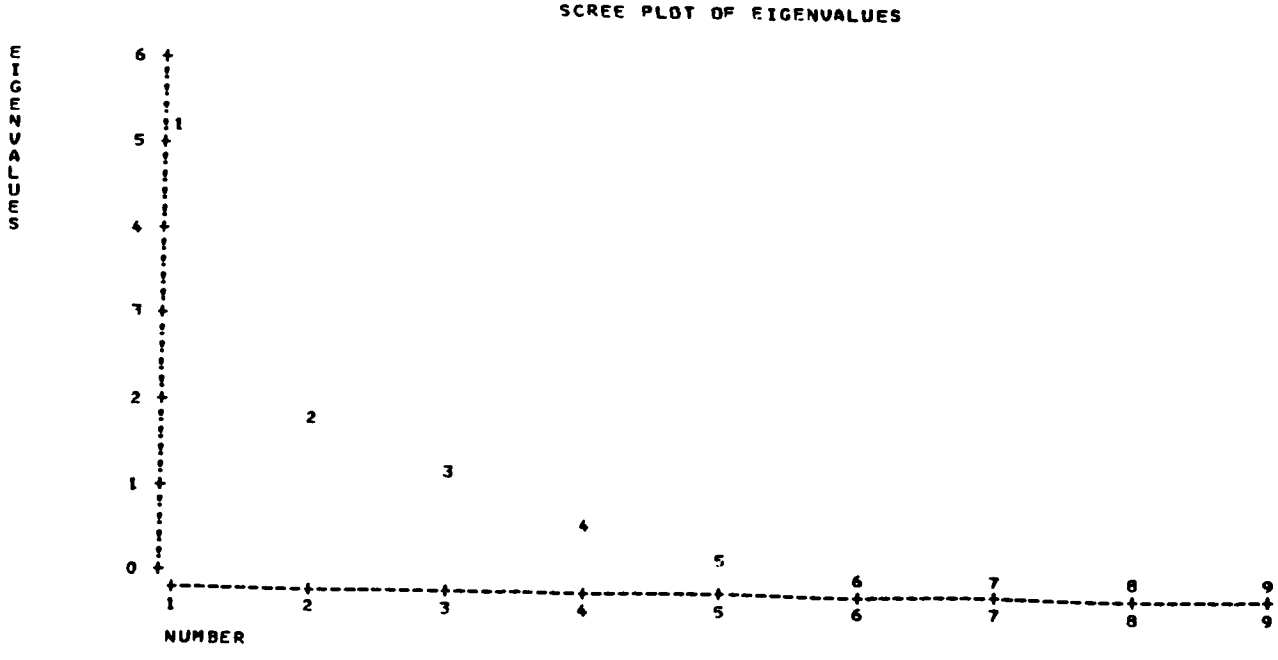
5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3
FACTEUR1	100 *		
FACTEUR2	43	100 *	
FACTEUR3	20	9	100 *

ANNEXE 7.π MARCHE SHAMP

1) VALEURS PROPRES ET POURCENTAGES D'INERTIE EXPLIQUEE

FACTEUR	% VARIANCE	% CUMULE
1	57.6	57.6
2	20.0	77.6
3	12.5	90.0
4	7.4	97.4
5	1.5	98.9
6	0.6	99.5
7	0.2	99.7
8	0.2	99.9
9	0.0	100



2) PATTERN DES FACTEURS APRES ROTATION ORTHOGONALE (VARIMAX)
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (90% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
FRASA	29	93 *	5
FRASB	36	95 *	7
LEM	96 *	20	5
MERA	9	10	95 *
MERR	5	1	97 *
SABA	85 *	28	0
RAO	73 *	15	12
IND	95 *	23	5
ACT	92 *	35	5

3) ESTIMATEURS FINAUX DES COMMUNALITES (dans l'espace à 3 facteurs) TOTAL = 8.1

FRASA	FRASB	LEM	MERA	MERR	SABA	RAO	IND	ACT
.99	.99	.97	.93	.94	.80	.57	.95	.97

4) PATTERN DES FACTEURS APRES ROTATION OBLIQUE
DANS L'ESPACE DES TROIS PREMIERS FACTEURS (90% de la variance)

	FACTEUR1	FACTEUR2	FACTEUR3
FRASA	52	99 *	14
FRASB	57	99 *	12
LEM	98 *	48	14
MERA	17	18	96 *
MERR	11	8	97 *
SABA	89 *	52	7
RAO	75 *	37	18
IND	98 *	51	12
ACT	98 *	61	12

5) CORRELATIONS ENTRE LES FACTEURS

	FACTEUR1	FACTEUR2	FACTEUR3
FACTEUR1	100 *		
FACTEUR2	53	100 *	
FACTEUR3	14	12	100 *

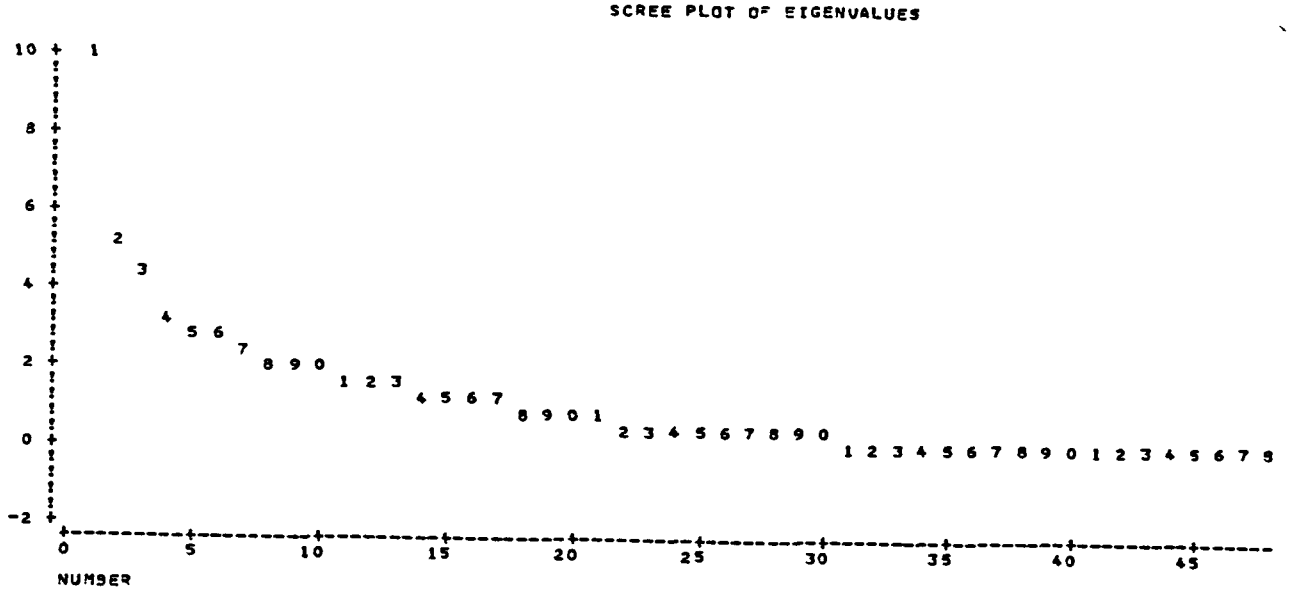
ANNEXE 8 MATRICE MULTI-TRAITS MULTI-METHODES (coeffcient de corrélation de Pearson).

ANNEXE 9 ANALYSE DE LA VALIDITE DE TRAITS ANALYSE FACTORIELLE (coefficient de Pearson)

1) TABLEAU DES VARIANCES EXPLIQUEES PAR LES PREMIERS FACTEURS ET REPRESENTATION GRAPHIQUE

FACTEUR	POURCENTAGE VARIANCE	POURCENTAGE CUMULEE
1	19	19
2	9.8	29
3	8.5	37.5
4	6.3	43.8
5	5.7	49.5
6	5.0	54.6
7	4.7	59.3
8	4.1	63.4
9	3.8	67.2
10	3.7	71.0
11	3.3	74.3
12	3.1	77.4
13	2.8	80.2
14	2.5	82.7
...	...	...

REPRESENTATION DE LA VARIANCE EXPLIQUEE PAR LES FACTEURS



INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8	FACTOR9
CTOR1	100 *	14	10	37 *	13	10	11	32 *	10
CTOR2	14	100 *	6	25	15	14	14	4	25
CTOR3	10	6	100 *	9	22	18	11	13	10
CTOR4	37 *	25	9	100 *	8	4	15	17	14
CTOR5	13	15	22	8	100 *	28 *	13	2	12
CTOR6	10	14	18	4	28 *	100 *	11	11	15
CTOR7	11	14	11	15	13	11	100 *	12	6
CTOR8	32 *	4	13	17	2	11	12	100 *	12
CTOR9	10	25	10	14	12	15	6	12	100 *
CTOR10	15	15	20	11	13	12	7	7	25
CTOR11	3	-5	-2	10	6	5	13	17	-4
CTOR12	9	14	11	3	14	16	12	4	-7
CTOR13	6	19	6	9	14	18	3	15	10
CTOR14	10	-7	20	10	-1	6	11	18	-8
CTOR15	9	16	6	6	7	4	9	10	-8
CTOR16	-1	-7	-4	5	-11	2	6	6	-20
CTOR17	-4	9	2	10	18	10	5	8	16

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8	FACTOR9
BRAAHUI	3	13	-6	5	5	25	19	-4	-2
FRASAHUI	10	4	4	8	19	11	11	11	9
FRASBHUI	9	16	0	10	28 *	18	5	19	-2
LEMHUI	11	8	26	11	94 *	31 *	16	7	5
MERAHUI	33 *	3	14	24	-1	9	13	85 *	2
MERRHUI	39 *	10	13	23	-7	10	18	88 *	-3
SABHUI	17	5	2	2	16	33 *	11	62 *	10
TOTHUI	3	9	29 *	7	26	75 *	20	9	2
TRTOHUI	12	13	19	5	28 *	76 *	42	22	14
RAOHUI	36 *	1	12	0	25	3	18	69 *	18
ELROHUI	27 *	6	32 *	5	29 *	39 *	19	24	7
TRELHUI	26	4	7	19	27 *	22	18	18	32 *
INDHUI	7	8	28 *	13	94 *	28 *	19	8	2
ACTHUI	6	8	39 *	13	94 *	24	18	7	5
BRAABOI	7	6	10	11	9	8	7	7	23
BRABBOI	10	6	7	1	15	1	13	5	69 *
FRASABOI	5	45 *	11	20	4	30 *	5	13	86 *
FRASBBOI	8	52 *	15	21	2	18	1	10	75 *
LEMBOI	19	97 *	0	25	9	10	22	3	31 *
MERABOI	27 *	8	17	1	20	92 *	17	19	10
MERRBOI	5	7	14	9	20	89 *	16	4	17
SABBOI	12	94 *	3	8	13	4	11	13	2
TOTBOI	4	12	21	11	-6	12	88 *	18	4
TRTOBOI	6	19	4	14	28 *	18	95 *	3	-4
RAOBOI	11	62 *	16	25	8	3	55 *	8	37 *
ELROBOI	67 *	47 *	16	27 *	29 *	14	15	2	16
TRELBOI	17	20	5	14	25	13	90 *	16	5
INDBOI	15	96 *	0	29 *	13	13	9	6	-27 *
ACTBOI	7	95 *	6	27 *	7	13	17	4	-27 *
BRAAGOU	2	53 *	7	8	19	12	0	0	-14
BRABGOU	12	13	2	9	12	15	13	28 *	7
FRASAGOU	33 *	14	2	66 *	14	3	26	40 *	23
FRASBGOU	27 *	40 *	9	79 *	17	6	16	27 *	20
LEMGOU	47 *	12	12	94 *	8	1	16	12	0
MERRGOU	21	19	26 *	10	19	15	14	18	19
SABGOU	8	16	12	19	24	7	20	5	0
TOTGOU	95 *	10	7	43 *	6	9	11	45 *	11
TRTOGOU	95 *	7	7	42 *	5	8	9	43 *	9
RAOGOU	38 *	0	15	74 *	-4	-1	21	11	-9
ELROGOU	99 *	16	11	42 *	4	12	11	38 *	4
TRELGOU	97 *	10	10	37 *	3	10	10	32 *	2
INDGOU	33 *	48 *	3	78 *	21	12	3	8	32 *
ACTGOU	34 *	37 *	6	89 *	18	10	3	8	22
FRASASMA	17	11	43 *	15	4	4	7	18	19
FRASBSMA	21	8	47 *	15	5	5	9	21	20
LEMSMA	3	10	98 *	10	29 *	21	10	9	10
MERASMA	3	13	11	9	5	12	5	13	9
MERRSMA	5	17	7	11	11	17	4	16	6
SABSMA	7	13	47 *	15	15	35 *	8	12	22
RAOSMA	12	5	82 *	11	3	4	8	30 *	13
INDSMA	15	10	93 *	7	43 *	30 *	11	1	11
ACTSMA	11	7	93 *	4	45 *	25	13	4	13

2) PATTERN DES 9 PREMIERS FACTEURS APRES ROTATION ORTHOGONALE DANS L'ESPACE DES DIX SEPT PREMIERS FACTEURS DONT LA VALEUR PROPRE EST SUPERIEURE A 1 (88.8 % de la variance).

ROTATED FACTOR PATTERN

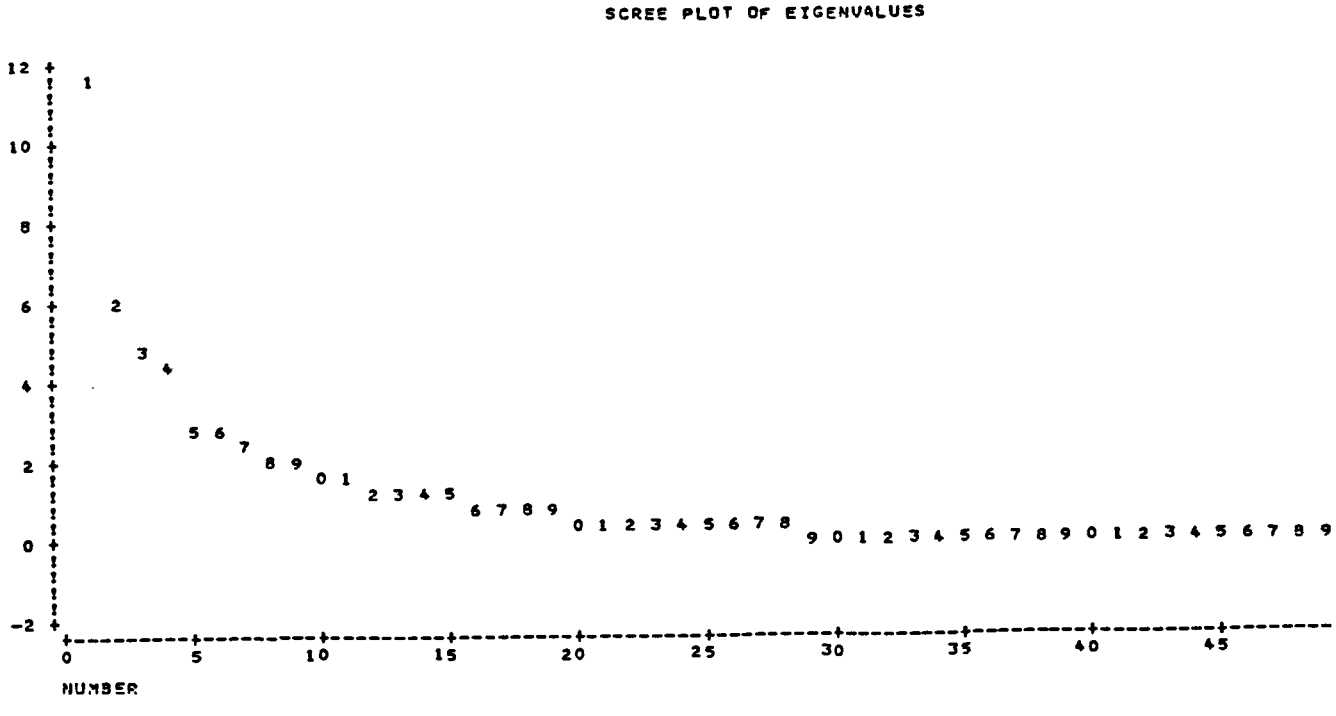
	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8	FACTOR9
BRAAHUI	-1	6	-10	0	-1	20	15	-8	-3
FRASAHUI	8	1	3	1	13	4	4	2	9
FRASBHUI	4	14	-3	3	23 *	12	-4	11	-5
LEMHUI	3	2	14	6	92 *	18	7	3	1
MERAHUI	21	-1	8	16	-3	4	6	82 *	-2
MERRHUI	28 *	7	6	11	-9	6	10	82 *	-6
SABHUI	8	1	-1	-2	9	28 *	6	59 *	3
TOTHUI	-3	4	17	3	15	74 *	13	1	-3
TRTOHUI	6	7	9	0	20	74 *	-9	14	9
RAOHUI	28 *	-1	5	-9	22	-4	0	65 *	12
ELROHUI	22	1	16	-7	25 *	36 *	6	13	4
TRELHUI	21	-2	-9	10	27 *	20	3	7	31 *
INDHUI	-1	1	18	9	93 *	12	10	4	-2
ACTHUI	-2	1	30 *	10	92 *	9	9	3	1
BRAABOI	-3	-4	7	8	3	4	2	-1	27 *
BRABBOI	5	-3	1	-4	9	-7	10	-2	74 *
FRASABOI	-3	35 *	7	12	-5	24 *	-1	5	84 *
FRASBBOI	2	43 *	11	11	-5	12	-5	1	74 *
LEMBOI	11	95 *	-1	10	-2	1	15	-2	16
MERABOI	22	1	10	-6	4	89 *	10	9	3
MERRBOI	-2	0	5	6	8	88 *	9	-3	11
SABBOI	6	52 *	-1	0	10	-1	5	12	-1
TOTBOI	-2	7	14	5	-12	10	87 *	12	1
TRTOBOI	0	11	-2	7	21	10	94 *	-3	-1
RAOBOI	3	58 *	13	12	-2	-5	49 *	1	25 *
ELROBOI	60 *	39 *	11	12	18	4	7	-11	5
TRELBOI	11	12	-1	4	18	3	88 *	10	-1
INDBOI	6	95 *	-1	16	3	4	1	2	11
ACTBOI	-3	94 *	4	15	-4	5	9	1	10
BRAAGOU	-2	50 *	3	0	14	5	-8	-2	-19
BRABGOU	3	5	-4	5	6	7	6	29 *	11
FRASAGOU	19	4	-5	58 *	9	-2	17	27 *	16
FRASBGOU	10	27 *	4	74 *	9	0	8	19	8
LEMGOU	33 *	0	6	90 *	3	-1	8	-2	-7
MERRGOU	19	11	24 *	-1	18	7	10	11	24 *
SABGOU	4	10	9	13	16	-1	16	-2	-7
TOTGOU	91 *	2	0	23	1	4	2	24 *	6
TRTOGOU	91 *	0	-1	22	1	4	0	22	5
RAOGOU	30 *	-11	10	69 *	-4	0	14	-3	-6
ELROGOU	95 *	9	6	21	-2	8	3	17	-3
TRELGOU	95 *	3	5	17	-3	7	2	11	-4
INDGOU	16	38 *	-2	73 *	11	4	-5	-1	13
ACTGOU	17	25 *	1	85 *	10	4	-5	-2	7
FRASASHA	8	5	37 *	9	-6	-3	1	11	6
FRASBSHA	12	3	39 *	8	-5	-2	2	13	7
LEMSHA	-4	4	96 *	5	16	10	3	3	1
MERASHA	0	5	8	4	-3	4	2	4	5
MERRSHA	0	9	5	7	3	8	2	9	2
SABSHA	-4	7	43 *	13	-2	26 *	2	9	10
RAOSHA	7	-1	82 *	4	-7	-5	1	21	8
INDSHA	8	3	89 *	1	29 *	17	4	-6	1
ACTSHA	5	1	89 *	-1	32 *	12	6	-3	4

ANNEXE 10 ANALYSE DE LA VALIDITE DE TRAITS
ANALYSE FACTORIELLE (coefficient de Spearman)

1) TABLEAU DES VARIANCES EXPLIQUEES PAR LES PREMIERS FACTEURS ET REPRESENTATION GRAPHIQUE

FACTEUR	POURCENTAGE VARIANCE	POURCENTAGE CUMULEE
1	22.3	22.3
2	11.7	34.0
3	9.0	43.0
4	8.6	51.6
5	5.4	57.0
6	5.3	62.3
7	4.7	67.0
8	3.8	70.8
9	3.6	74.4
10	3.3	77.8
11	3.0	80.8
12	2.6	83.3
13	2.2	85.5
14	2.1	87.6
...	...	...

REPRESENTATION DE LA VARIANCE EXPLIQUEE PAR LES FACTEURS



2) PATTERN DES 9 PREMIERS FACTEURS APRES ROTATION ORTHOGONALE DANS L'ESPACE DES 15 PREMIERS FACTEURS DONT LA VALEUR PROPRE EST SUPERIEURE A 1 (89.7 % de la variance).

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8
BRAAHUI	-1	9	20	27 *	3	21	25 *	55 *
FRASAHUI	10	13	1	3	2	1	-1	3
FRASBHUI	13	20	1	8	2	7	0	16
LEMHUI	0	85 *	-1	8	0	-4	-1	-7
MERAHUI	23	5	8	7	6	83 *	1	8
MERRHUI	25 *	3	10	17	6	84 *	0	2
SABHUI	21	4	6	-9	6	46 *	16	18
TOTHUI	0	84 *	21	2	4	6	6	6
TRTOHUI	16	82 *	6	5	-1	25 *	13	-6
RAOHUI	13	38 *	5	9	2	22	5	6
ELROHUI	4	89 *	19	-4	6	2	10	14
TRELHUI	2	91 *	0	2	3	0	10	0
INDHUI	3	82 *	3	12	0	-5	-5	0
ACTHUI	4	82 *	3	10	0	-5	-4	0
BRAABOI	3	8	6	-3	2	-2	15	11
BRABBOI	0	8	4	2	3	6	15	10
FRASABOI	7	18	4	34 *	3	-2	85 *	-4
FRASBBOI	7	12	1	42 *	0	4	80 *	1
LEMBOI	20	2	1	93 *	12	-4	15	10
MERABOI	17	24	22	4	4	2	-2	17
MERRBOI	3	44 *	26 *	-2	12	-1	13	8
SABBOI	6	1	4	94 *	2	16	11	12
TOTBOI	1	6	19	3	86 *	18	4	3
TRTOBOI	12	4	7	1	96 *	-2	1	4
RAOBOI	23	-5	14	42 *	54 *	-21	26 *	-2
ELROBOI	60 *	1	-4	40 *	31 *	-5	-8	-2
TRELBOI	2	4	-5	8	93 *	3	-4	9
INDBOI	15	10	0	96 *	1	7	12	10
ACTBOI	11	7	1	95 *	5	4	17	14
BRAAGOU	-1	5	6	37 *	2	3	-7	82 *
BRABGOU	9	0	2	9	14	1	-4	86 *
FRASAGOU	71 *	-1	-1	-2	22	6	38 *	7
FRASBGOU	72 *	2	15	18	15	-6	31 *	18
LEMGOU	96 *	4	10	6	7	3	2	7
MERRGOU	5	1	-4	8	14	24	38 *	10
SABGOU	11	4	18	7	5	1	-1	8
TOTGOU	79 *	12	2	3	-1	24	-11	-7
TRTOGOU	80 *	9	4	2	0	26 *	-9	-10
RAOGOU	91 *	4	4	-6	8	9	-3	8
ELROGOU	79 *	6	-2	9	-1	29 *	-18	-3
TRELGOU	82 *	3	-1	8	-2	30 *	-14	-4
INDGOU	88 *	0	16	22	-5	-5	18	2
ACTGOU	89 *	-1	16	20	-4	-4	20	0
FRASASHA	11	8	64 *	-4	6	-1	5	29 *
FRASBSHA	12	12	68 *	-4	3	3	4	25 *
LEMSHA	12	4	95 *	2	4	13	-2	4
MERASHA	7	2	12	2	0	4	2	0
MERRSHA	7	5	4	5	0	14	1	11
SABSHA	-1	12	88 *	6	5	-5	5	-9
RAOSHA	17	5	69 *	7	5	37 *	6	-2
INDSHA	6	10	95 *	0	6	-5	0	-1

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8
FACTOR1	100 *	6	16	26	19	22	11	5
FACTOR2	6	100 *	19	12	4	7	11	6
FACTOR3	16	19	100 *	2	14	10	21	28
FACTOR4	26	12	2	100 *	16	2	22	15
FACTOR5	19	4	14	16	100 *	-1	20	18
FACTOR6	22	7	10	2	-1	100 *	6	11
FACTOR7	11	11	21	22	20	6	100 *	32 *
FACTOR8	5	6	28	15	18	11	32 *	100 *
FACTOR9	25	18	2	19	6	-1	-27	-10
FACTOR10	16	22	6	10	8	4	3	12
FACTOR11	14	2	18	3	11	3	24	27
FACTOR12	11	27	1	9	7	2	-14	-15
FACTOR13	15	9	-3	19	3	18	-31 *	-11
FACTOR14	1	12	-8	5	1	6	15	4
FACTOR15	12	11	8	14	3	-7	-31 *	-8

	FACTOR9	FACTOR10	FACTOR11	FACTOR12	FACTOR13	FACTOR14	FACTOR15
FACTOR1	25	16	14	11	15	1	12
FACTOR2	18	22	2	27	9	12	11
FACTOR3	2	6	18	1	-3	-8	8
FACTOR4	19	10	3	9	19	5	14
FACTOR5	6	8	11	7	3	1	3
FACTOR6	-1	4	3	2	18	6	-7
FACTOR7	-27	3	24	-14	-31 *	15	-31 *
FACTOR8	-10	12	27	-15	-11	4	-8
FACTOR9	100 *	18	-6	10	48 *	-9	61 *
FACTOR10	18	100 *	6	4	3	15	6
FACTOR11	-6	6	100 *	-1	-19	1	-8
FACTOR12	10	4	-1	100 *	24	5	10
FACTOR13	48 *	3	-19	24	100 *	-14	50 *
FACTOR14	-9	15	1	5	-14	100 *	-30 *
FACTOR15	61 *	6	-8	10	50 *	-30	100 *

NOTE: PRINTED VALUES ARE MULTIPLIED BY 100 AND ROUNDED TO THE NEAREST INTEGER
VALUES GREATER THAN 0.304719 HAVE BEEN FLAGGED BY AN '*'

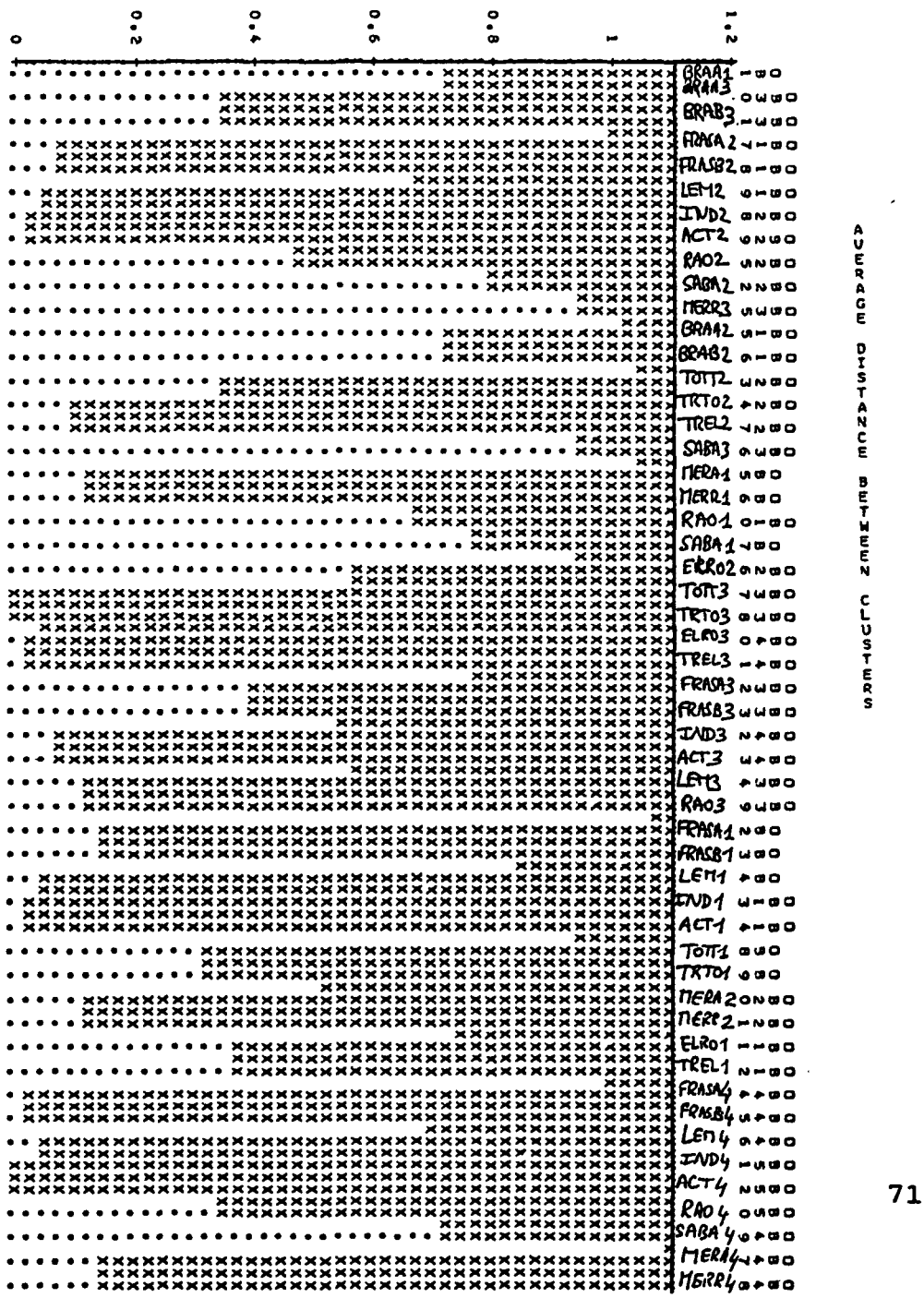
FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6	FACTOR7	FACTOR8
BRAAMUI	7	16	27	31 *	12	28	35 *	63 *
FRASAMUI	15	20	4	8	6	4	2	10
FRASBMUI	19	27	7	13	8	12	8	24
LEMMUI	5	86 *	5	19	2	-7	-5	-10
MERAMUI	33 *	10	14	13	9	84 *	2	14
MERRMUI	37 *	9	15	25	11	84 *	-2	9
SABMUI	29 *	9	12	1	10	43 *	8	18
TOTMUI	4	86 *	29 *	6	8	11 *	16	13
TRTOMUI	21	86 *	15	12	4	31 *	19	0
RADMUI	21	40 *	10	20	6	16	-2	3
ELROMUI	8	90 *	28	2	10	6	17	19
TRELMUI	6	90 *	9	11	6	1	10	0
INDMUI	8	84 *	8	23	4	-6	-5	-2
ACTMUI	9	84 *	9	21	4	-5	-3	-2
BRAABOI	7	14	9	5	7	2	16	16
BRABBOI	2	11	8	8	6	0	11	12
FRASABOI	16	25	12	42 *	14	3	83 *	12
FRASBBOI	16	19	8	49 *	11	9	80 *	16
LEMBOI	31 *	7	7	96 *	24	-1	34 *	24
MERABOI	22	32 *	26	13	11	10	1	16
MERRBOI	9	52 *	31 *	7	18	3	12	8
SABBOI	18	9	9	93 *	13	18	31 *	26
TOTBOI	10	9	24	8	86 *	16	14	14
TRTOBOI	19	6	14	9	97 *	-1	12	14
RADBOI	33 *	0	21	49 *	63 *	-19	37 *	13
ELROBOI	65 *	2	3	52 *	40 *	-1	1	2
TRELBOI	9	5	0	15	92 *	2	4	14
INDBOI	26	15	6	98 *	13	10	33 *	25
ACTBOI	23	13	7	97 *	17	8	38 *	29 *
BRAAGOU	7	11	15	40 *	11	7	9	84 *
BRABGOU	14	7	10	18	22	5	5	81 *
FRASAGOU	75 *	5	10	10	33 *	18	44 *	18
FRASBGOU	76 *	7	25	28 *	29 *	10	46 *	29 *
LEMGOU	96 *	7	18	19	18	19	12	12
MERRGOU	12	5	0	17	19	32 *	33 *	15
SABGOU	16	9	19	16	11	12	-1	9
TOTGOU	82 *	13	7	21	6	29 *	-20	-14
TRTOGOU	83 *	11	8	20	8	32 *	-18	-17
RADGOU	91 *	6	13	7	18	25 *	4	10
ELROGOU	82 *	7	3	26	6	36 *	-24	-11
TRELGOU	84 *	5	4	25	6	37 *	-19	-11
INDGOU	90 *	4	25	32 *	10	10	31 *	12
ACTGOU	91 *	3	25	30 *	10	11	32 *	10
FRASASHA	18	14	68 *	1	13	-3	8	38 *
FRASBSHA	19	18	72 *	2	11	1	6	34 *
LEMSHA	21	15	95 *	2	12	22	12	22
MERASHA	13	6	18	3	6	8	14	16
MERRSHA	14	9	12	6	6	21	16	26
SABSHA	8	23	97 *	7	13	0	14	7
RADSHA	26	14	71 *	8	12	44 *	18	17
INDSHA	14	21	95 *	1	14	4	13	16

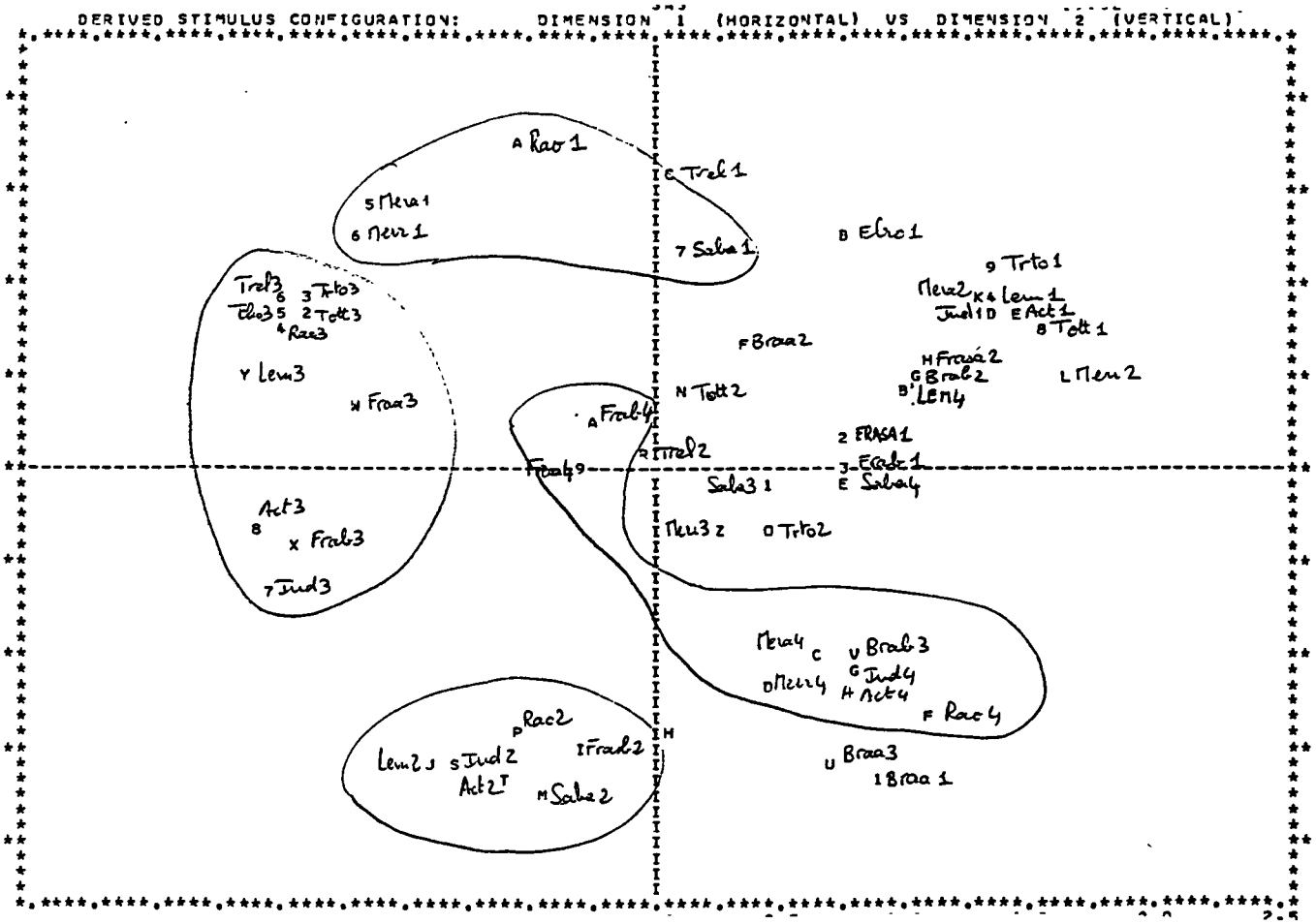
ANNEXE 11 ANALYSE DE LA VALIDITE DE TRAIT : UTILISATION DE LA TYPOLOGIE HIERARCHIQUE ET DE L'ANALYSE DES SIMILARITES

ANNEXE 11 A 52 INDICES - COEFFICIENT DE CORRELATION DE PEARSON

1) RESULTATS DE L'ANALYSE TYPOLOGIQUE HIERARCHIQUE METHODE DU LIEN MOYEN
(Chaque indice est affecté d'un suffixe correspondant au marché sur lequel il a été calculé. 1 pour HUILE 2 pour BOISSON 3 pour PARFUM 4 pour GOUT. Par exemple LEM4 correspond à l'indice LEM calculé sur le marché SHAMP.)

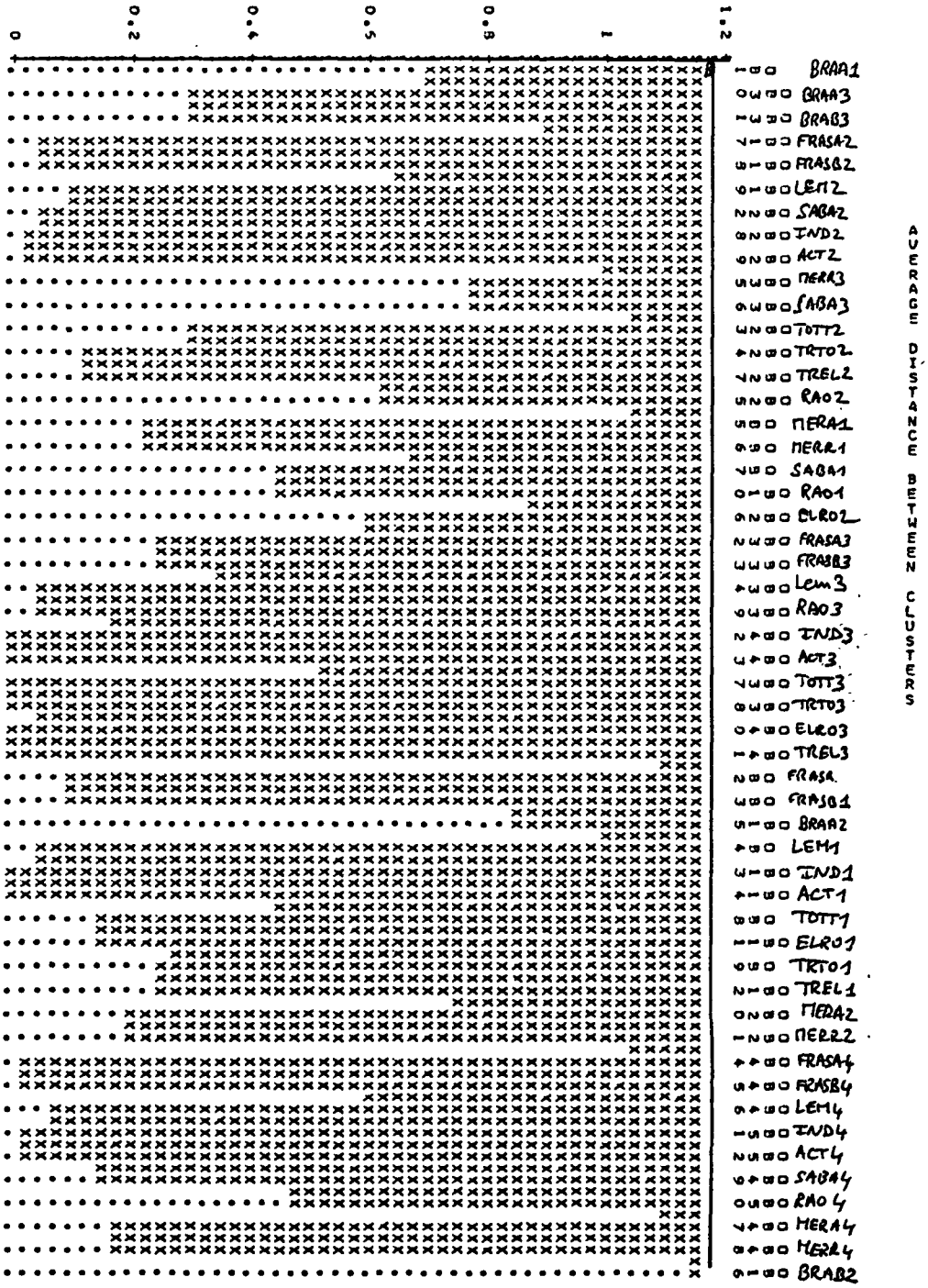


2) REPRESENTATION DES INDICES AVEC LE MODELE INDSCAL (version non métrique).
La carte suivante correspond à la représentation des indices selon les deux premieres dimensions, dans l'espace à trois dimension (Stress = 0.25).

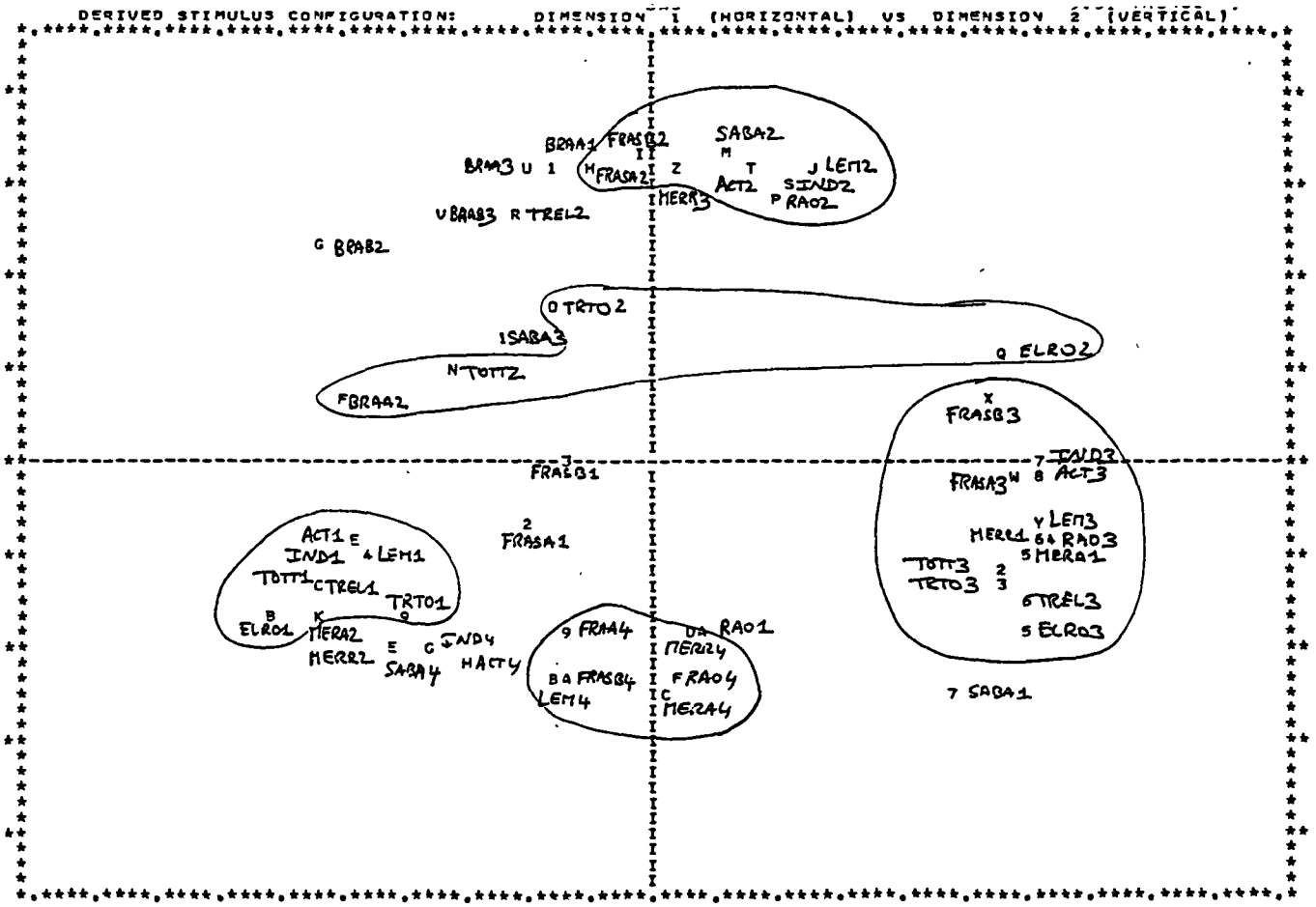


ANNEXE 11 B 52 INDICES - COEFFICIENT DE RANG DE SPEARMAN

1) RESULTATS DE L'ANALYSE TYPOLOGIQUE HIERARCHIQUE
METHODE DU LIEN MOYEN
(Chaque indice est affecté d'un suffixe correspondant au marché sur lequel il a été calculé. 1 pour HUILE 2 pour BOISSON 3 pour PARFUM 4 pour GOUT. Par exemple LEM4 correspond à l'indice LEM calculé sur le marché SHAMP.)

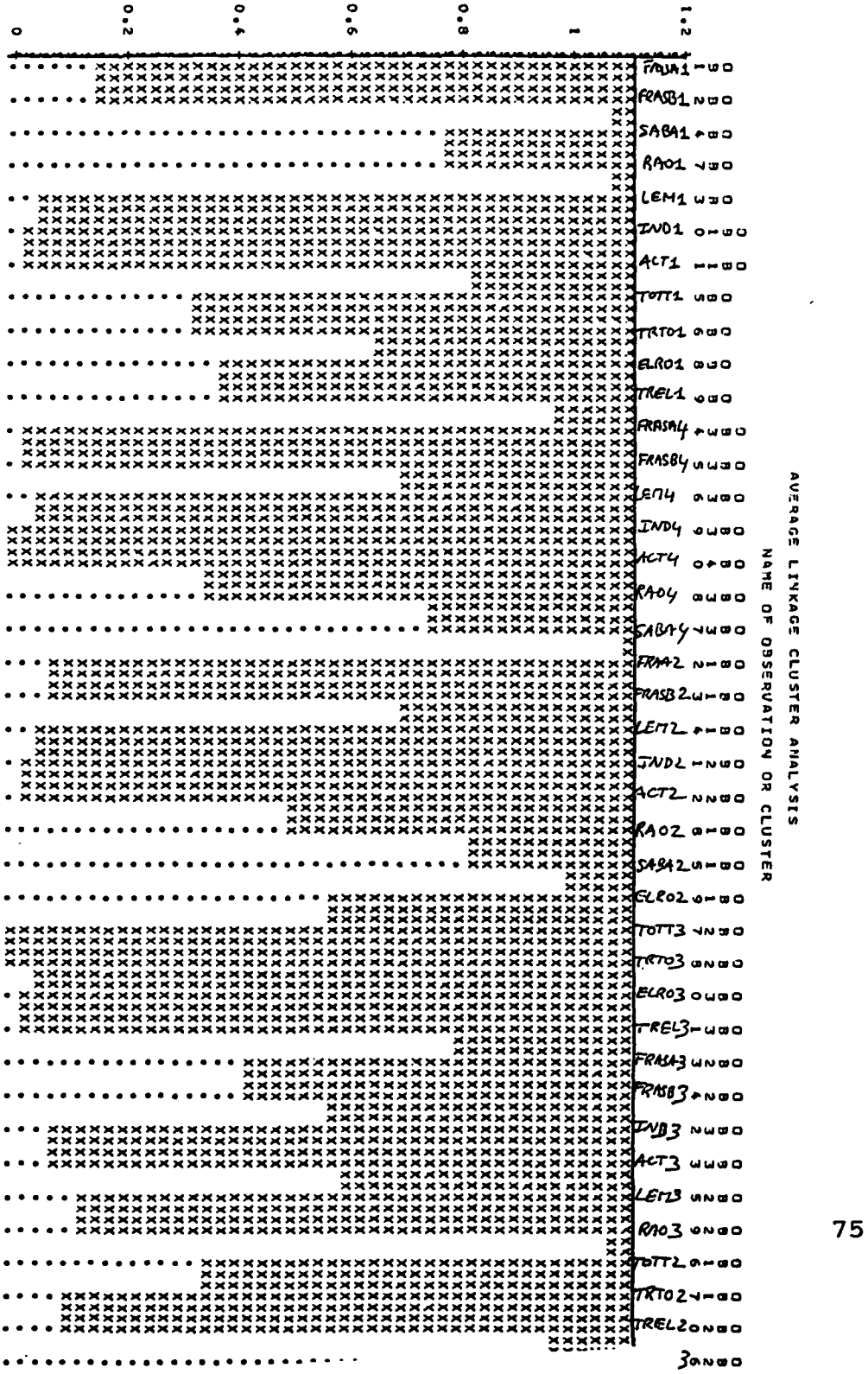


2) REPRESENTATION DES INDICES AVEC LE MODELE INDSCAL (version non métrique).
La carte suivante correspond à la représentation des indices selon les deux premieres dimensions, dans l'espace à trois dimension (Stress = 0.22).

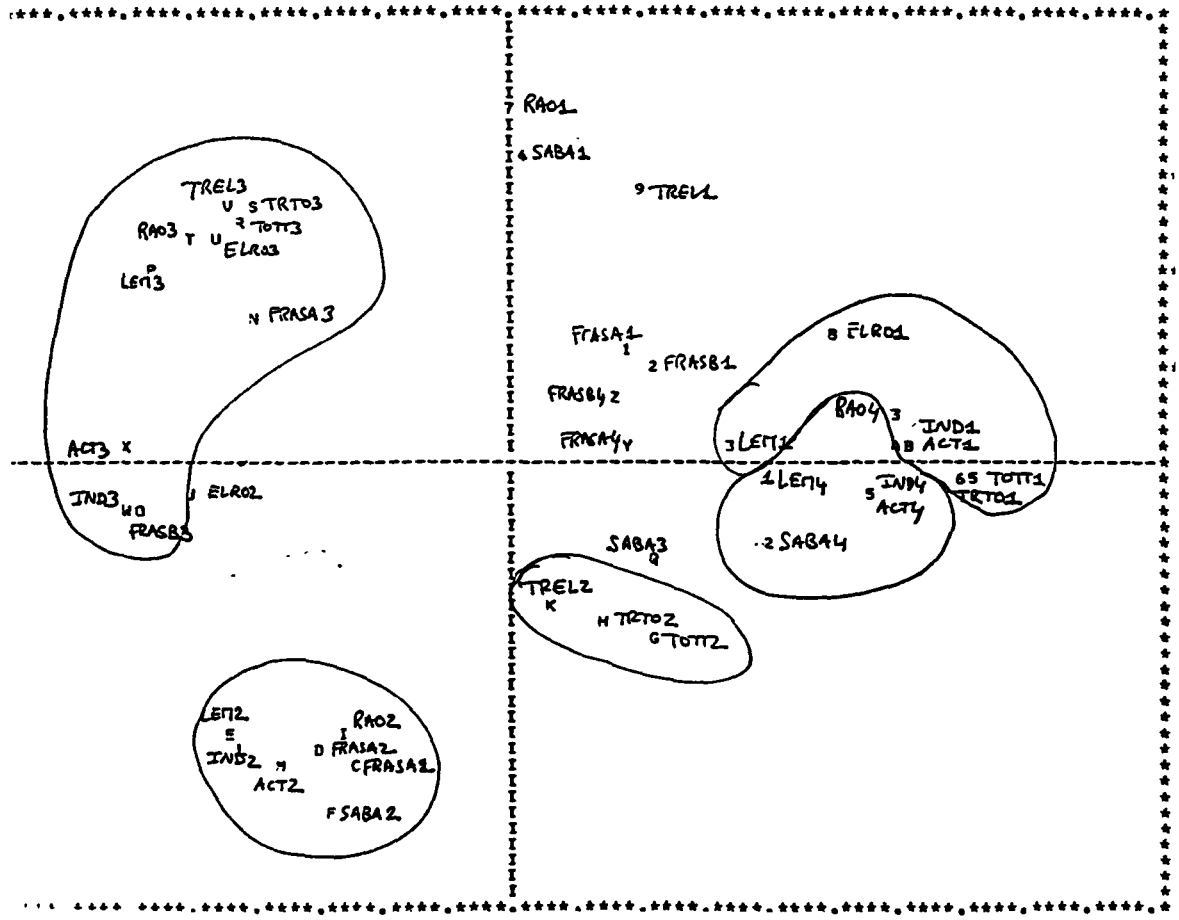


ANNEXE 11 C 40 INDICES - COEFFICIENT DE CORRELATION DE PEARSON (BRAA, BRAB, MERR, MERA ont été supprimés).

1) RESULTATS DE L'ANALYSE TYPOLOGIQUE HIERARCHIQUE METHODE DU LIEN MOYEN
(Chaque indice est affecté du suffixe correspondant au marché sur lequel il a été calculé. 1 pour HUILE 2 pour BOISSON 3 pour PARFUM 4 pour GOUT. Par exemple LEM4 correspond à l'indice LEM calculé sur le marché SHAMP.)

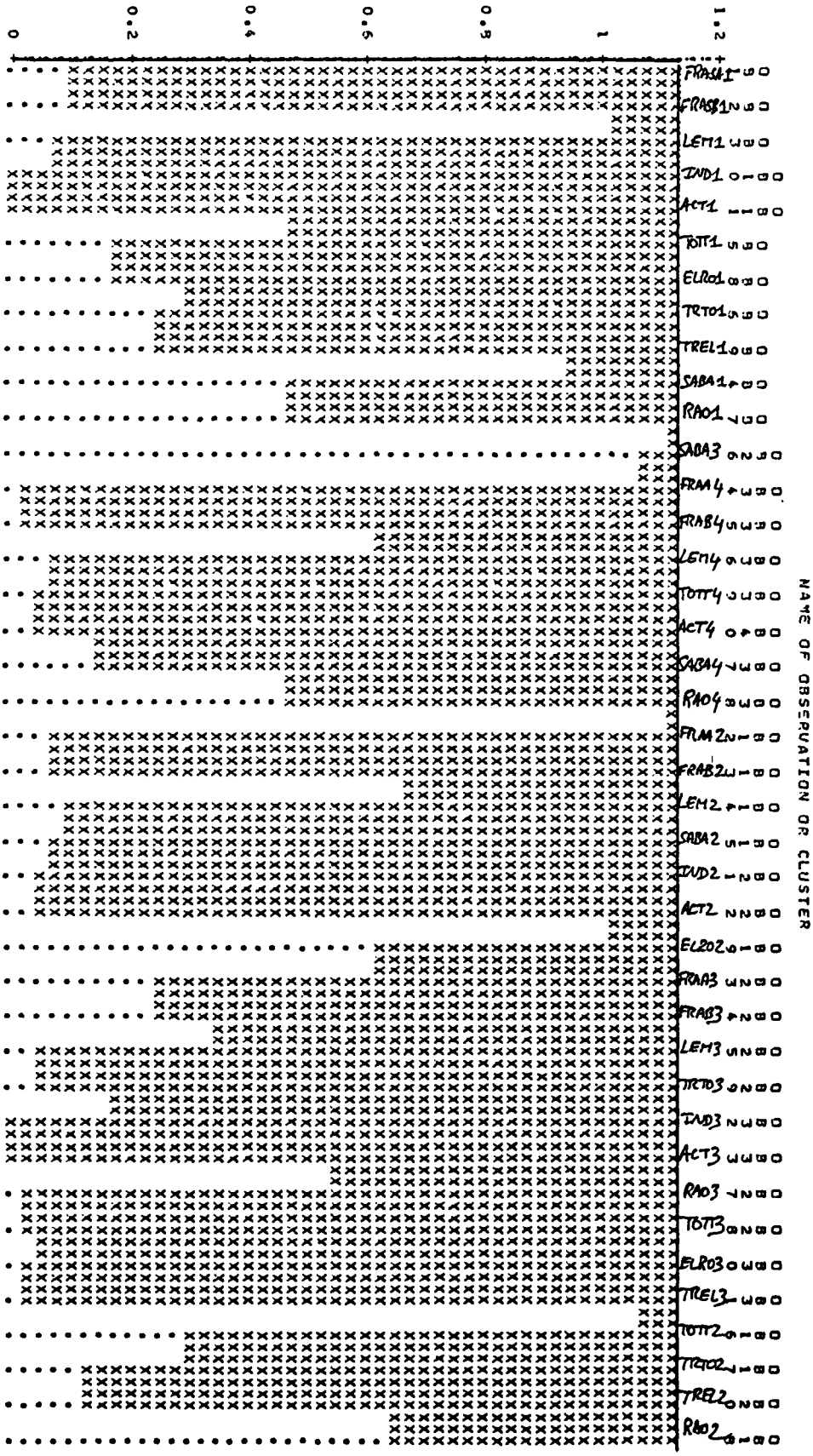


2) REPRESENTATION DES INDICES AVEC LE MODELE INDSCAL (version non métrique).
La carte suivante correspond à la représentation des indices selon les deux premieres dimensions, dans l'espace à trois dimension (Stress = 0.19).

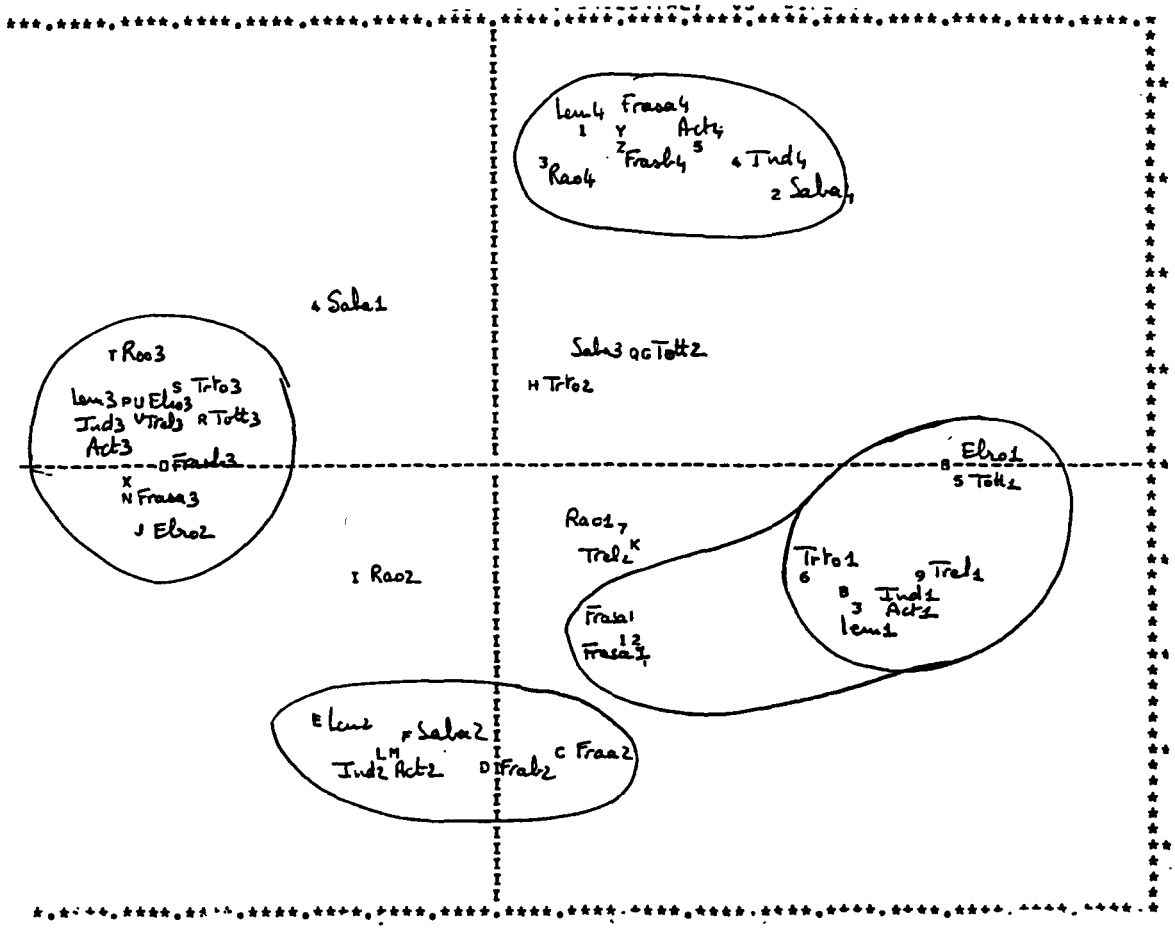


ANNEXE 11 D 40 INDICES - COEFFICIENT DE RANG DE SPEARMAN

1) RESULTATS DE L'ANALYSE TYPOLOGIQUE HIERARCHIQUE
METHODE DU LIEN MOYEN
(Chaque indice est affecté du suffixe correspondant au marché sur lequel il a été calculé. 1 pour HUILE 2 pour BOISSON 3 pour PARFUM 4 pour GOUT. Par exemple LEM4 correspond à l'indice LEM calculé sur le marché SHAMP.)



2) REPRESENTATION DES INDICES AVEC LE MODELE INDSCAL (version non métrique).
La carte suivante correspond à la représentation des indices selon les deux premieres dimensions, dans l'espace à trois dimension (Stress = 0.18).



ANNEXE 12.A ANALYSE DE LA VALIDITE DE TRAITS
ANALYSE FACTORIELLE CONFIRMATOIRE
METHODE DES MOINDRES CARRES NON-PONDERES
(coefficient de Pearson, 52 indices)

1) REPRESENTATION DE LA MATRICE CIBLE POUR LA ROTATION DANS L'ESPACE DES QUATRE PREMIERS FACTEURS (ces quatre facteurs représentent 43.8% de la variance, voir Annexe 9 pour les pourcentages de variance expliqués par les facteurs).

TARGET MATRIX FOR PROCRUSTEAN TRANSFORMATION

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
BRAAHUI	100 *	0	0	0
FRASAHUI	100 *	0	0	0
FRASBHUI	100 *	0	0	0
LEMHUI	100 *	0	0	0
MERAHUI	100 *	0	0	0
MERRHUI	100 *	0	0	0
SABHUI	100 *	0	0	0
TOTHUI	100 *	0	0	0
TRTOHUI	100 *	0	0	0
RAOHUI	100 *	0	0	0
ELROHUI	100 *	0	0	0
TRELHUI	100 *	0	0	0
INDHUI	100 *	0	0	0
ACTHUI	100 *	0	0	0
BRAABOI	0	100 *	0	0
BRABBOI	0	100 *	0	0
FRASABOI	0	100 *	0	0
FRASBBOI	0	100 *	0	0
LEMBOI	0	100 *	0	0
MERABOI	0	100 *	0	0
MERRBOI	0	100 *	0	0
SABBOI	0	100 *	0	0
TOTBOI	0	100 *	0	0
TRTOBOI	0	100 *	0	0
RAOBOI	0	100 *	0	0
ELROBOI	0	100 *	0	0
TRELBOI	0	100 *	0	0
INDBOI	0	100 *	0	0
ACTBOI	0	100 *	0	0
BRAAGOU	0	0	100 *	0
BRABGOU	0	0	100 *	0
FRASAGOU	0	0	100 *	0
FRASBGOU	0	0	100 *	0
LEMGOU	0	0	100 *	0
MERRGOU	0	0	100 *	0
SABGOU	0	0	100 *	0
TOTGOU	0	0	100 *	0
TRTOGOU	0	0	100 *	0
RAOGOU	0	0	100 *	0
ELROGOU	0	0	100 *	0
TRELGOU	0	0	100 *	0
INDGOU	0	0	100 *	0
ACTGOU	0	0	100 *	0
FRASASHA	0	0	0	100 *
FRASBSHA	0	0	0	100 *
LEMSHA	0	0	0	100 *
MERASHA	0	0	0	100 *
MERRSHA	0	0	0	100 *
SABSHA	0	0	0	100 *
RAOSHA	0	0	0	100 *
INDSHA	0	0	0	100 *

2) PATTERN DES FACTEURS APRES ROTATION SOUS CONTRAINTES

CORRELATION ENTRE LES FACTEURS

	F1	F2	F3	F4
F1	100 *			
F2	-11	100 *		
F3	-3	-6	100 *	
F4	30	12	8	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
BRAAHUI	14	21	6	17
FRASAHUI	29	12	8	3
FRASBHUI	35 *	18	10	3
LEMHUI	76 *	17	-7	28
MERAHUI	23	-7	46 *	17
MERRHUI	25	-5	52 *	15
SABHUI	27	4	18	6
TOTHUI	61 *	15	-4	25
TRTOHUI	53 *	13	4	17
RAOHUI	28	-9	32	17
ELROHUI	58 *	1	18	37 *
TRELHUI	40 *	5	24	18
INDHUI	72 *	18	-8	29
ACTHUI	70 *	18	-9	38 *
BRAABOI	13	16	10	16
BRABBOI	11	23	7	19
FRASABOI	1	53 *	12	17
FRASBBOI	-5	55 *	16	20
LEMBOI	-26	88 *	24	4
MERABOI	54 *	10	10	17
MERRBOI	52 *	17	-2	20
SABBOI	-2	39 *	14	9
TOTBOI	20	21	9	14
TRTOBOI	29	35 *	3	5
RAOBOI	-3	62 *	17	17
ELROBOI	2	28	48 *	22
TREL8OI	27	30	13	5
INDBOI	-23	85 *	22	6
ACTBOI	-25	88 *	16	11
BRAAGOU	12	44 *	5	15
BRABGOU	22	19	16	14
FRASAGOU	16	17	51 *	5
FRASBGOU	-3	38 *	49 *	14
LEMGOU	4	9	66 *	11
MERRGOU	21	19	16	23
SABGOU	19	19	9	15
TOTGOU	20	-17	92 *	11
TRTOGOU	21	-19	91 *	11
RAOGOU	12	1	52 *	7
ELROGOU	14	-14	91 *	10
TRELGOU	16	-19	86 *	10
INDGOU	-15	43 *	50 *	17
ACTGOU	-9	34 *	54 *	16
FRASASHA	5	7	24	64 *
FRASBSHA	10	4	27	67 *
LEMSHA	33 *	12	-3	95 *
MERASHA	11	17	7	14
MERRSHA	10	20	9	14
SABSHA	21	18	6	48 *
RAOSHA	20	2	14	71 *
INDSHA	42 *	10	-1	90 *
ACTSHA	42 *	8	-4	92 *

ANNEXE 12.B ANALYSE DE LA VALIDITE DE TRAITS
ANALYSE FACTORIELLE CONFIRMATOIRE
METHODE DES MOINDRES CARRES NON-PONDERES
(coefficient de Spearman, 52 indices)

ROTATION DANS L'ESPACE DES QUATRE PREMIERS FACTEURS (ces quatre facteurs représentent 51.6 % de la variance, voir Annexe 10, pour les pourcentages de variance expliqués par les facteurs).

CORRELATION ENTRE LES FACTEURS

	F1	F2	F3	F4
F1	100 *			
F2	-4	100 *		
F3	2	7	100 *	
F4	20	-1	9	100 *

FACTOR STRUCTURE (CORRELATIONS)

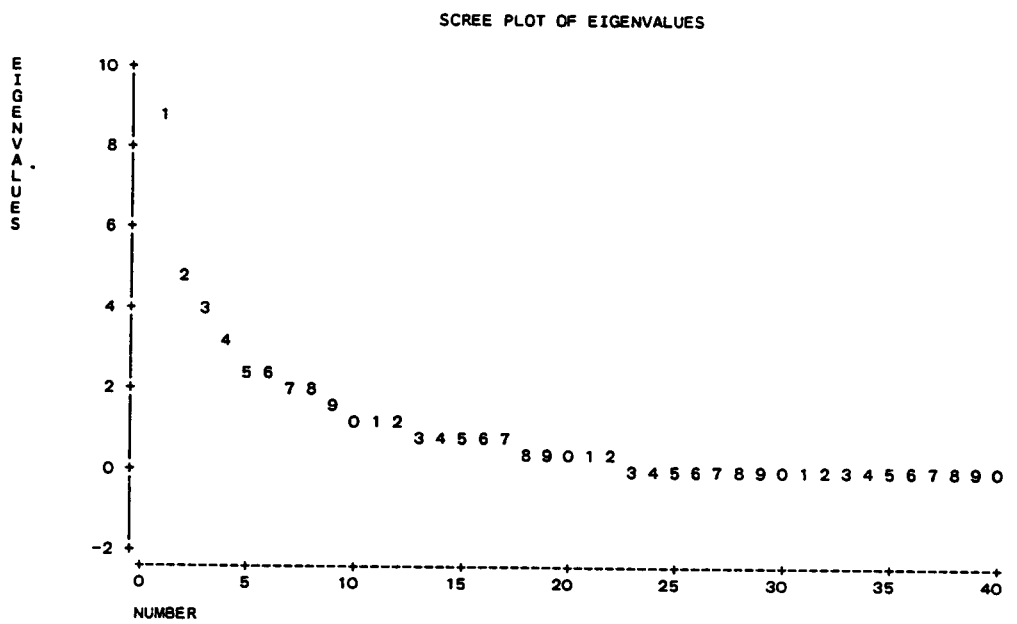
	FACTOR1	FACTOR2	FACTOR3	FACTOR4
BRAAHUI	10	45 *	7	21
FRASAHUI	29	9	14	4
FRASBHUI	32	16	17	4
LEMHUI	93 *	6	-1	5
MERAHUI	16	11	41 *	17
MERRHUI	18	17	48 *	17
SABHUI	24	3	37 *	15
TOTHUI	72 *	7	-7	29
TRTOHUI	77 *	9	11	18
RAOHUI	50 *	10	25	9
ELROHUI	77 *	6	-4	29
TRELHUI	83 *	7	-4	11
INDHUI	90 *	11	0	8
ACTHUI	90 *	10	1	9
BRAABOI	23	13	7	10
BRABBOI	13	13	2	10
FRASABOI	20	57 *	10	5
FRASBBOI	14	62 *	12	0
LEMBOI	-3	91 *	27	-7
MERABOI	39 *	11	16	34
MERRBOI	56 *	9	2	38 *
SABBOI	-2	87 *	15	-3
TOTBOI	7	22	11	27
TRTOBOI	6	23	18	19
RAOBOI	-1	53 *	28	14
ELROBOI	3	38 *	59 *	2
TRELBOI	6	26	12	7
INDBOI	3	90 *	21	-8
ACTBOI	0	95 *	18	-7
BRAAGOU	8	47 *	6	14
BRABGOU	14	30	14	16
FRASAGOU	9	20	68 *	7
FRASBGOU	4	35	65 *	18
LEMGOU	7	10	89 *	14
MERRGOU	6	28	16	3
SABGOU	13	15	17	21
TOTGOU	27	-3	86 *	8
TRTOGOU	25	-5	88 *	10
RAOGOU	8	0	86 *	12
ELROGOU	20	0	87 *	5
TRELGOU	17	1	89 *	5
INDGOU	1	25	81 *	16
ACTGOU	1	24	82 *	16
FRASASHA	15	9	17	62 *
FRASBSHA	20	8	19	66 *
LEMSHA	14	7	17	92 *
MERASHA	9	10	13	20
MERRSHA	9	15	14	15
SABSHA	24	10	1	86 *
RAOSHA	13	12	27	60 *
INDSHA	20	6	7	96 *
ACTSHA	20	7	10	98

ANNEXE 12.C ANALYSE DE LA VALIDITE DE TRAITS
ANALYSE FACTORIELLE CONFIRMATOIRE
METHODE DES MOINDRES CARRES NON-PONDERES
(coefficient de Pearson, 40 indices)

POURCENTAGE DE VARIANCE EXPLIQUES PAR LES FACTEURS

FACTEUR	% VARIANCE EXPLIQUEE	% VARIANCE CUMULEE
1	21.7	21.7
2	12.3	34.0
3	10.2	44.2
4	7.9	52.2
5	6.2	58.4
6	6.1	64.4
7	5.0	69.4
8	4.5	73.9
9	3.7	77.6
10	3.3	80.9
11	2.9	83.8
12	2.6	86.4
13	2.2	88.6
14	2.0	90.5
..	...	...

REPRESENTATION GRAPHIQUE DU POUVOIR EXPLICATIF DES FACTEURS



PATTERN DES FACTEURS APRES ROTATION SOUS CONTRAINTES DANS L'ESPACE DES QUATRE PREMIERS FACTEURS QUI REPRESENTENT 52 % DE LA VARIANCE TOTALE

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FRASAHUI	31	9	10	0
FRASBHUI	37 *	15	12	0
LEMHUI	94 *	13	2	23
SABHUI	18	6	16	7
TOTHUI	48 *	13	5	28
TRTOHUI	38 *	11	9	21
RAOHUI	22	-4	29	20
ELROHUI	47 *	1	24	41 *
TRELHUI	36	7	30	21
INDHUI	92 *	13	-1	22
ACTHUI	90 *	14	-3	31
FRASABOI	2	54 *	13	18
FRASBBOI	-2	55 *	16	20
LEMBOI	-9	92 *	21	4
SABBOI	2	39 *	12	9
TOTBOI	21	25	8	15
TRTOBOI	39 *	36	5	2
RAOBOI	12	66 *	16	16
ELROBOI	6	34	50 *	25
TRELBOI	36	33	13	2
INDBOI	-6	89 *	19	5
ACTBOI	-8	91 *	13	10
FRASAGOU	22	24	51 *	5
FRASBGOU	10	45 *	49 *	12
LEMGOU	13	19	68 *	11
SABGOU	27	19	10	12
TOTGOU	8	-4	94 *	19
TRTOGOU	9	-8	94 *	19
RAOGOU	15	7	55 *	8
ELROGOU	4	-1	93 *	19
TRELGOU	5	-7	89 *	18
INDGOU	-1	51 *	50 *	16
ACTGOU	3	42 *	55 *	16
FRASASHA	3	10	23	65 *
FRASBSHA	7	8	26	69 *
LEMSHA	39 *	13	-1	92 *
SABSHA	16	18	7	49 *
RAOSHA	21	6	14	71 *
INDSHA	46 *	11	3	88 *
ACTSHA	47 *	9	0	89 *

INTER-FACTOR CORRELATIONS

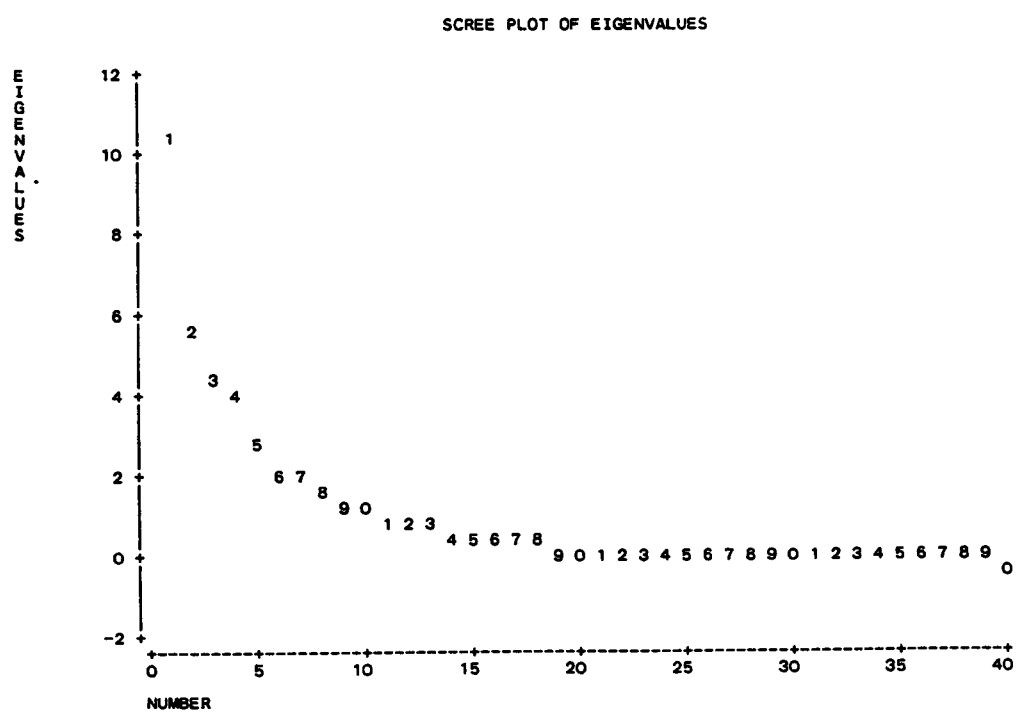
	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FACTOR1	100 *	15	10	16
FACTOR2	15	100 *	22	7
FACTOR3	10	22	100 *	13
FACTOR4	16	7	13	100 *

ANNEXE 12.D ANALYSE DE LA VALIDITE DE TRAITS
ANALYSE FACTORIELLE CONFIRMATOIRE
METHODE DES MOINDRES CARRES NON-PONDERES
(coefficient de Spearman, 40 indices)

POURCENTAGE DE VARIANCE EXPLIQUES PAR LES FACTEURS

FACTEUR	% VARIANCE EXPLIQUEE	% VARIANCE CUMULE
1	25.8	21.7
2	14.2	34.0
3	11.4	44.2
4	10.5	52.2
5	6.8	58.4
6	5.2	64.4
7	4.7	69.4
8	3.8	73.9
9	3.1	77.6
10	2.6	80.9
11	2.4	83.8
12	2.1	86.4
13	1.8	88.6
14	1.3	90.5
..	...	...

REPRESENTATION GRAPHIQUE DU POUVOIR EXPLICATIF DES FACTEURS



PATTERN DES FACTEURS APRES ROTATION SOUS CONTRAINTES DANS L'ESPACE DES QUATRE PREMIERS FACTEURS QUI REPRESENTENT 62 % DE LA VARIANCE TOTALE

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FRASAHUI	29	10	16	4
FRASBHUI	34	15	19	4
LEMHUI	95 *	13	5	6
SABHUI	23	6	35	14
TOTHUI	73 *	12	-2	29
TRTOHUI	76 *	17	15	17
RAOHUI	54 *	15	26	11
ELROHUI	79 *	11	2	30
TRELHUI	85 *	13	2	12
INDHUI	91 *	16	7	9
ACTHUI	91 *	15	8	10
FRASABOI	26	59 *	11	8
FRASBBOI	22	63 *	13	4
LEMBOI	10	94 *	25	0
SABBOI	9	87 *	13	1
TOTBOI	8	23	11	28
TRTOBOI	7	25	19	20
RAOBOI	7	58 *	30	19
ELROBOI	7	49 *	60 *	5
TRELBOI	8	25	11	8
INDBOI	17	93 *	20	-2
ACTBOI	14	95 *	16	-1
FRASAGOU	11	31	69 *	8
FRASBGOU	8	45 *	67 *	21
LEMGOU	9	26	91 *	17
SABGOU	12	15	17	20
TOTGOU	27	12	87 *	10
TRTOGOU	25	11	89 *	11
RAOGOU	8	13	87 *	13
ELROGOU	19	14	86 *	6
TRELGOU	17	15	89 *	7
INDGOU	5	41 *	83 *	19
ACTGOU	5	40	84 *	19
FRASASHA	16	8	19	63 *
FRASBSHA	21	8	20	68 *
LEMSHA	11	10	19	94 *
SABSHA	20	14	4	86 *
RAOSHA	15	15	27	63 *
INDSHA	16	9	10	96 *
ACTSHA	17	11	13	99 *

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FACTOR1	100 *	5	-1	28
FACTOR2	5	100 *	7	11
FACTOR3	-1	7	100 *	17
FACTOR4	28	11	17	100 *

ANNEXE 14 ANALYSE DE LA VALIDITE DE TRAITS
UTLILISATION DU MODELE D'EQUATION SIMULTANES LISREL
ESTIMATION PAR LE MAXIMUM DE VRAISEMBLANCE

ANNEXE 14.A MODELE1: LIBRE ORTHOGONAL
STANDARDIZED SOLUTION

MATRICE LAMBDA X (corrélations entre les indices et les facteurs)

	KSI 1	KSI 2	KSI 3
FRASBBOI	.537	-.010	.103
LEMBOI	.975	.112	.011
RAOBOI	.631	.179	.130
ELROBOI	.506	.488	-.053
FRASBGOU	.307	.658	.252
LEMGOU	.153	.973	.159
RAOGOU	.042	.964	.097
ELROGOU	.163	.660	.091
FRASBSHA	.078	.068	.638
LEMSHA	.042	.077	.810
RAOSHA	.067	.205	.271

MATRICE PHI (corrélation entre les facteurs)

KSI 1	KSI 2	KSI 3
1.000	1.000	1.000

ANNEXE 14.B MODELE2: CONTRAINT ORTHOGONAL
STANDARDIZED SOLUTION

MATRICE LAMBDA X (corrélations entre les indices et les facteurs)

	KSI 1	KSI 2	KSI 3
FRASBBOI	.518	.000	.000
LEMBOI	1.004	.000	.000
RAOBOI	.637	.000	.000
ELROBOI	.548	.000	.000
FRASBGOU	.000	.727	.000
LEMGOU	.000	1.029	.000
RAOGOU	.000	.932	.000
ELROGOU	.000	.639	.000
FRASBSHA	.000	.000	.876
LEMSHA	.000	.000	.616
RAOSHA	.000	.000	.308

MATRICE LAMBDA X (corrélations entre les indices et les facteurs)

KSI 1	KSI 2	KSI 3
1.000	1.000	1.000

ANNEXE 14.C MODELE3: LIBRE OBLIQUE
STANDARDIZED SOLUTION

MATRICE LAMBDA X (corrélations entre les indices et les facteurs)

	KSI 1	KSI 2	KSI 3
FRASBBOI	.538	-.011	.104
LEMBOI	.974	.114	.011
RAOBOI	.630	.176	.128
ELROBOI	.500	.491	-.061
FRASBGOU	.301	.652	.241
LEMGOU	.143	.969	.142
RAOGOU	.032	.963	.079
ELROGOU	.156	.658	.079
FRASBSHA	.082	.051	.638
LEMSHA	.046	.055	.810
RAOSHA	.066	.198	.268

MATRICE LAMBDA X (corrélations entre les indices et les facteurs)

	KSI 1	KSI 2	KSI 3
KSI 1	1.000		
KSI 2	.009	1.000	
KSI 3	-.008	.045	1.000

ANNEXE 14.D MODELE3: CONTRAINT OBLIQUE
STANDARDIZED SOLUTION

MATRICE LAMBDA X (corrélations entre les indices et les facteurs)

	KSI 1	KSI 2	KSI 3
FRASBBOI	.538	.000	.000
LEMBOI	.937	.000	.000
RAOBOI	.681	.000	.000
ELROBOI	.587	.000	.000
FRASBGOU	.000	.719	.000
LEMGOU	.000	1.030	.000
RAOGOU	.000	.933	.000
ELROGOU	.000	.633	.000
FRASBSHA	.000	.000	.789
LEMSHA	.000	.000	.678
RAOSHA	.000	.000	.335

MATRICE PHI

	KSI 1	KSI 2	KSI 3
KSI 1	1.000		
KSI 2	.335	1.000	
KSI 3	.147	.279	1.000

ANNEXE 15 A ANALYSE DE LA VALIDITE FACIALE DES INDICES

1) PROTOCOLE POUR L'EVALUATION DE LA VALIDITE FACIALE DES INDICES (document remis à l'avis des experts de la société Nielsen).
"Classer, pour chacun des marchés suivants (Huile, Boisson1 et Boisson2, Shampooings), les cartes du marchés proposées, par ordre de crédibilité ou d'intérêt décroissant (en vous basant sur ce qu'une telle carte peut vous apporter en matière de connaissance de la structure concurrentielle du marché)".
"Affecter une note de crédibilité-intérêt (entre 1 et 10) à chacune de ces cartes".

Listes des marques et de leurs abréviations sur chacun des marchés

MARCHE DE L'HUILE

	MARQUE	ABREVIATION
1	LESSIEUR TOURNESOL	L TOURNESOL
2	LESSIEUR ARACHIDE	L ARACHIDE
3	LESSIEUR MAIS	L MAIS
4	LESSIEUR OLIVE	L OLIVE
5	FRUIT D'OR	F D'OR
6	ECLAT D'OR	ECLAT D'OR
7	EPI D'OR	EPI d'OR
8	PUJET	PUJET
9	AUTRES MARQUES	AUTRES
10	MARQUES DISTRIBUTEURS	DR

MARCHE BOISSON1

	MARQUE	ABREVIATION
1	ORANGINA STANDARD	ORANGINA
2	INDIAN TONIC STANDARD	I TONIC
3	CANADA DRY	C DRY
4	GINI	GINI
5	RICQLES	RICQLES
6	COCA COLA STANDARD	COCA STD
7	COCA COLA SANS CAFEINE	COCA SANS CAFE
8	PEPSI COLA	PEPSI
9	SEVEN UP	SEVEN UP
10	LIMONADES VERRE	LIM VERRE
11	LIMONADES PET	LIM PET
12	PERRIER ZESTE	PERRIER Z
13	BRUT DE POMME	BRUT DE P
14	AUTRES BAF ET SODAS	AUTRES BAF

MARCHE BOISSON2

	MARQUE	ABREVIATION
1	DRY ORANGE	DRY ORANGE
2	DRY CITRON	DRY CITRON
3	DRY EXOTIQUE	DRY EXO
4	DRY PAMPLEMOUSSE	DRY PPL
5	ORANGINA	ORANGINA
6	FANTA CITRON	F CITRON
7	FANTA ORANGE	F ORANGE
8	FANTA EXOTIQUE	F EX
9	FRUITE ORANGE	FRUITE ORANGE
10	FRUITE POMME & CASSIS	FRUITE P ET C
11	FRUITE ANANAS PAMPL	FRUITE AN PPL
12	DISTRIBUTEUR ORANGE	DR ORANGE
13	DISTRIBUTEUR CITRON	DR CITRON
14	AUTRES MARQUES ORANGE	AUTRES M ORANGE

MARCHE DES SHAMPOOINGS

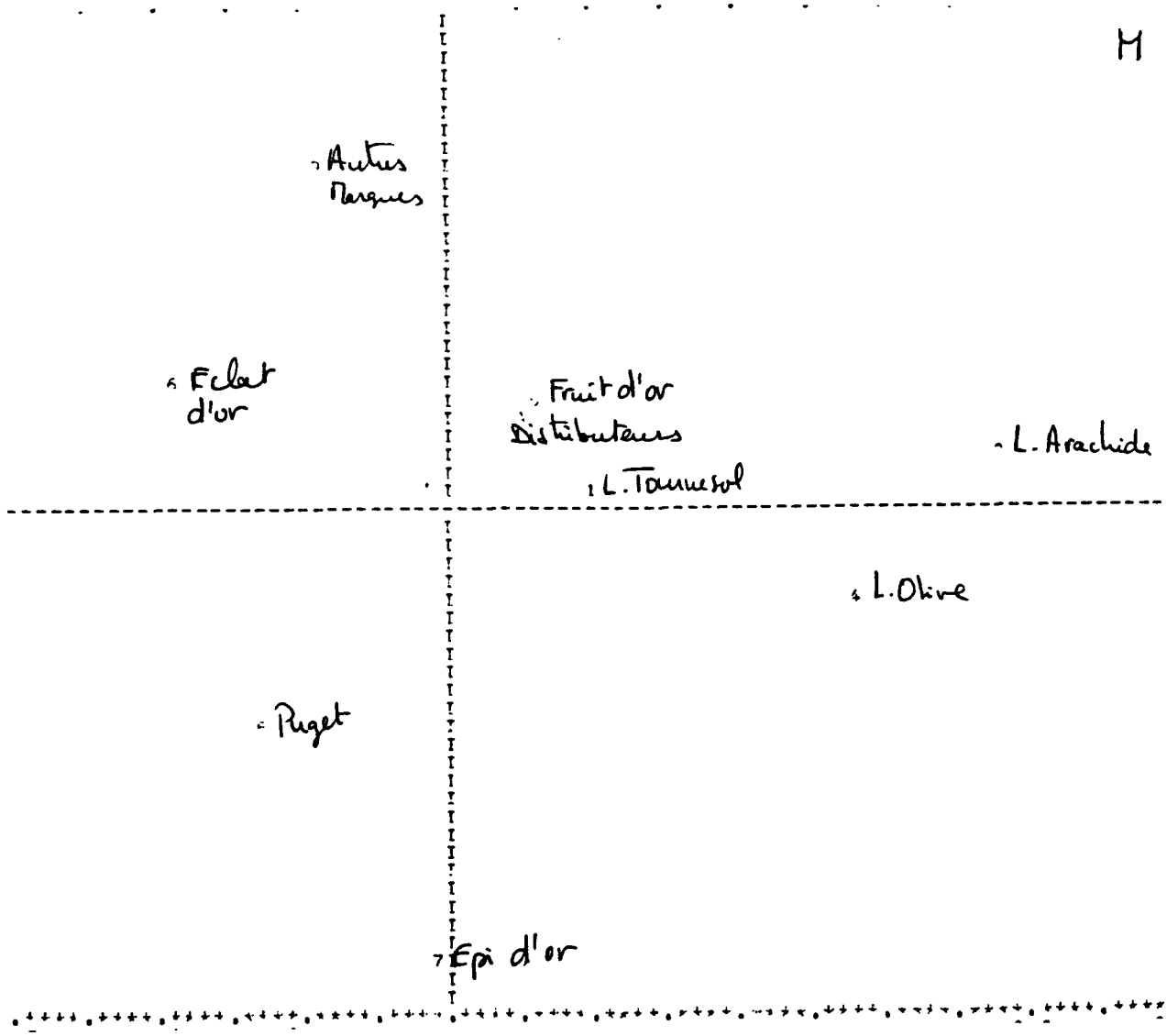
	MARQUE	ABREVIATION
1	MIXA	MIXA
2	BABYDOP	BABY D
3	BABIVEA	BIVEA
4	ULTRADOUX	ULTRA D
5	FLOREAL	FLOREAL
6	TIMOTEI	TIMOTEI
7	NIVEA	NIVEA
8	DIMENSION	DIME
9	ELSEVE	ELSEVE
10	LONGUEURS ET PONTES	L ET P
11	SOYANCE HAP	SOYANCE
12	ENERGANCE	ENERGANCE
13	PETROL HAHN	PETROL H
14	MOELLE ET GARNIER HAP	MOELLE ET G
15	DOP HAP	DOP
16	PALMOLIVE HAP	PALMOLIVE
17	EQUILIBRE	EQUILIBRE
18	HEAD & SHOULDERS	HEAD
19	CLEAR	CLEAR
20	TOTAL DISTRIBUTEURS	DISTRIBUTEURS
21	SOYANCE ANTIPEL	SOYANCE AP
22	MOELLE ET GARNIER ANTIP	MOELLE ET G AP
23	DOP ANTIPEL	DOP AP
24	PALMOLIVE ANTIPEL	PALMOLIVE AP

SIGNIFICATION DES INDICES UTILISES POUR CONSTRUIRE LES CARTES

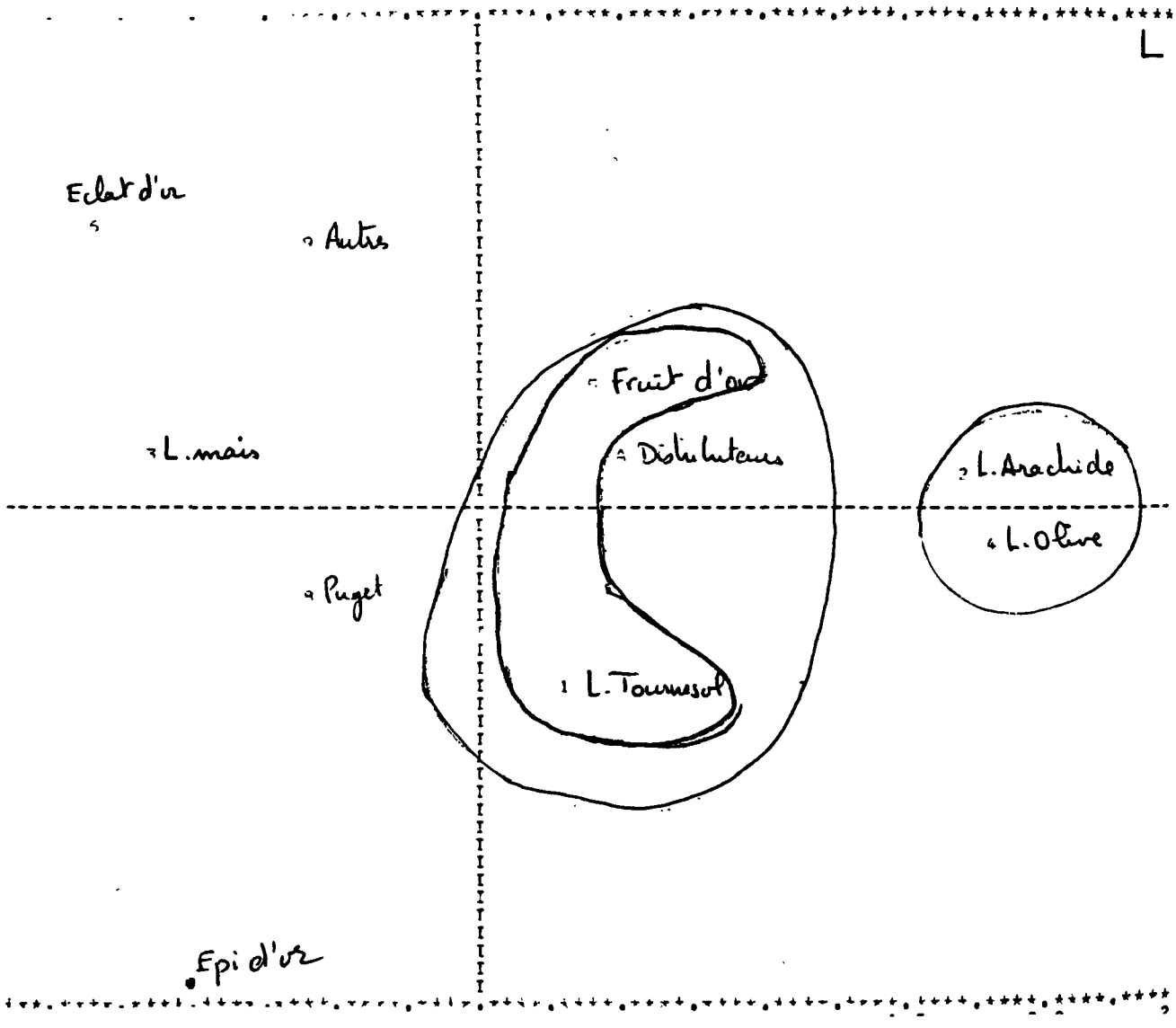
FRASA	Similarité construite à partir des intervalles inter-achats entre les marques
FRASB	idem
BRAA	idem
BRAB	idem
LEM	Similarité basée sur le nombre d'individus ayant effectué un transfert entre les deux marque, au cours de la période d'observation.
RO	idem
SABA	Similarité basée sur le nombre d'actes de transferts entre les deux marques au cours de la période.
TOTT	Similarité en terme de ce que les marques ont perdu au bénéfice des autres, pertes mesurées en nombre d'actes de transferts vers une autre marque.
ACTE1	idem
TRTO	Similarité en terme de ce que les marques ont gagné sur les autres, gain mesurés en nombre d'actes de transferts vers une autre marque.
ACTE2	idem
ELRO	Similarité en terme de ce que les marques ont perdu au bénéfice des autres, pertes mesurées en nombre d'individus effectuant un transfert vers une autre marque.
INDI1	idem
TREL	Similarité en terme de ce que les marques ont gagné sur les autres, gain mesurés en nombre d'individus ayant effectué un transfert vers une autre marque.
INDI2	idem
MERA	Similarité construite à partir des quantités relatives d'achat des marques
MERR	idem

ANNEXE 15.B CARTES REPRESENTATIVES DES MARCHES

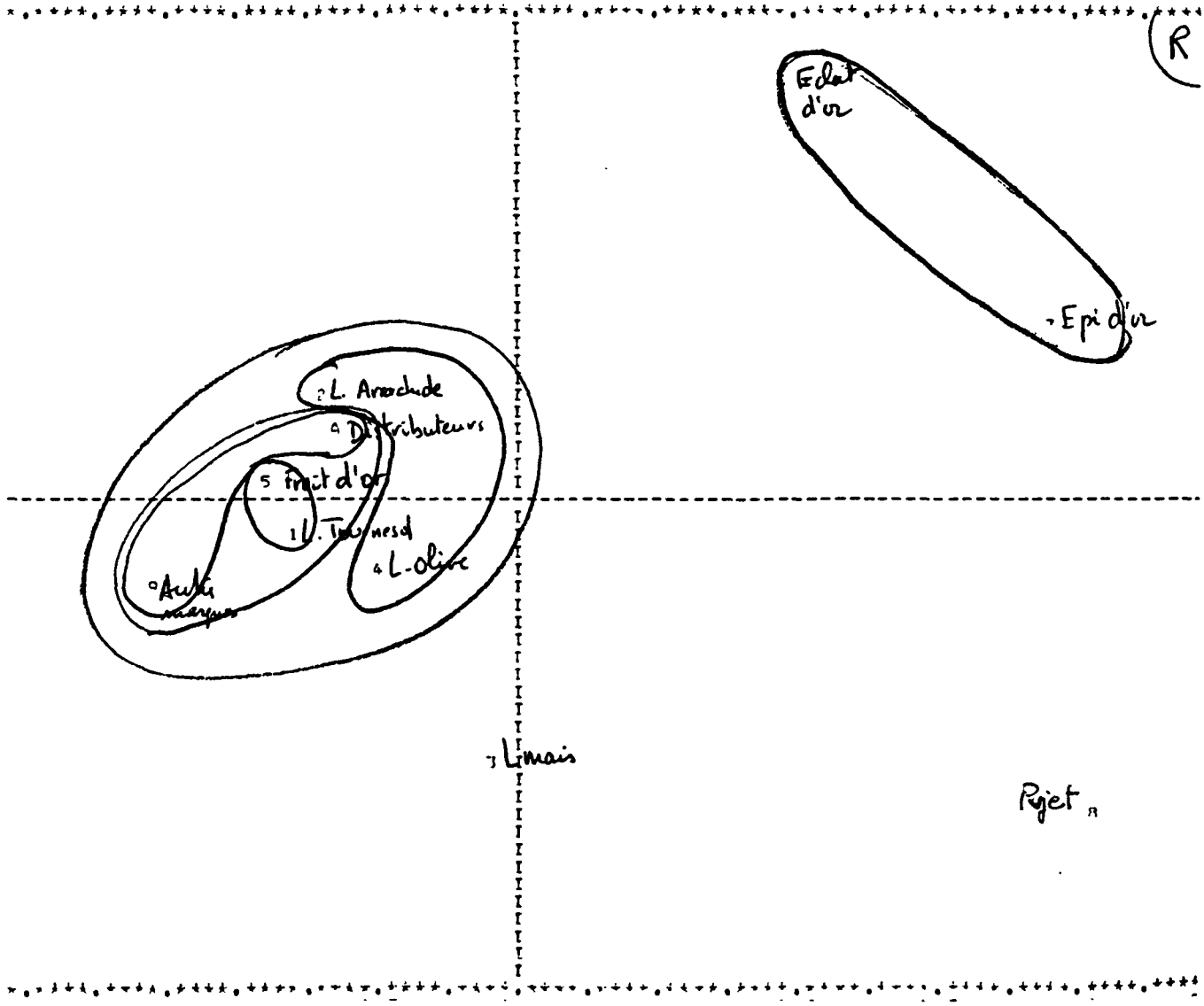
1) MARCHE HUILE INDICE FRASA



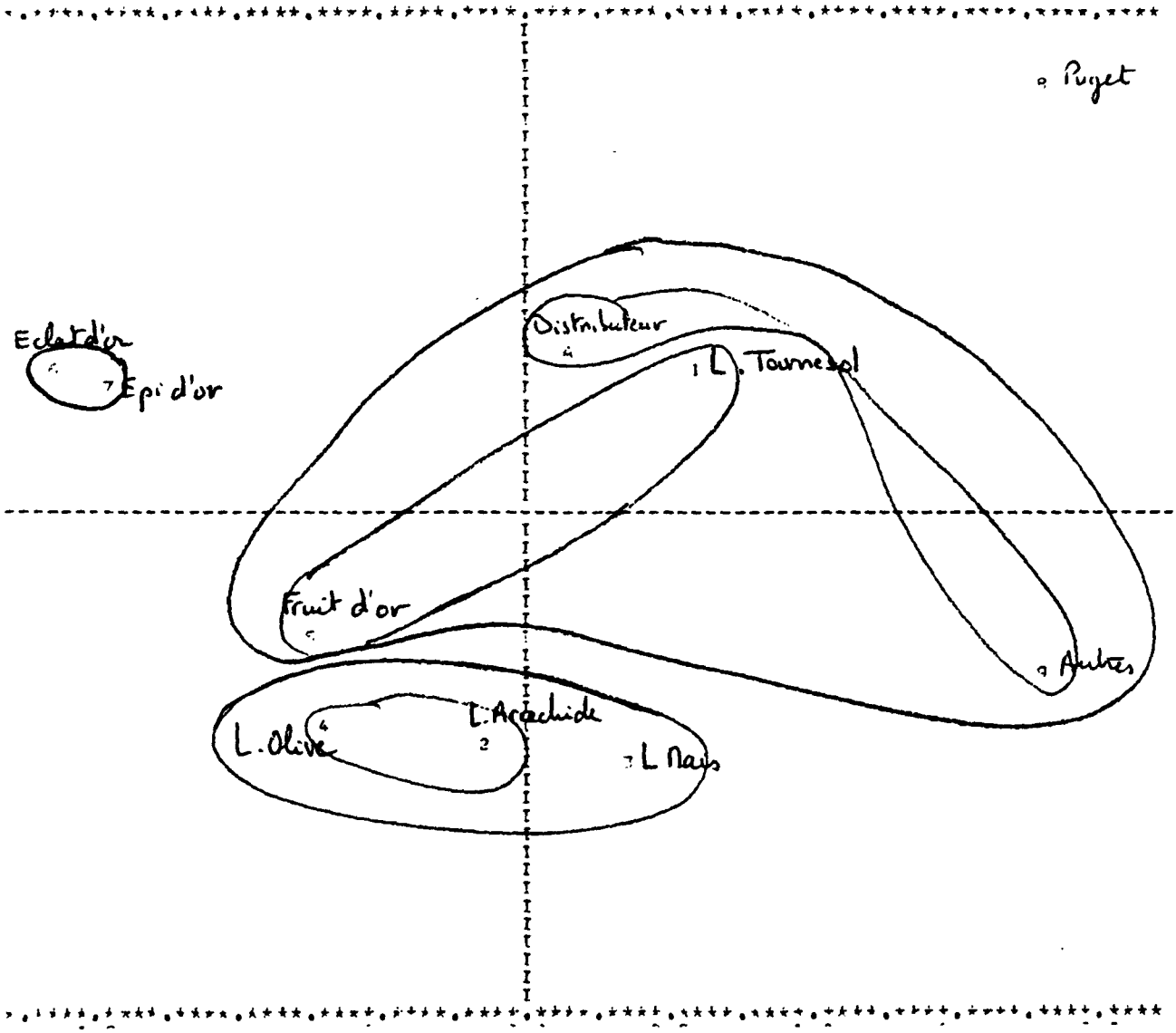
MARCHE HUILE INDICE FRASB



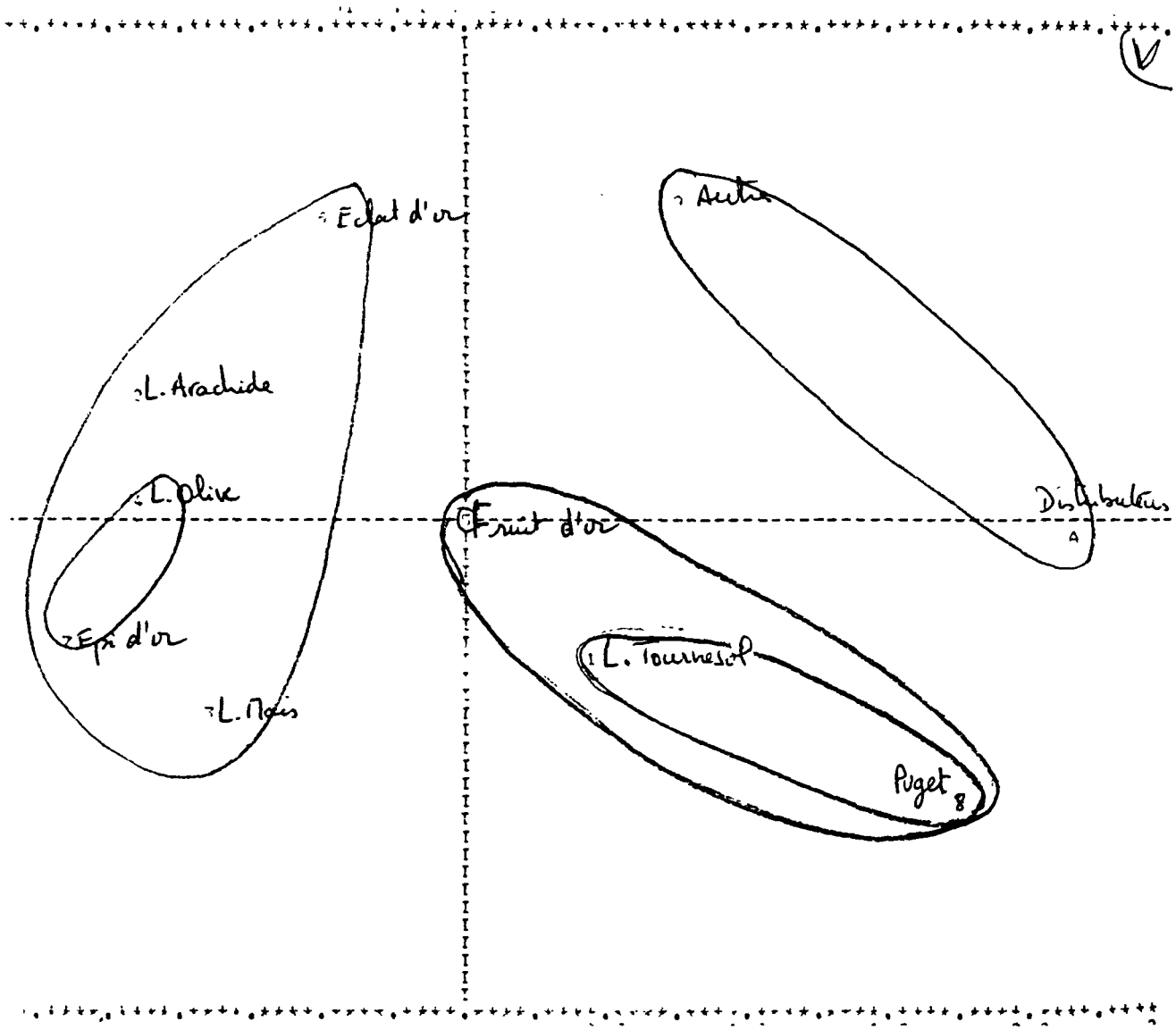
MARCHE HUILE INDICE LEM



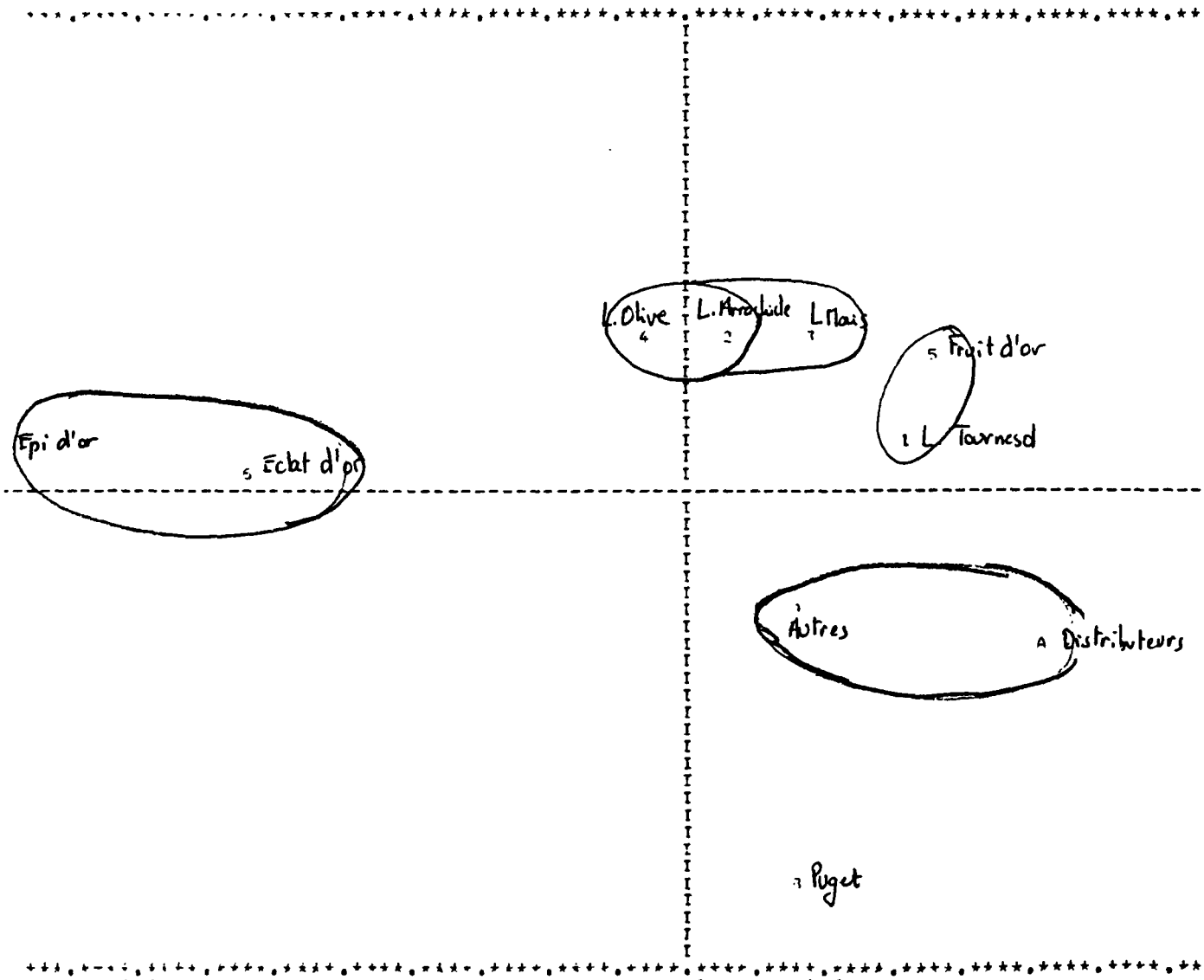
MARCHE HUILE INDICE RAO



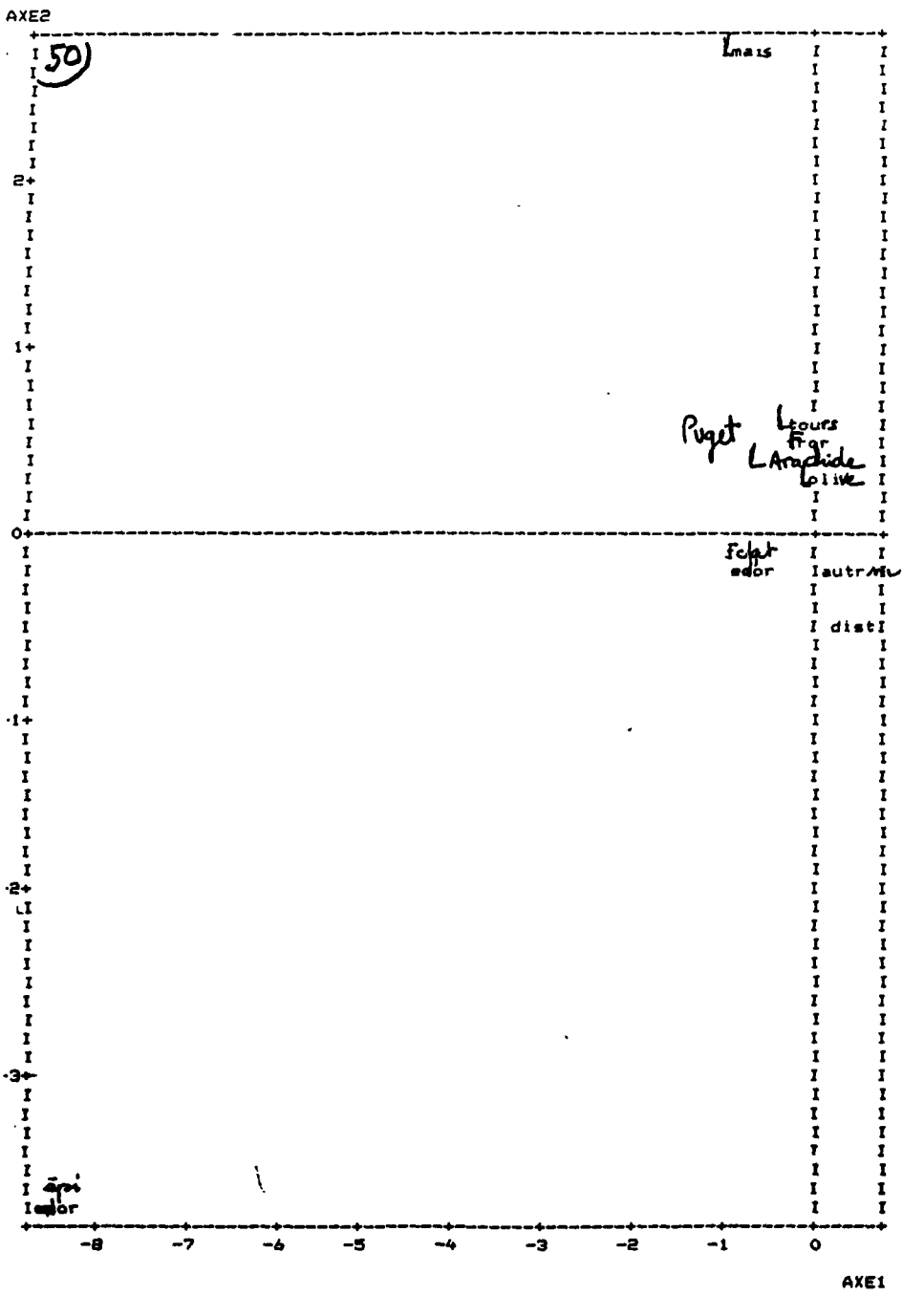
MARCHE HUILE INDICE SABA



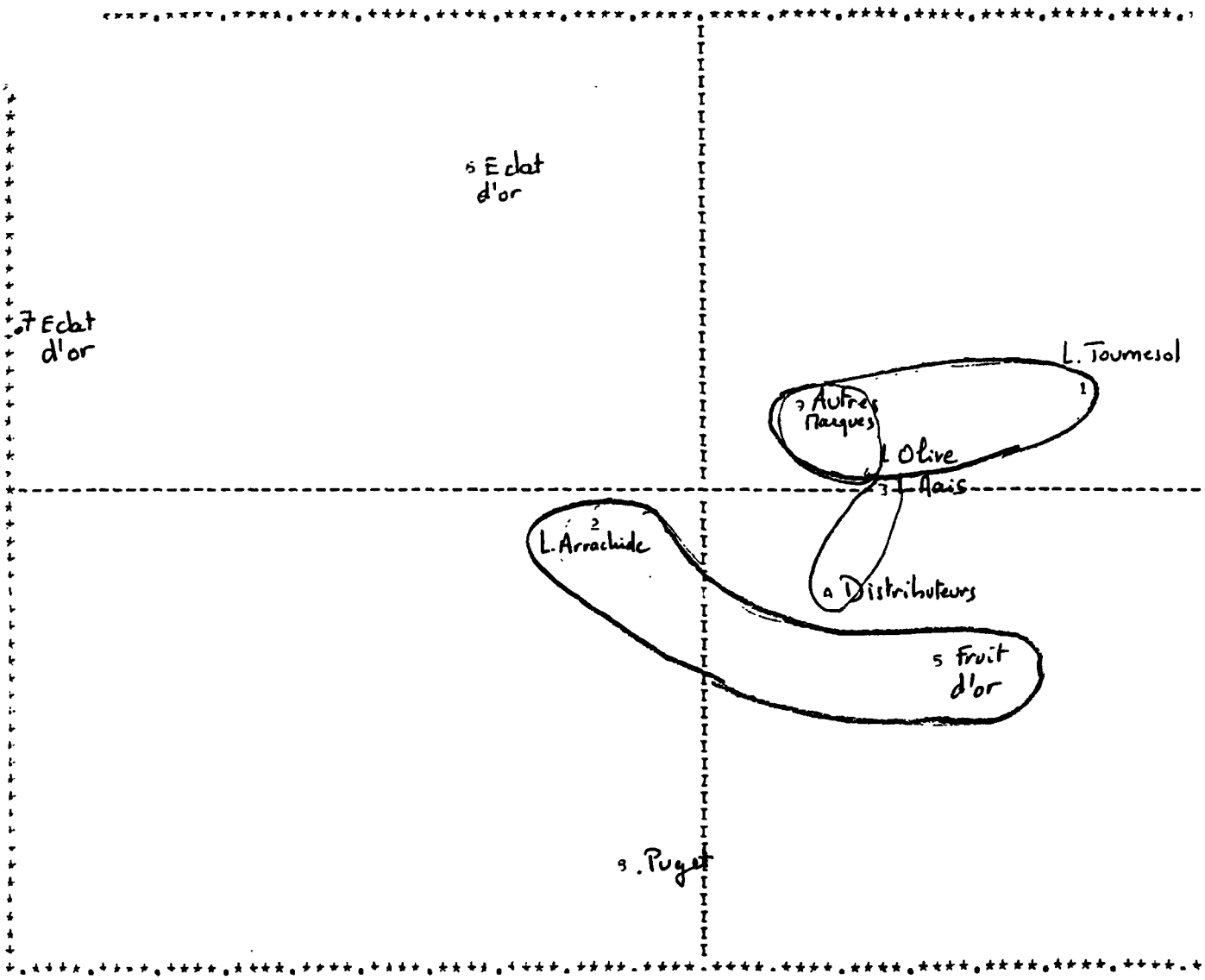
MARCHE HUILE INDICE TOTT



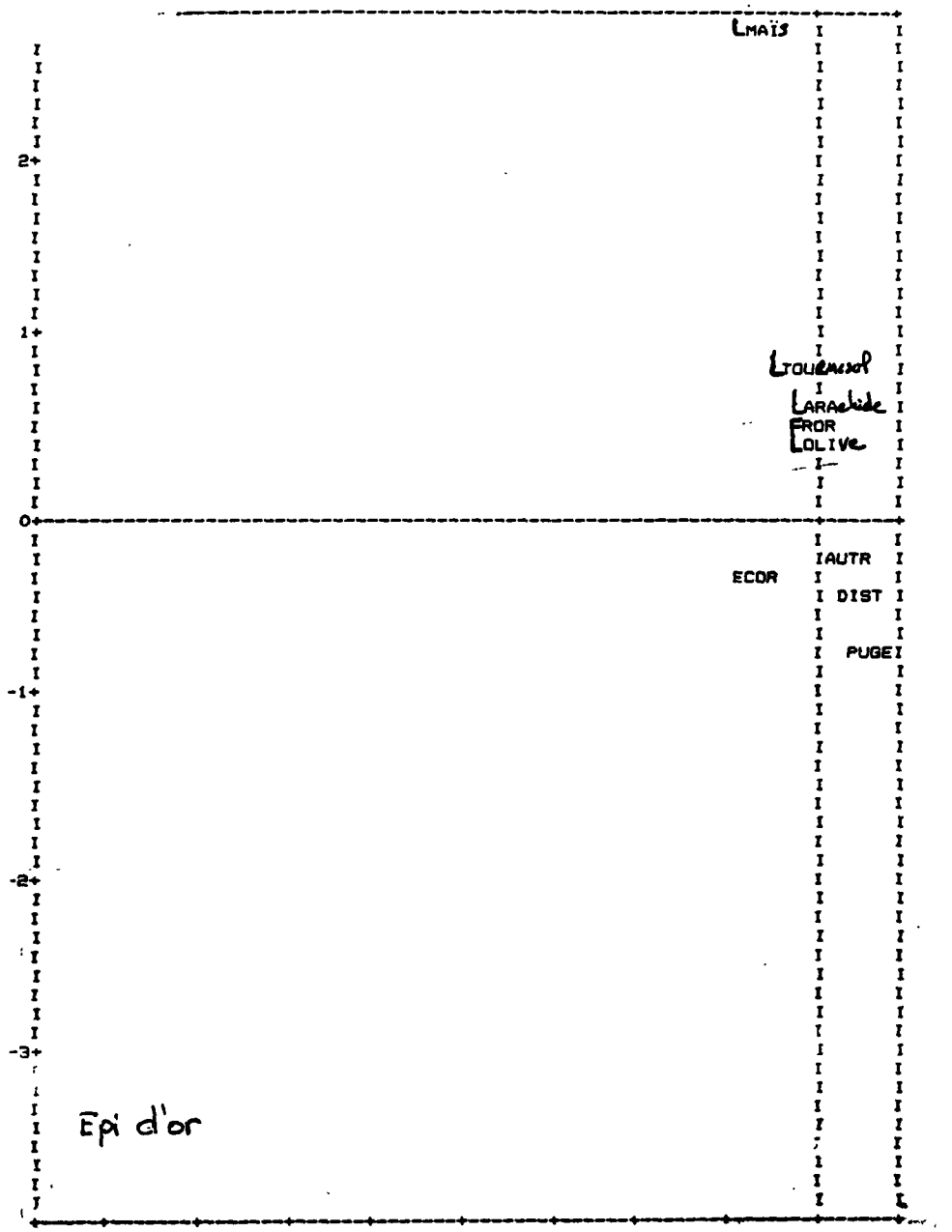
MARCHE HUILE INDICE ACTE1



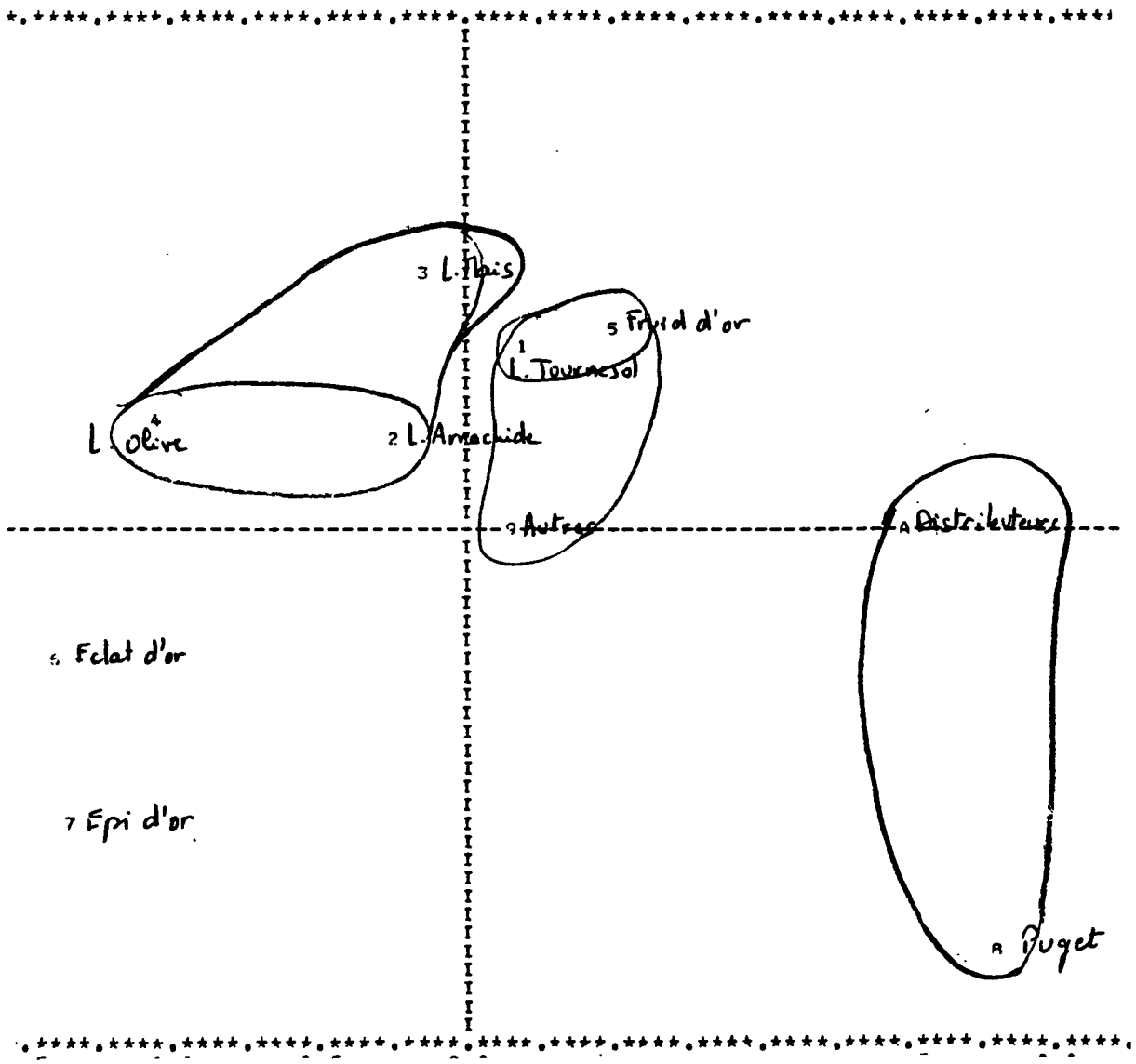
MARCHE HUILE INDICE TRTO



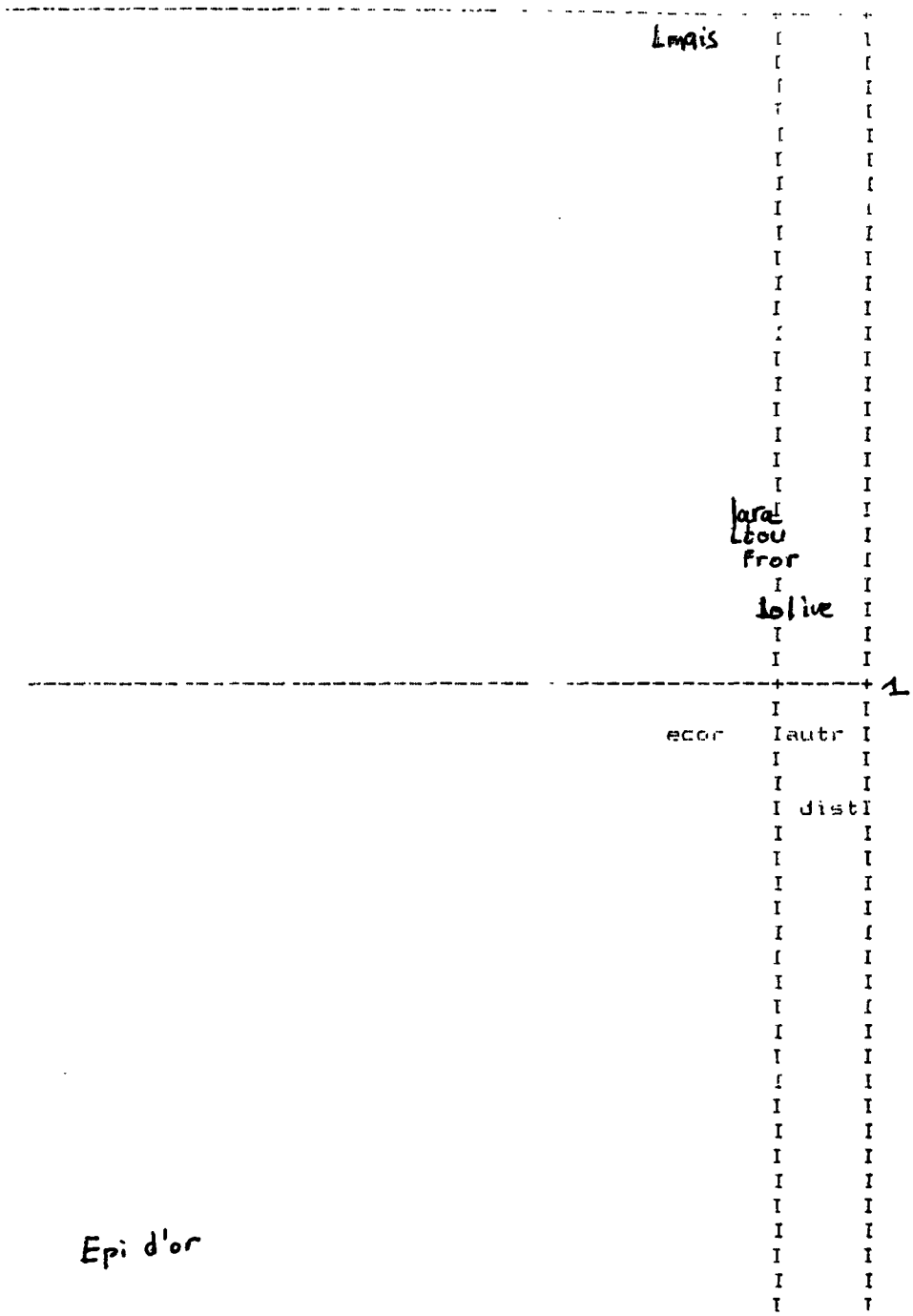
MARCHE HUILE INDICE ACTE2



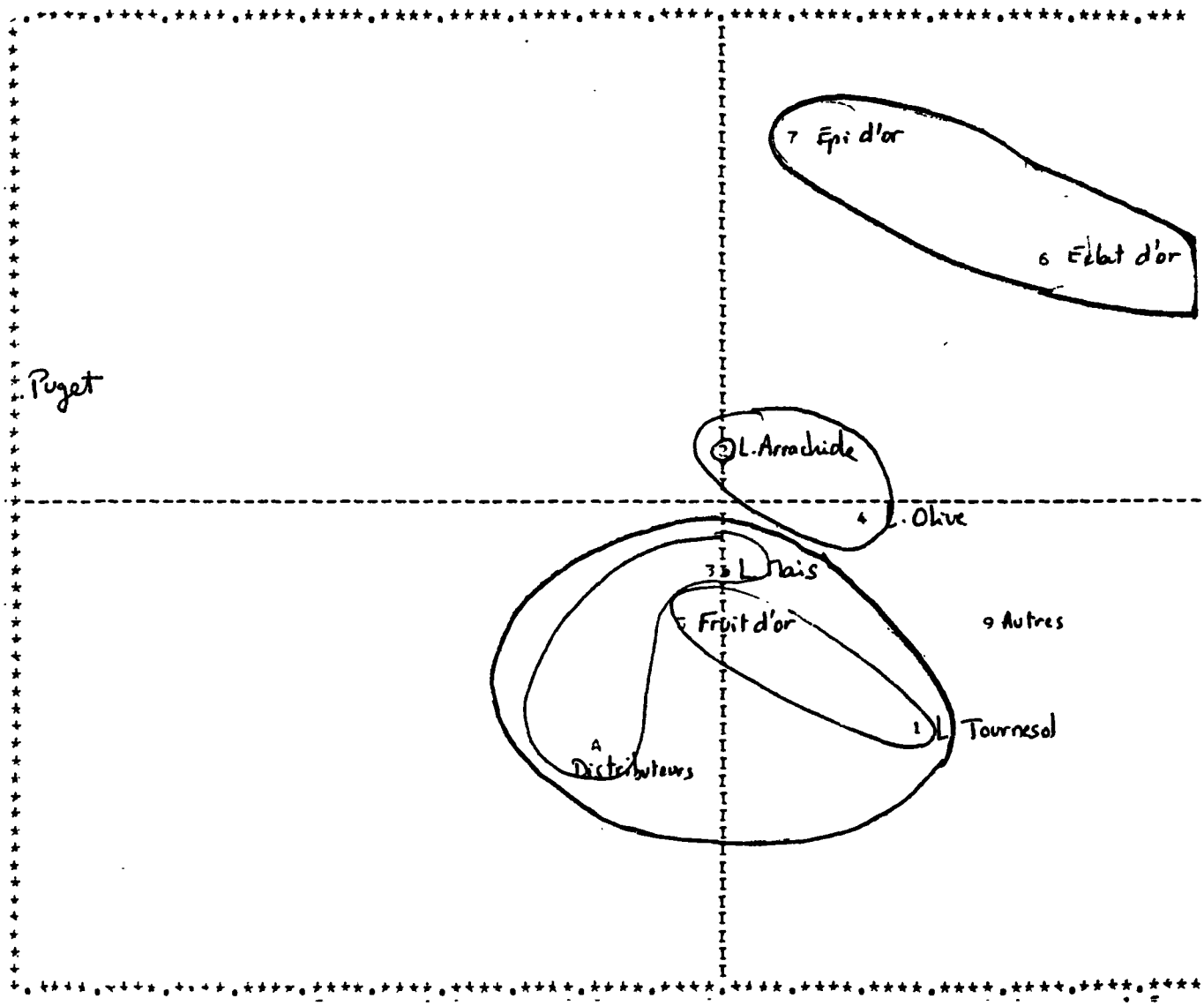
MARCHE HUILE INDICE ELRO



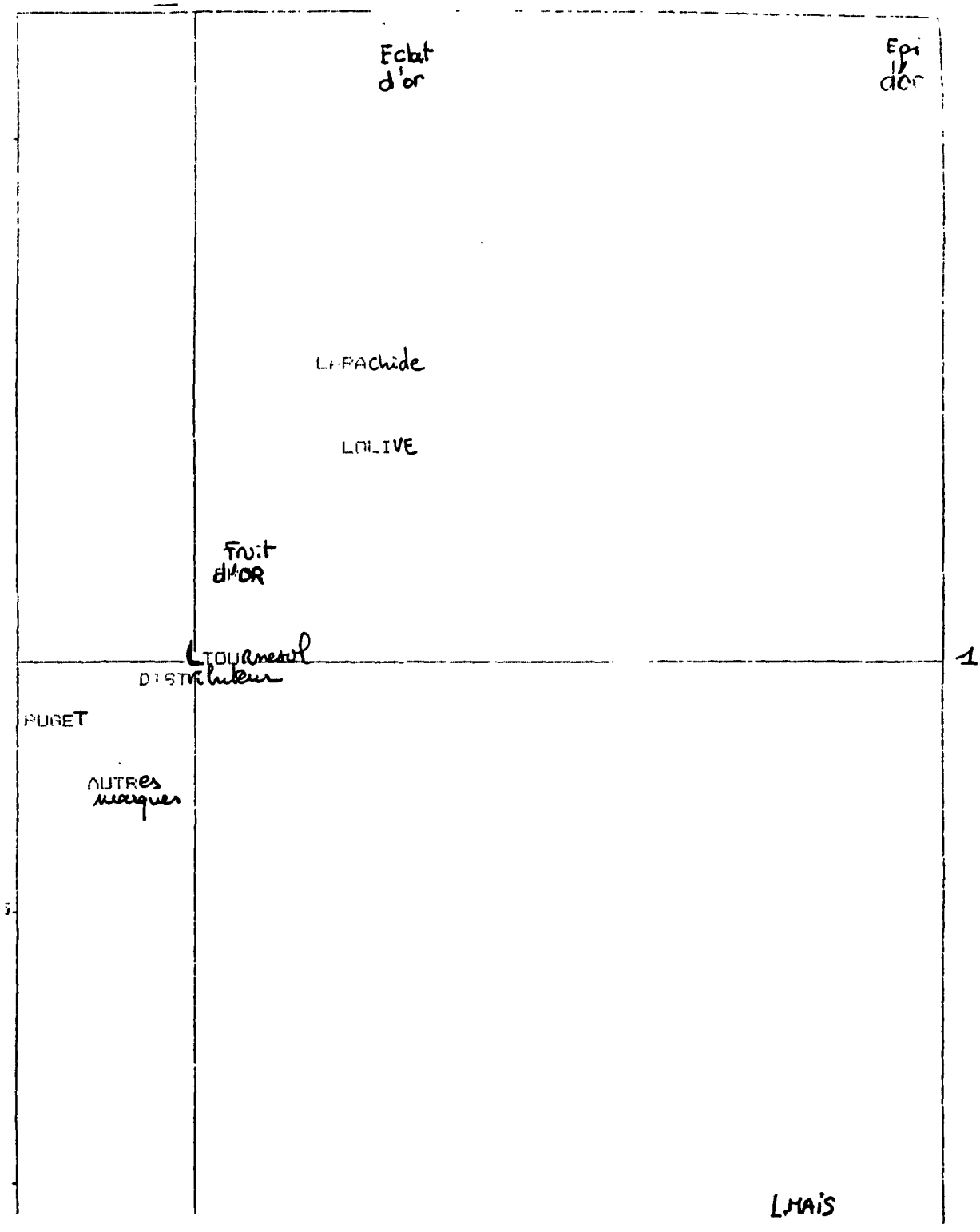
MARCHE HUILE INDICE INDI1



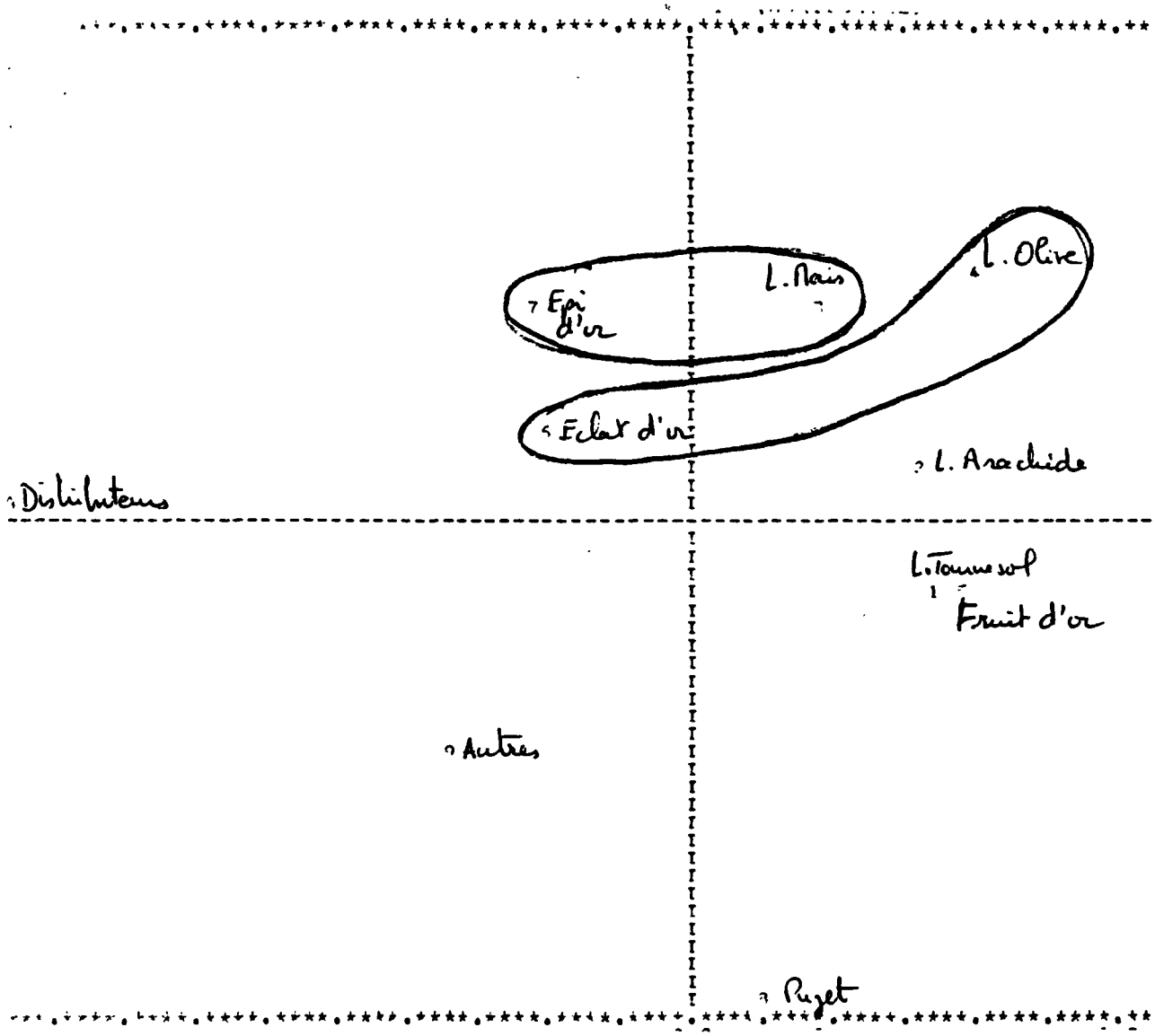
MARCHE HUILE INDICE TREL



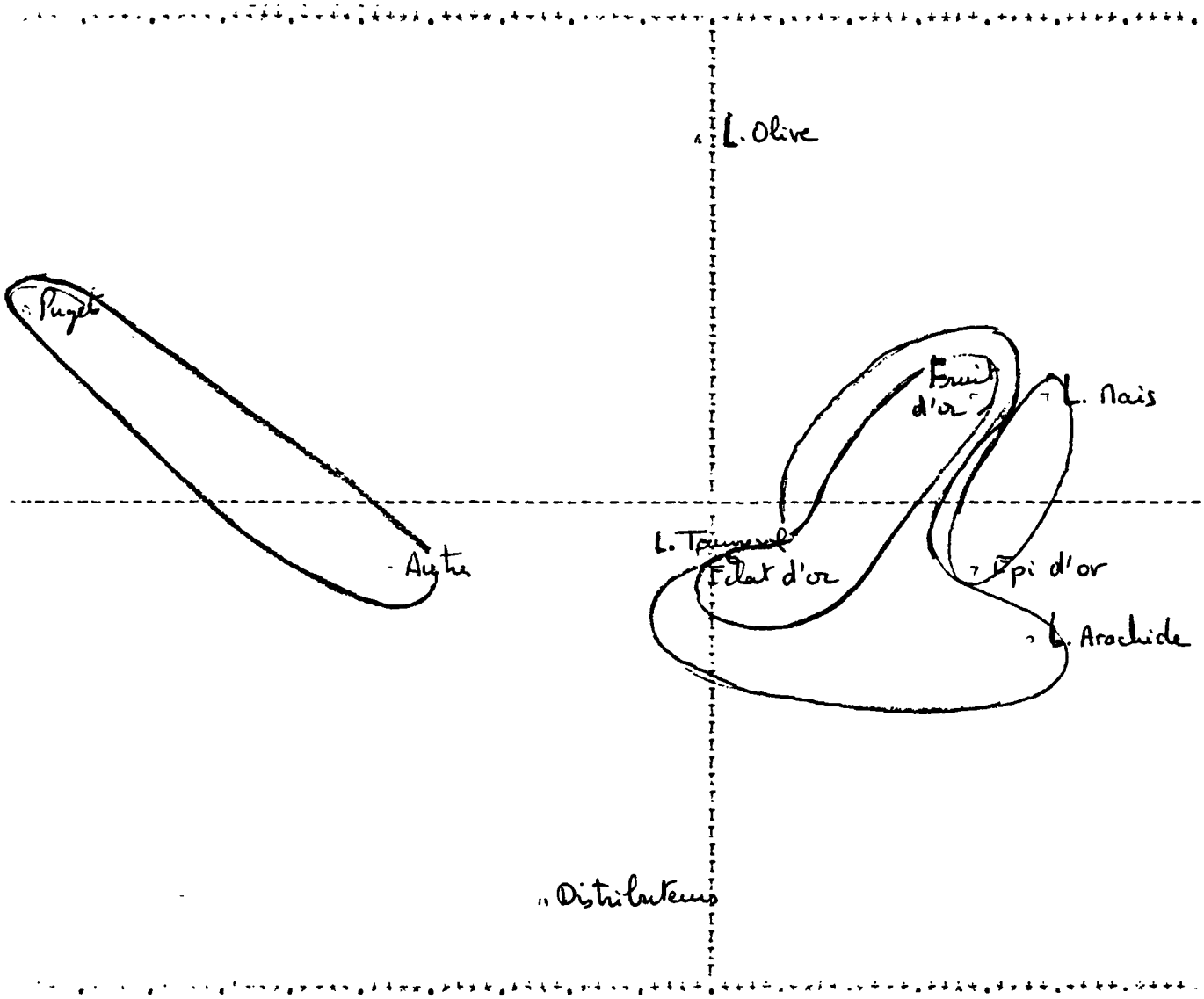
MARCHE HUILE INDICE INDI2



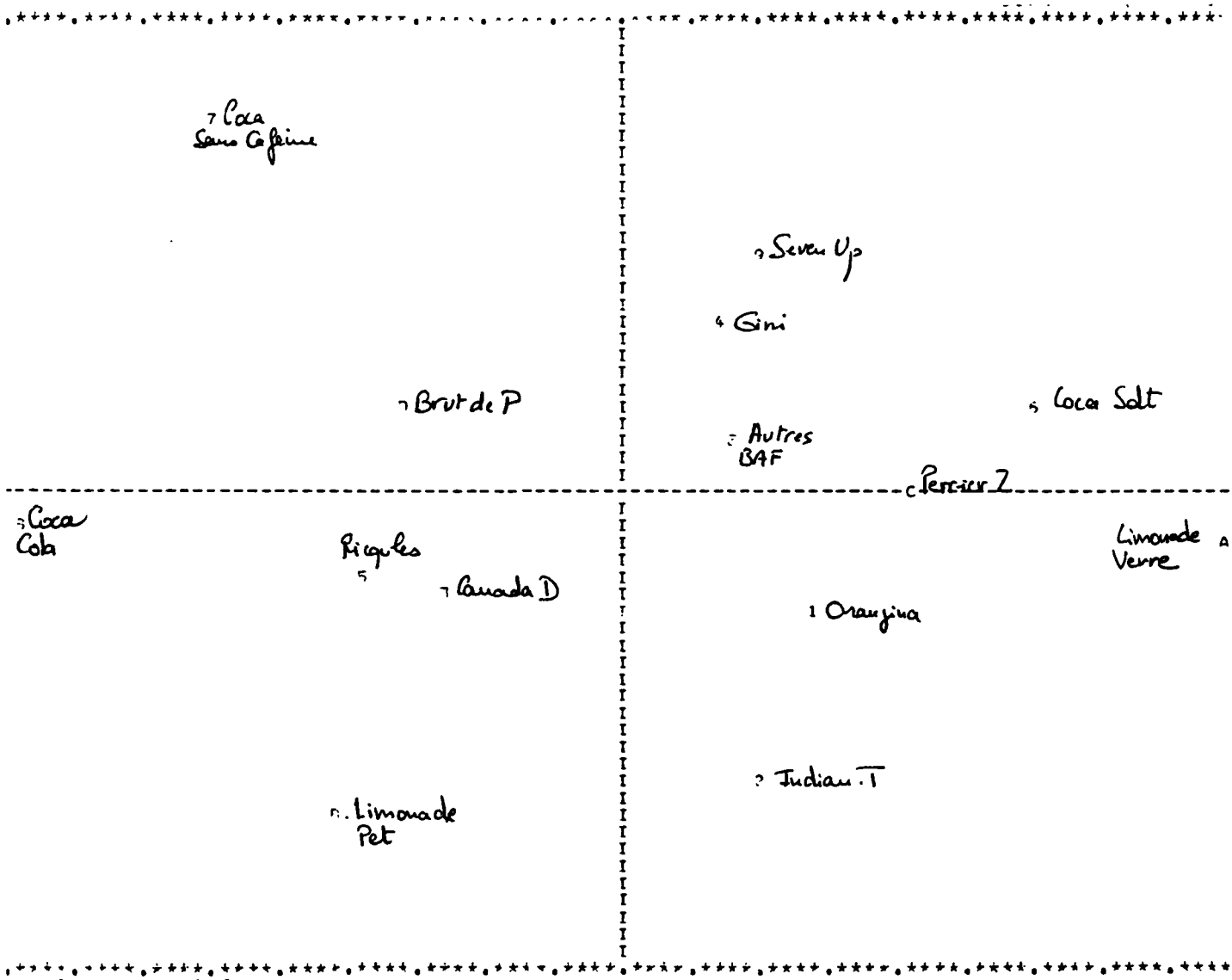
MARCHE HUILE INDICE MERA



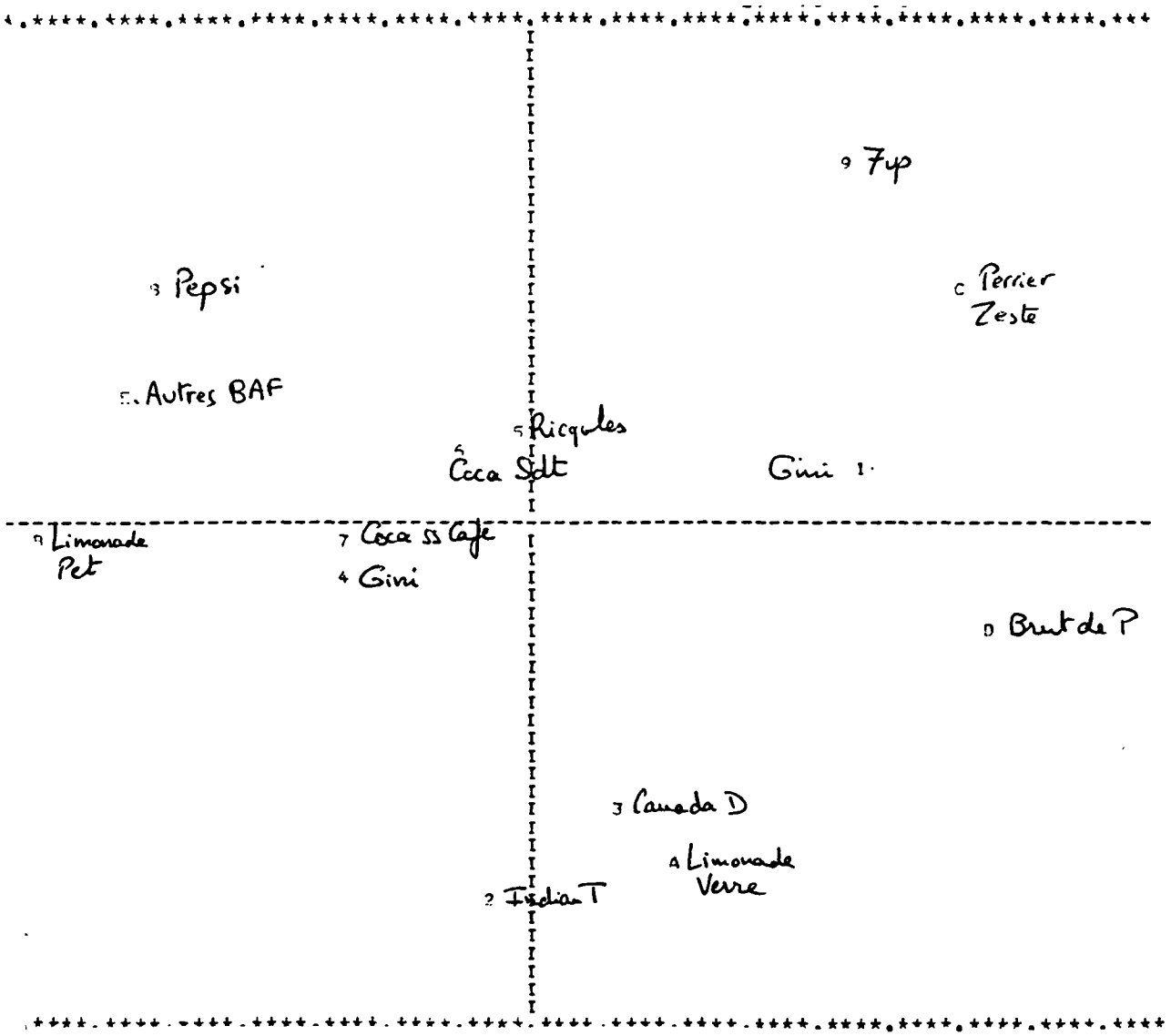
MARCHE HUILE INDICE MERR



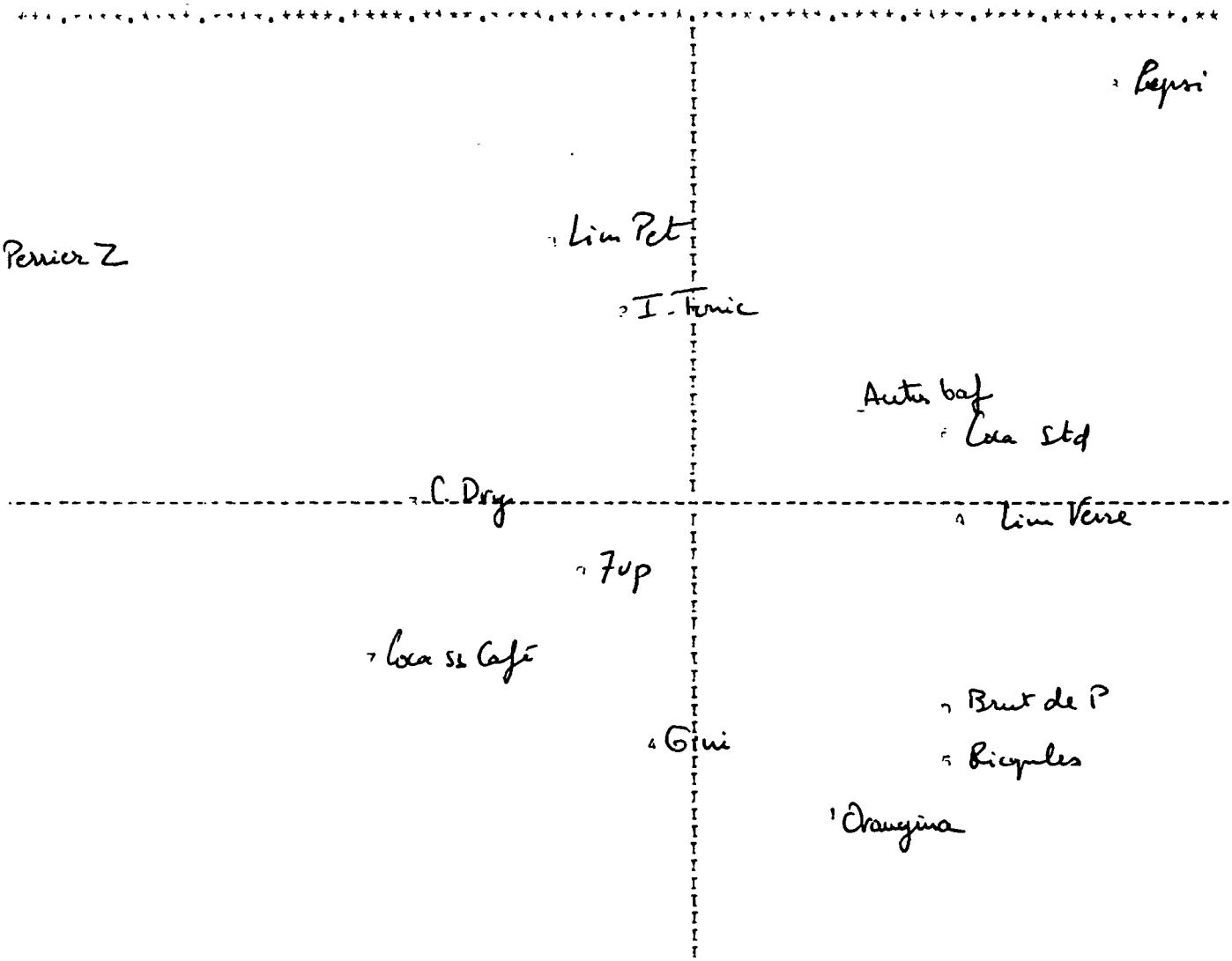
2) MARCHE BOISSON INDICE BRAA



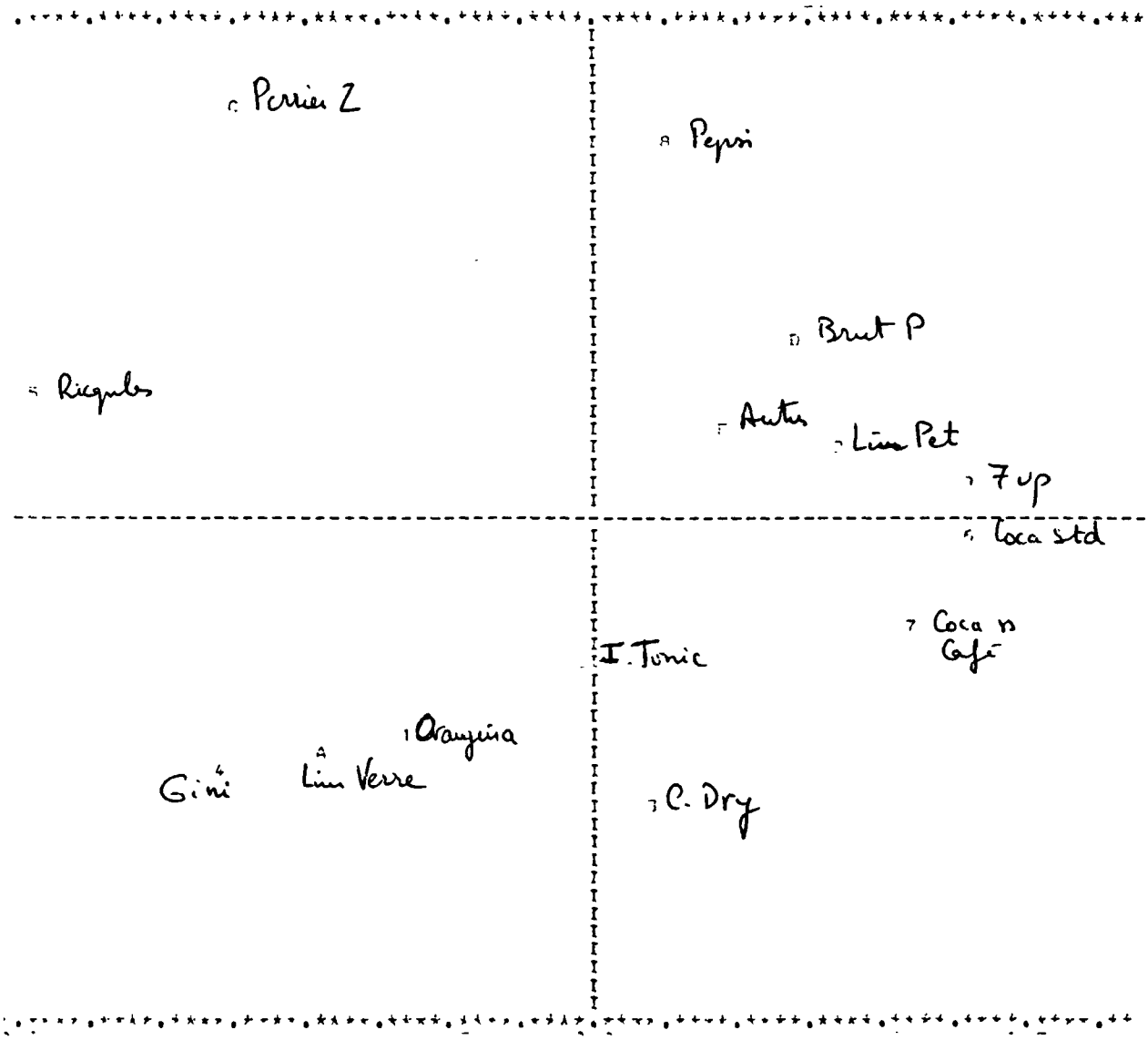
MARCHE BOISSON INDICE BRAB



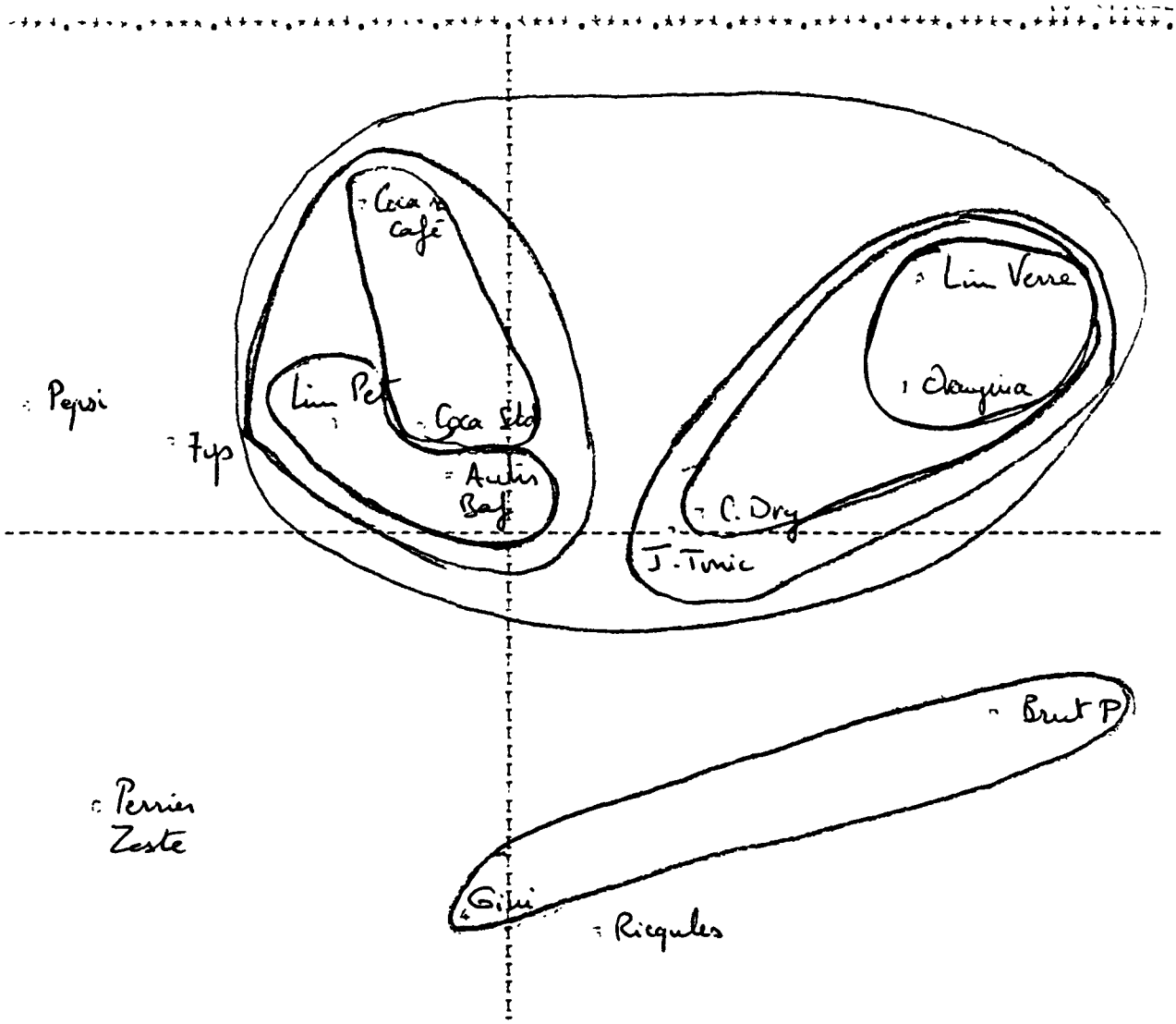
MARCHE BOISSON INDICE FRASA



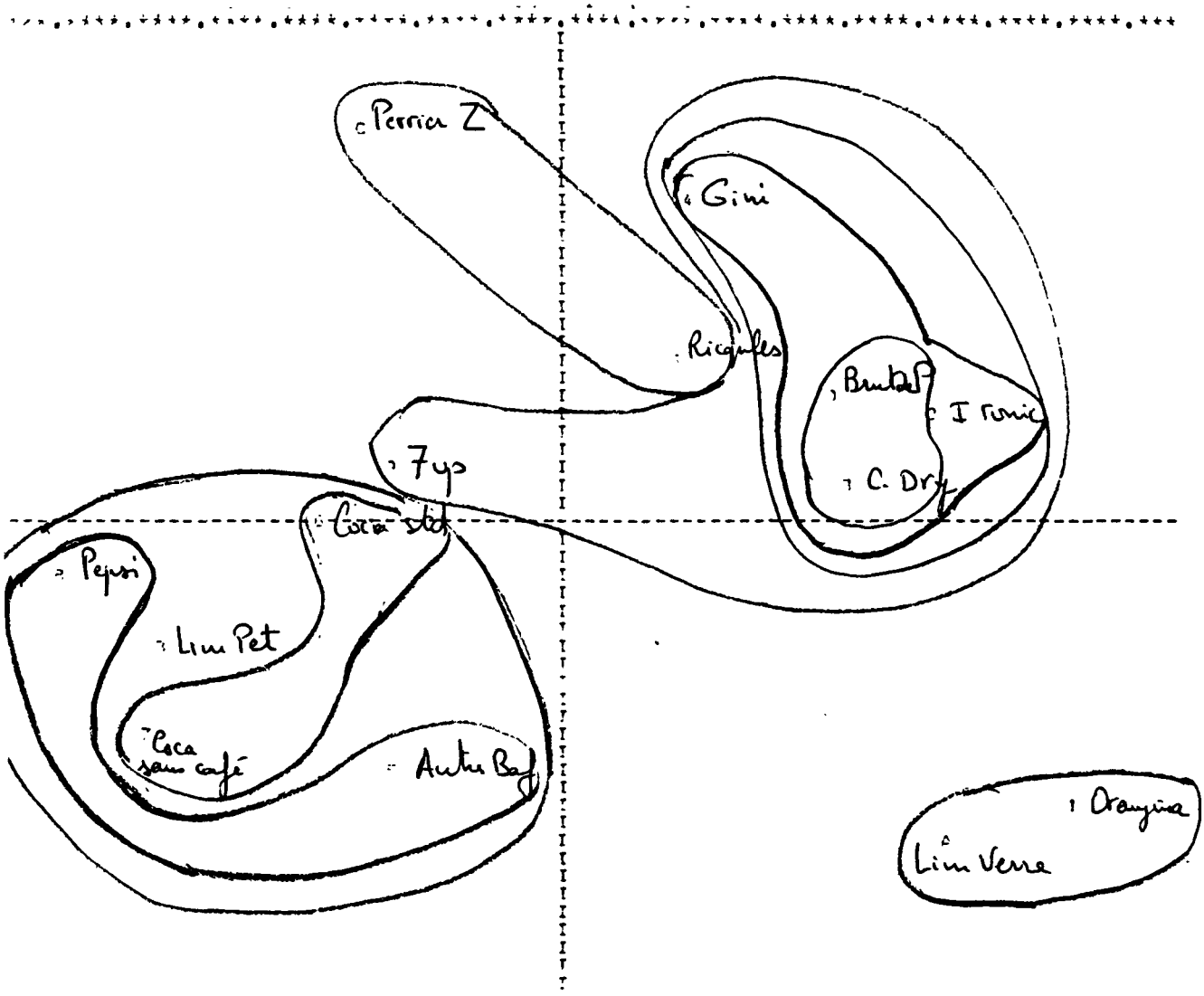
MARCHE BOISSON INDICE FRASB



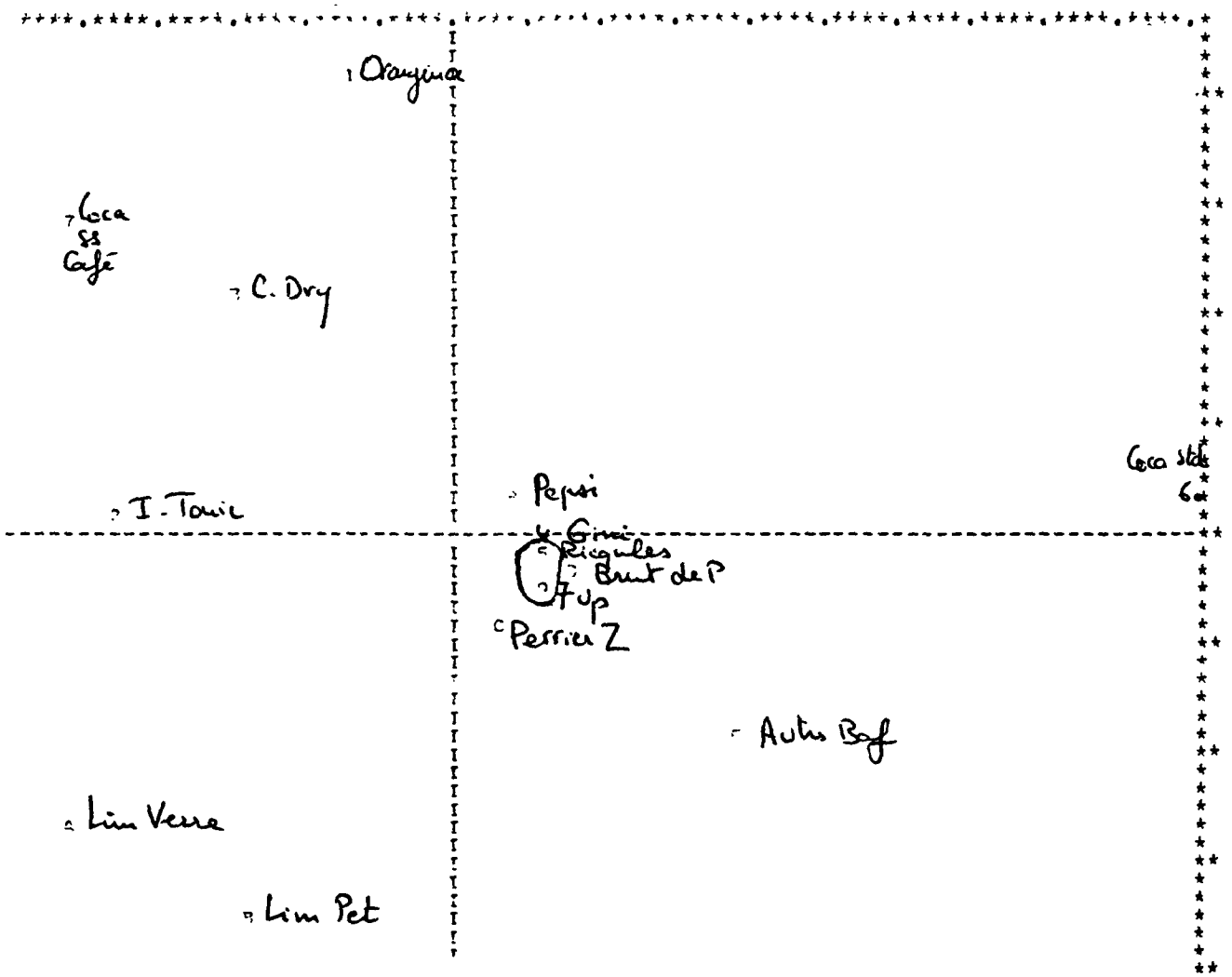
MARCHE BOISSON INDICE LEM



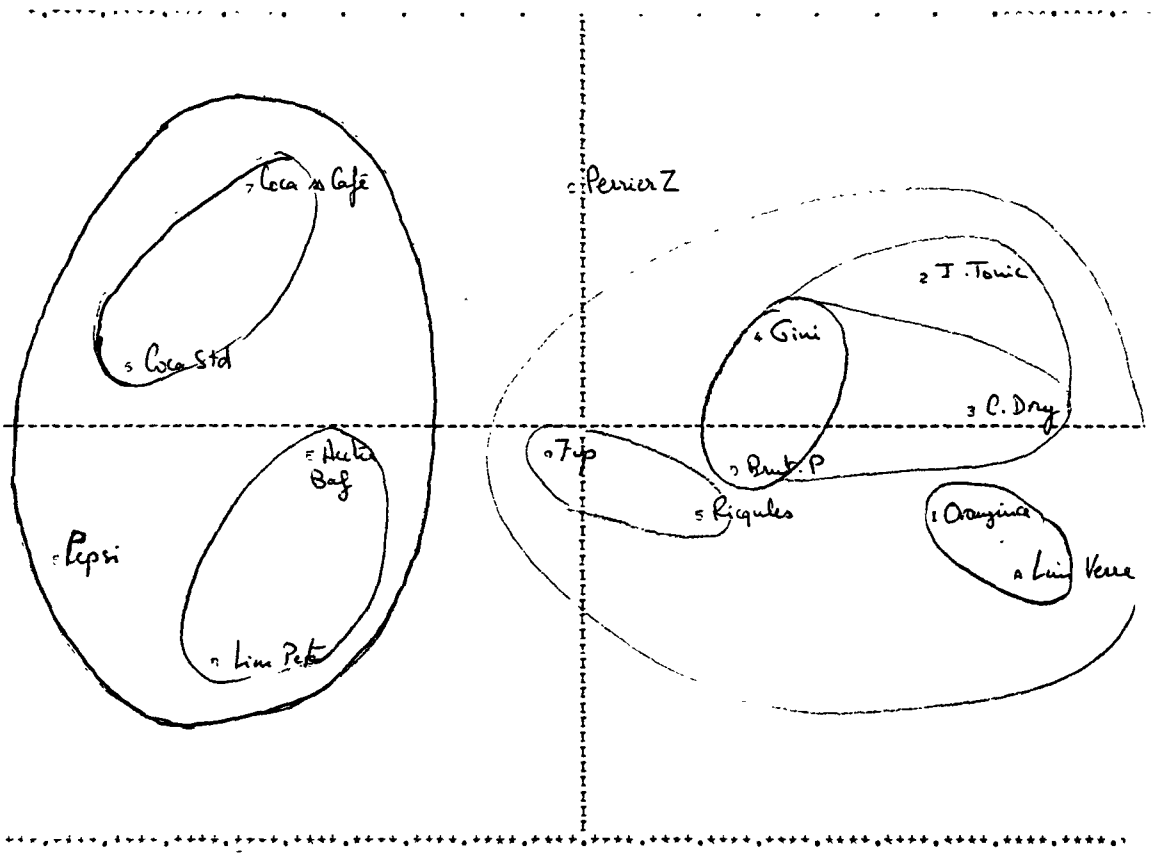
MARCHE BOISSON INDICE RAO



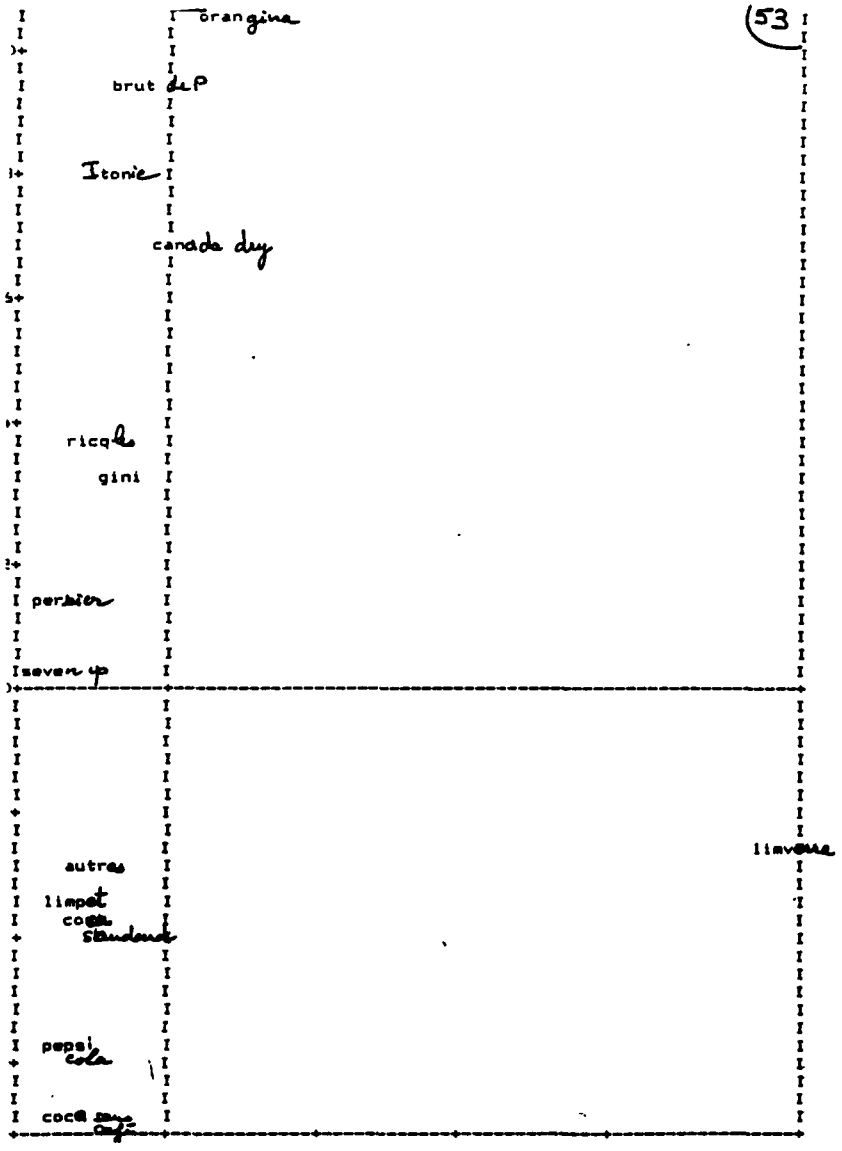
MARCHE BOISSON INDICE SABA



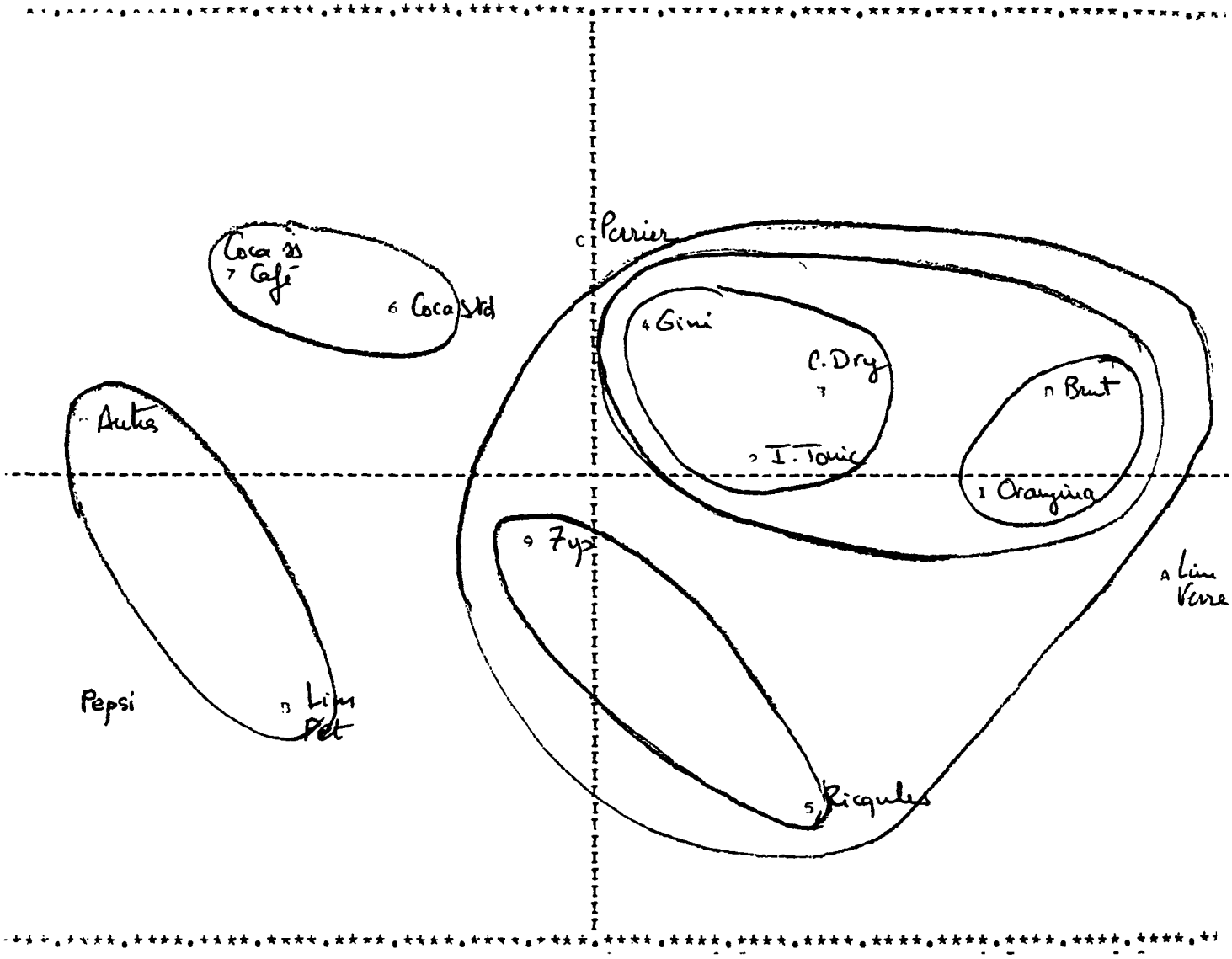
MARCHE BOISSON INDICE TOTT



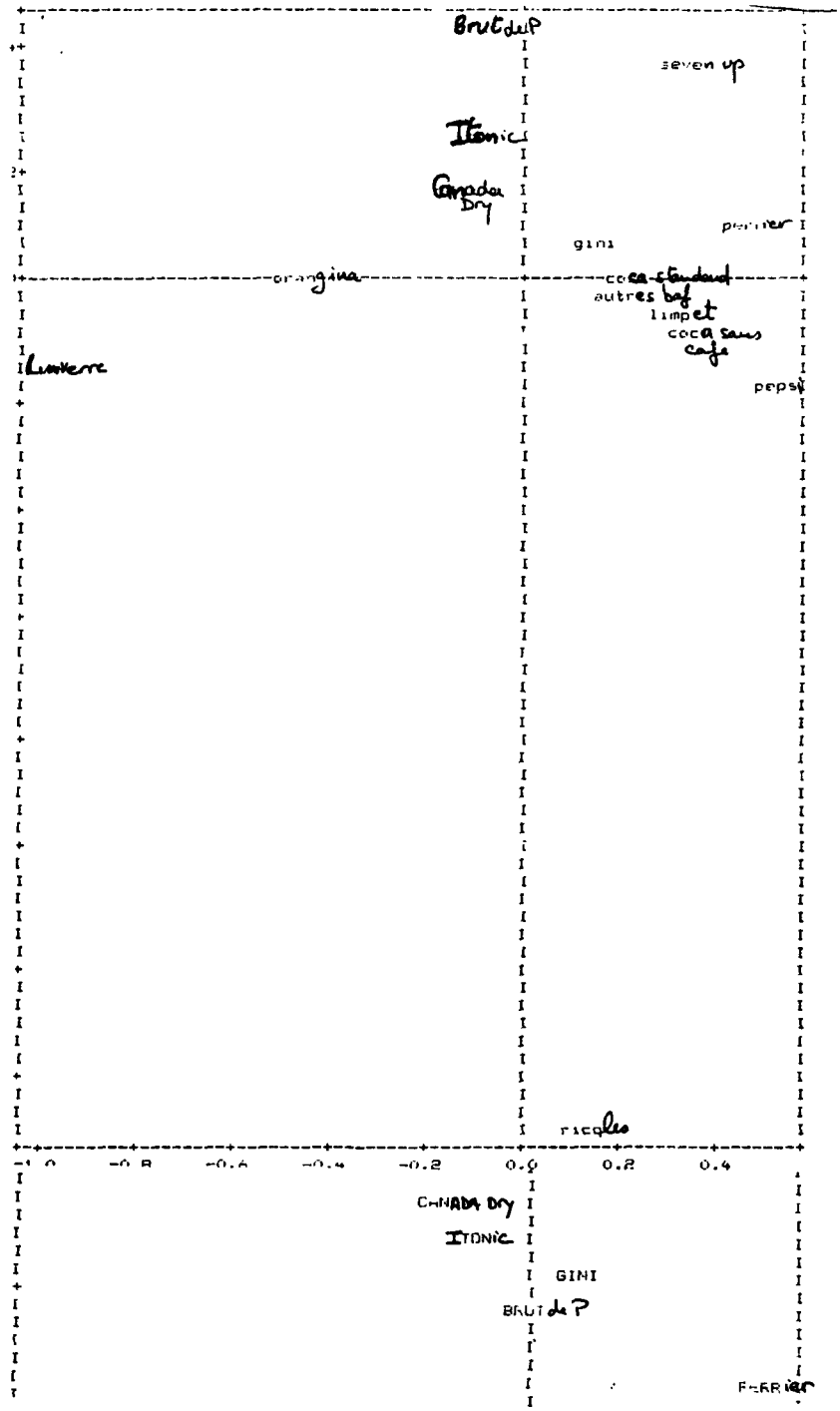
MARCHE BOISSON INDICE ACTE1



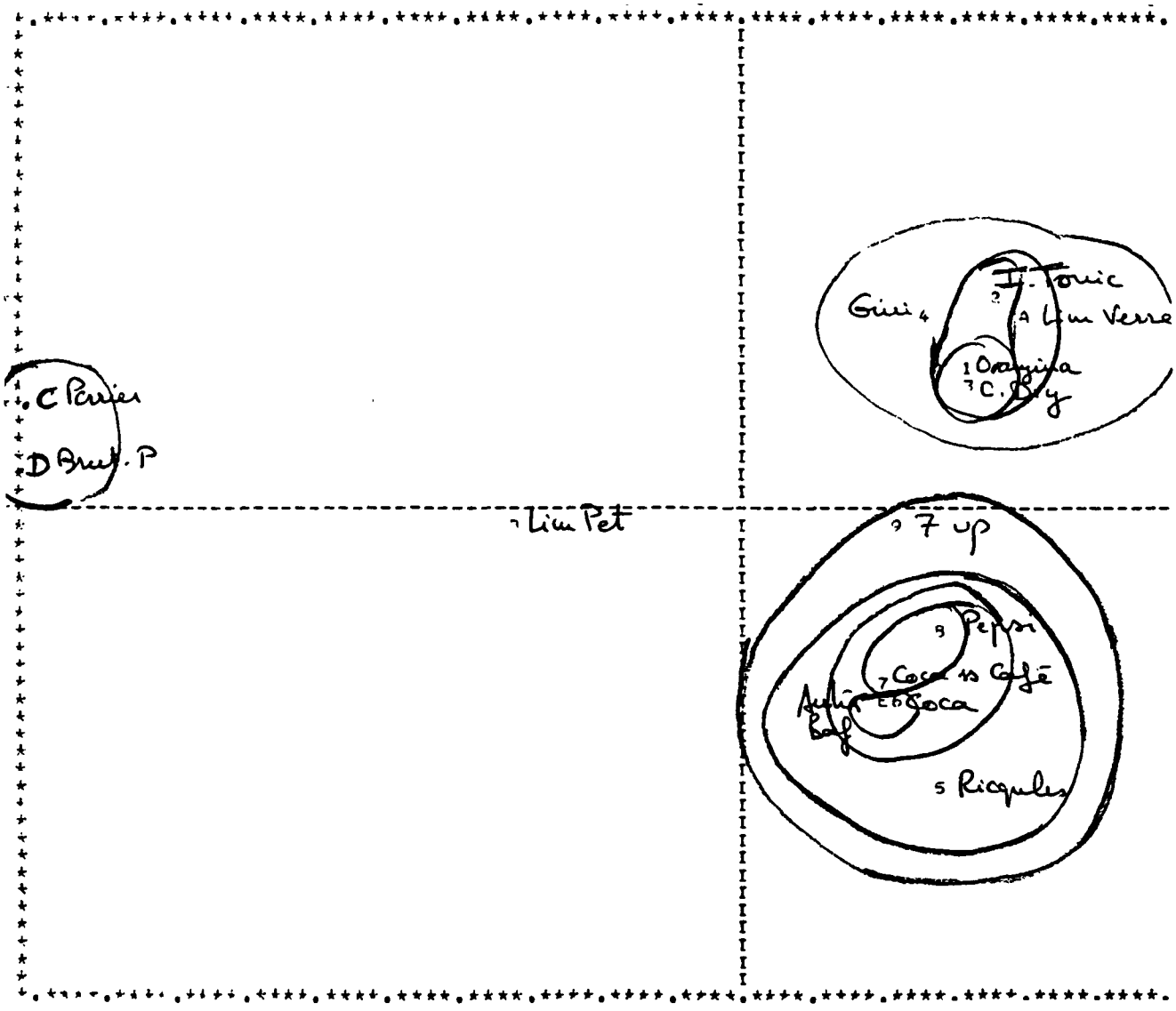
MARCHE BOISSON INDICE TRTO



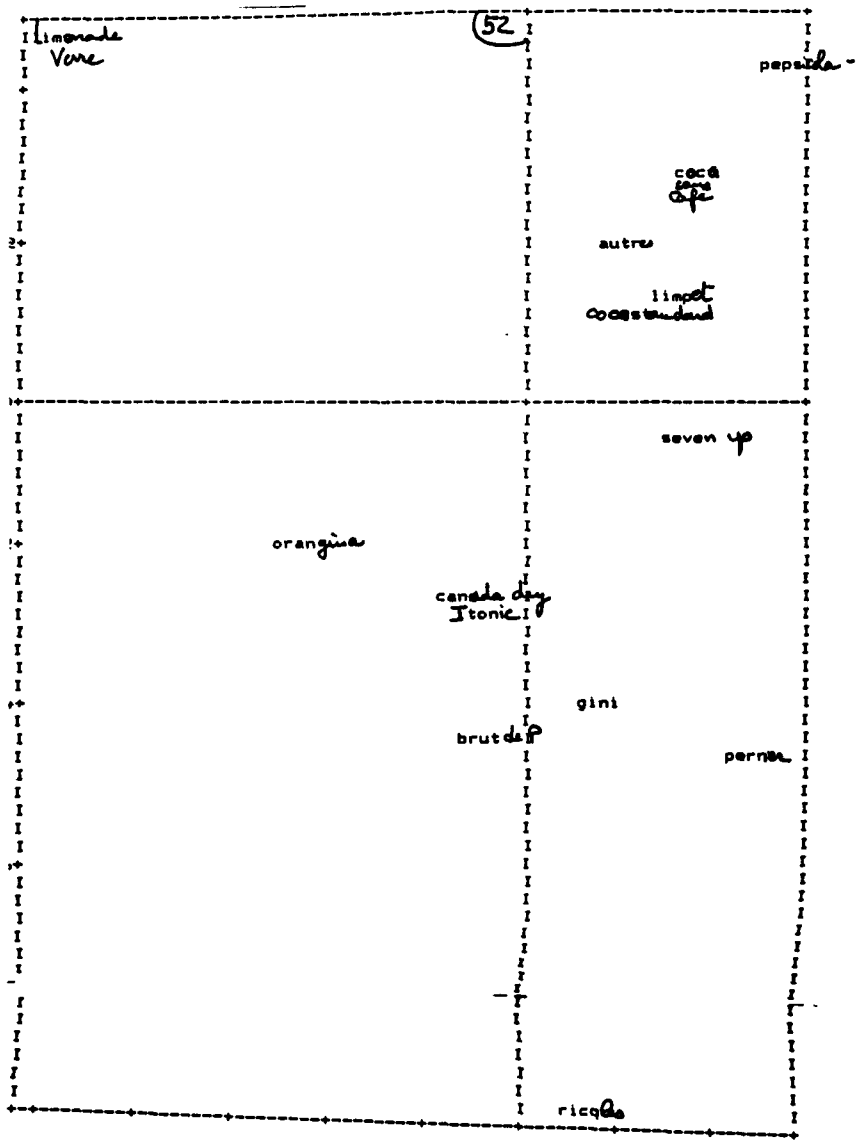
MARCHE BOISSON INDICE ACTE2



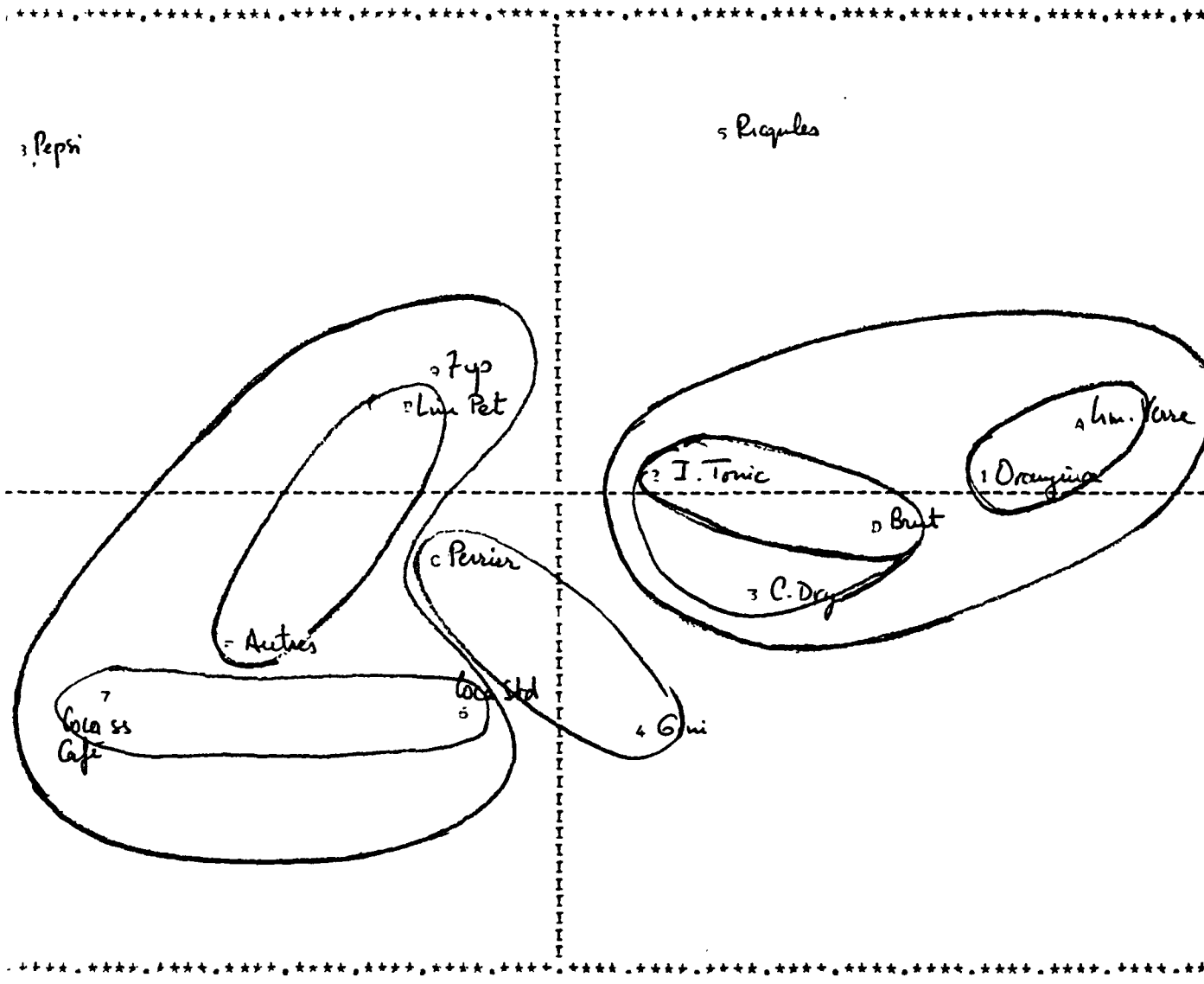
MARCHE BOISSON INDICE ELRO



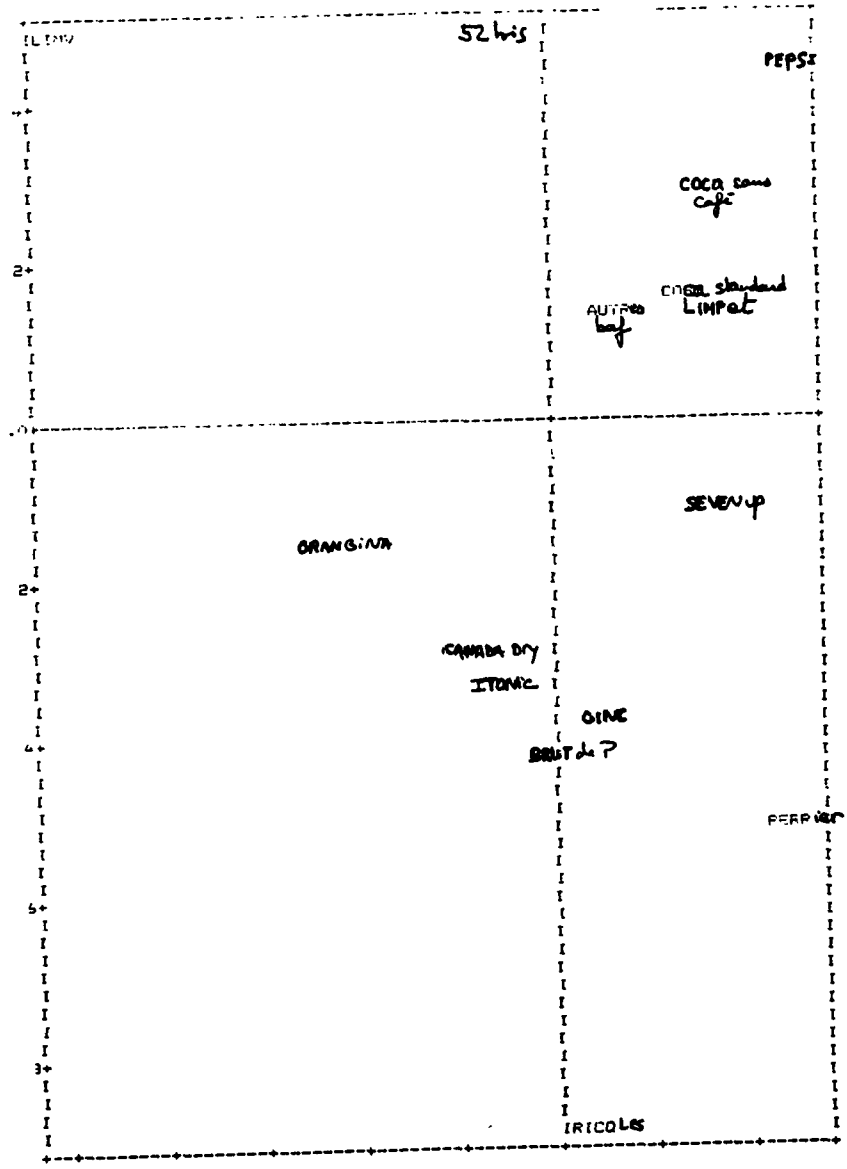
MARCHE BOISSON INDICE INDI1



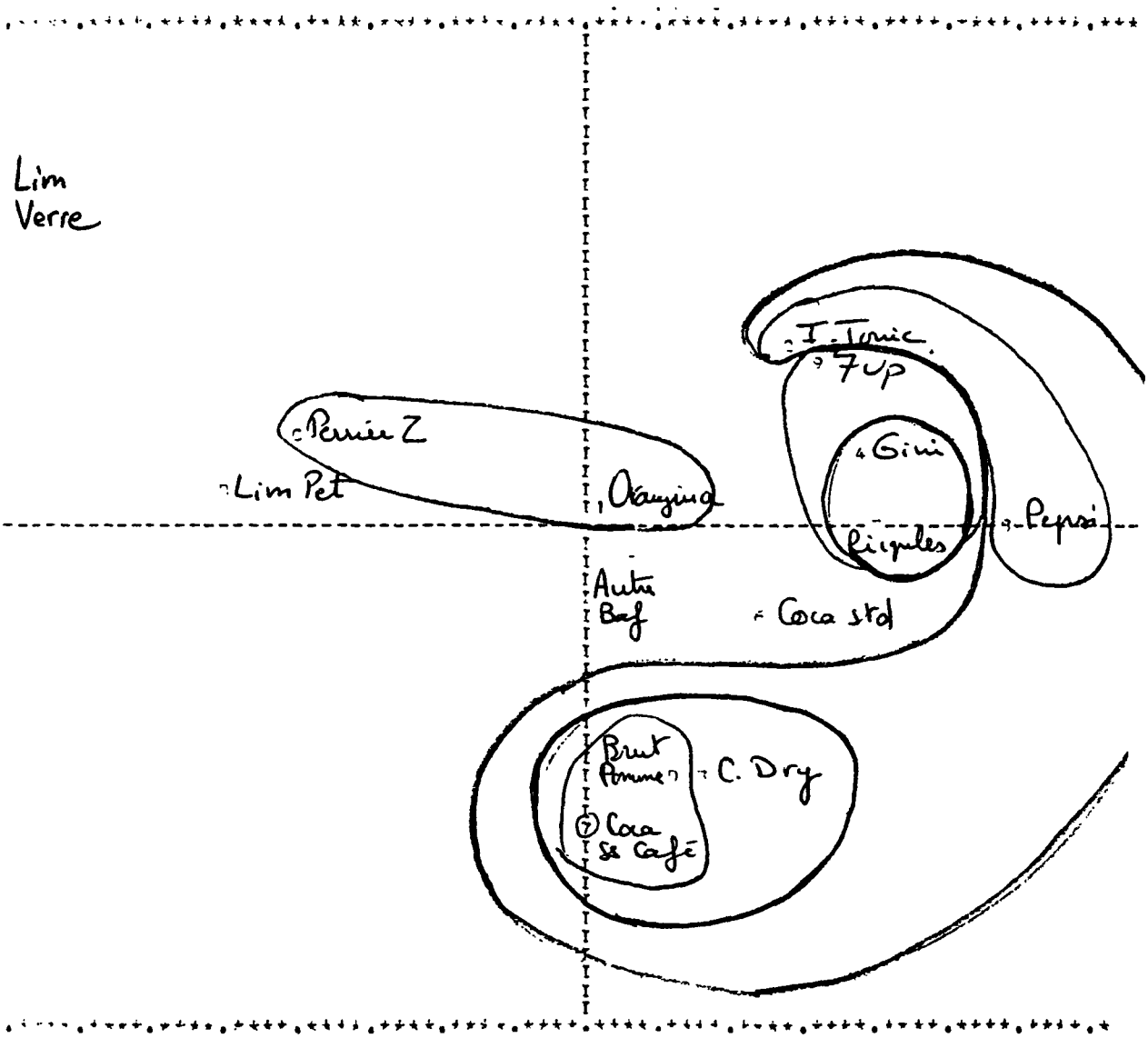
MARCHE BOISSON INDICE TREL



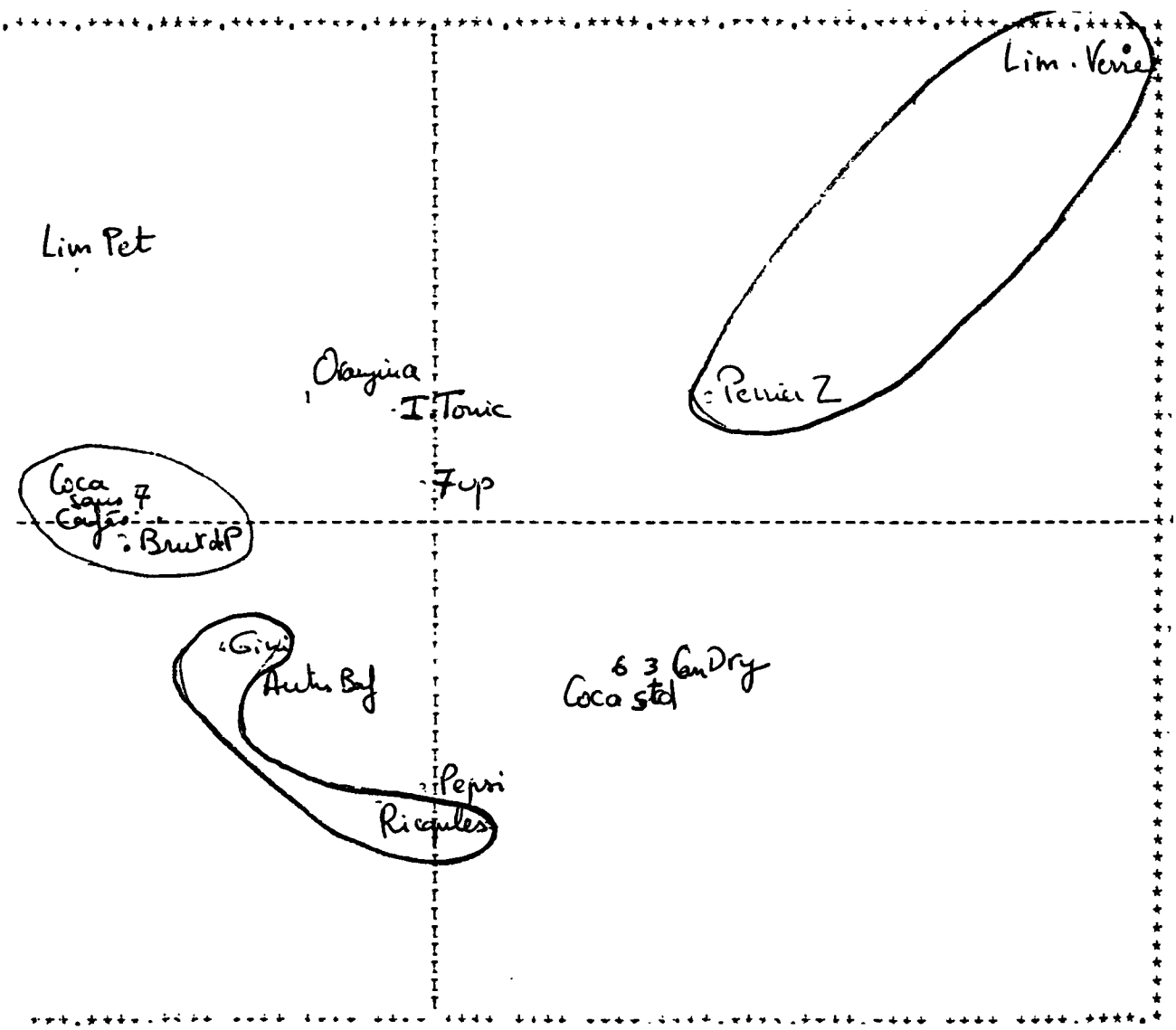
MARCHE BOISSON INDICE INDI2



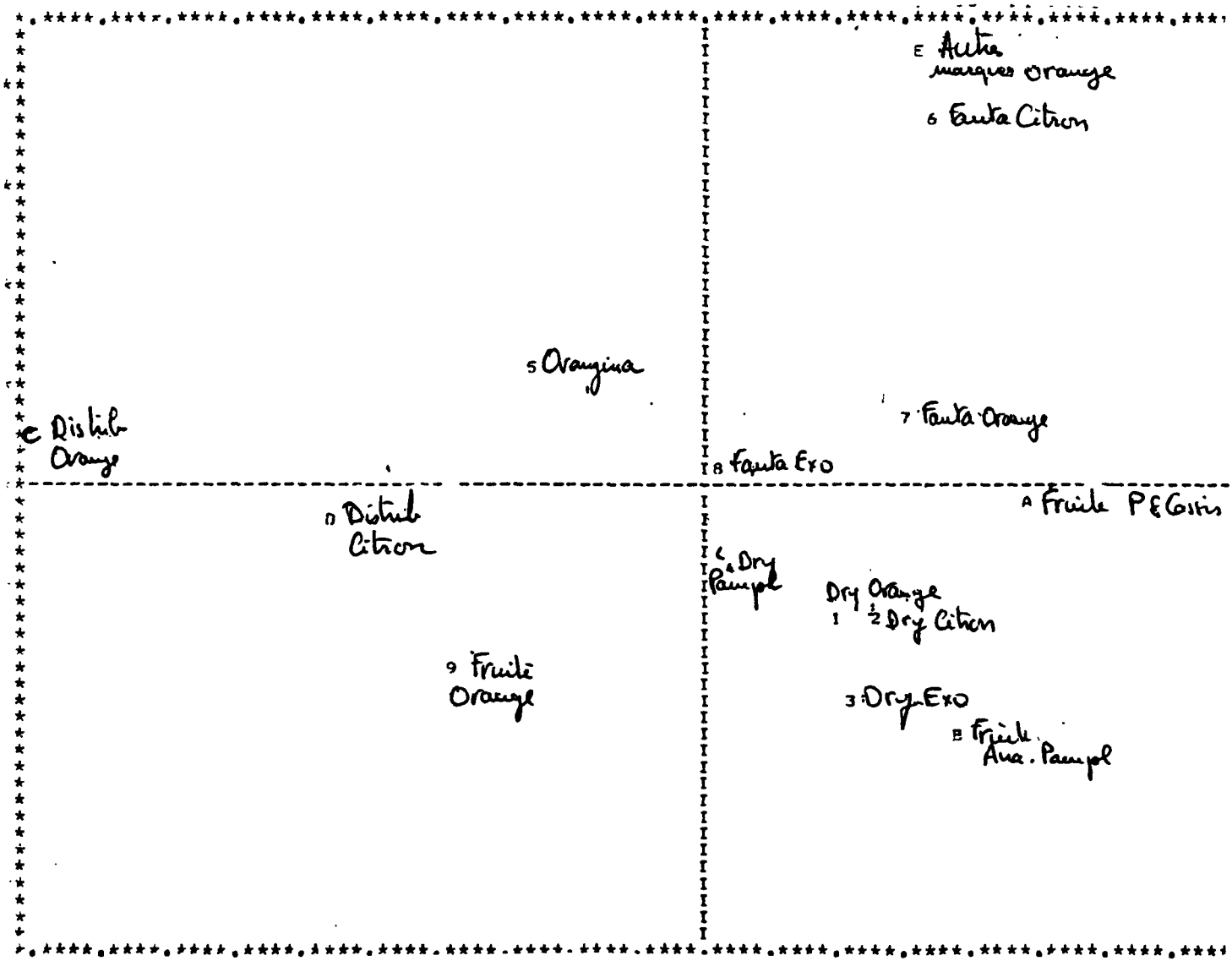
MARCHE BOISSON INDICE MERA



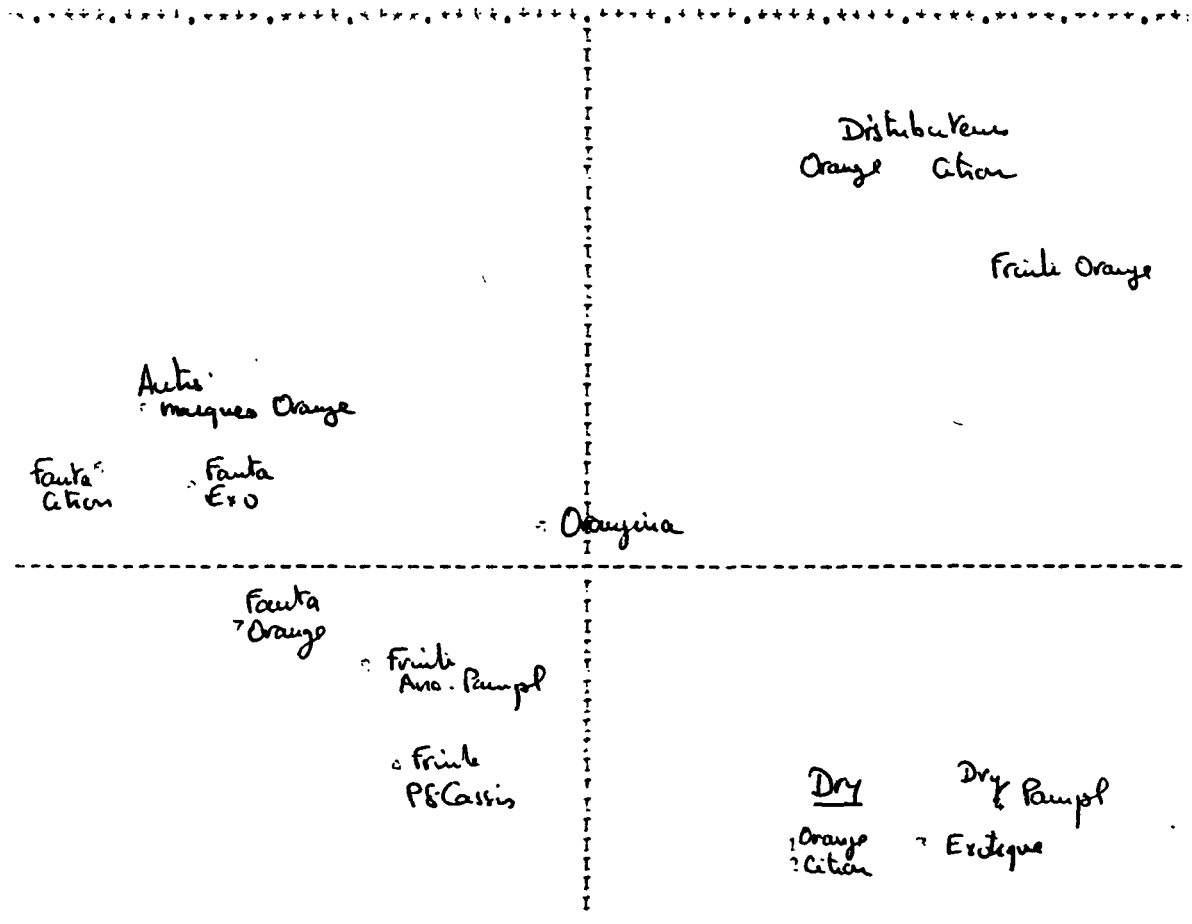
MARCHE BOISSON INDICE MERR



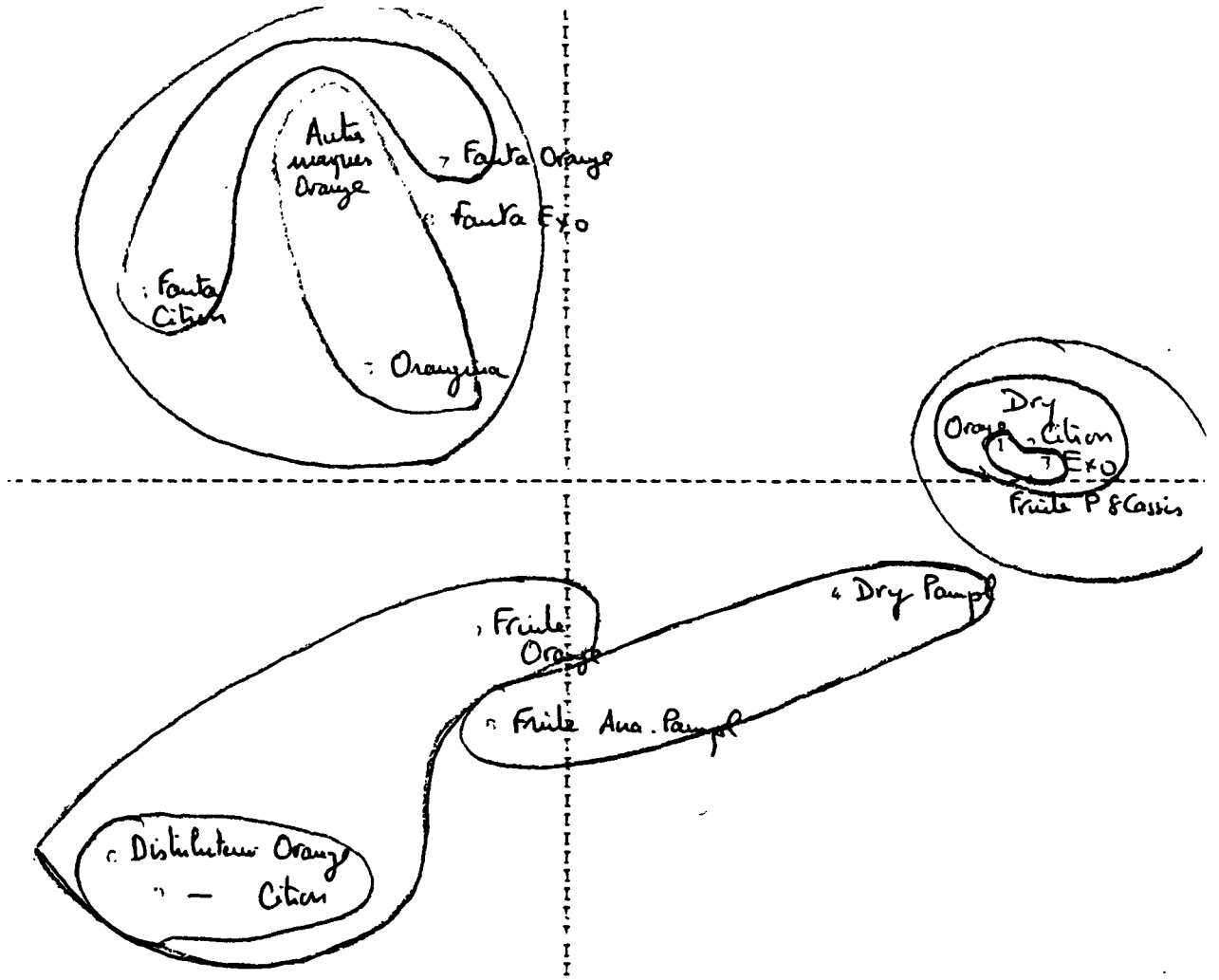
3) MARCHE PARFUM INDICE FRASA



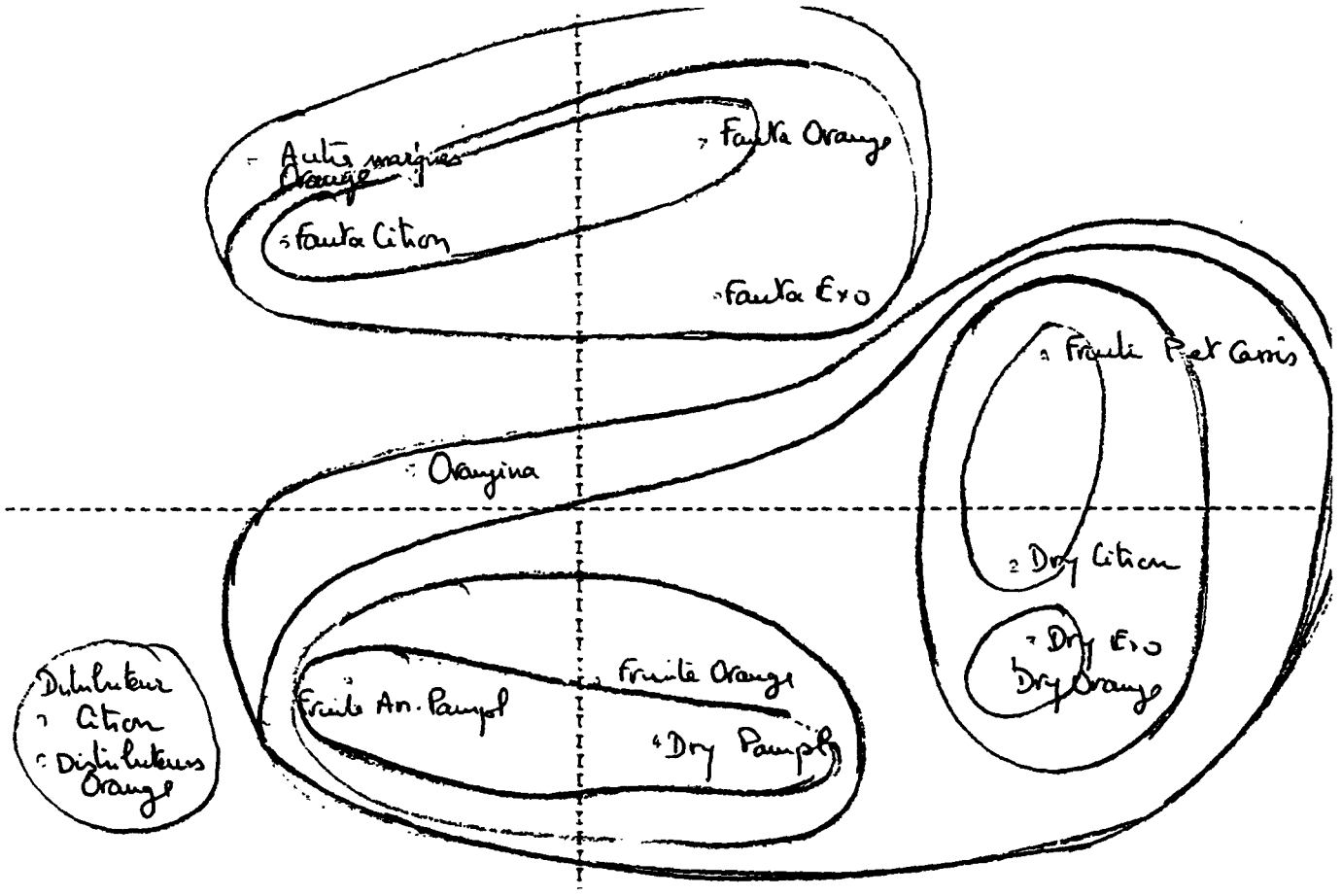
MARCHE PARFUM INDICE FRASB



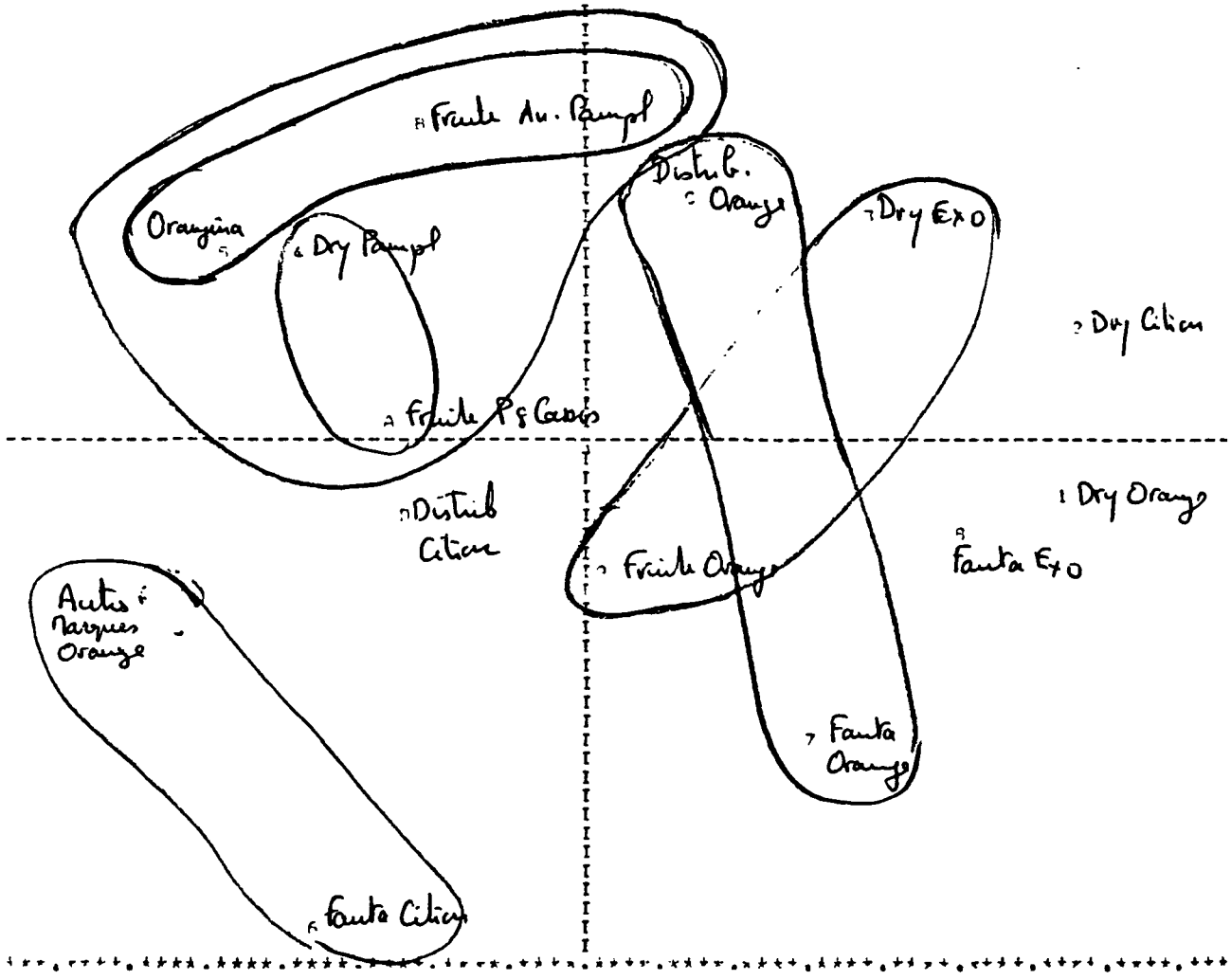
MARCHE PARFUM INDICE LEM



MARCHE PARFUM INDICE RAO

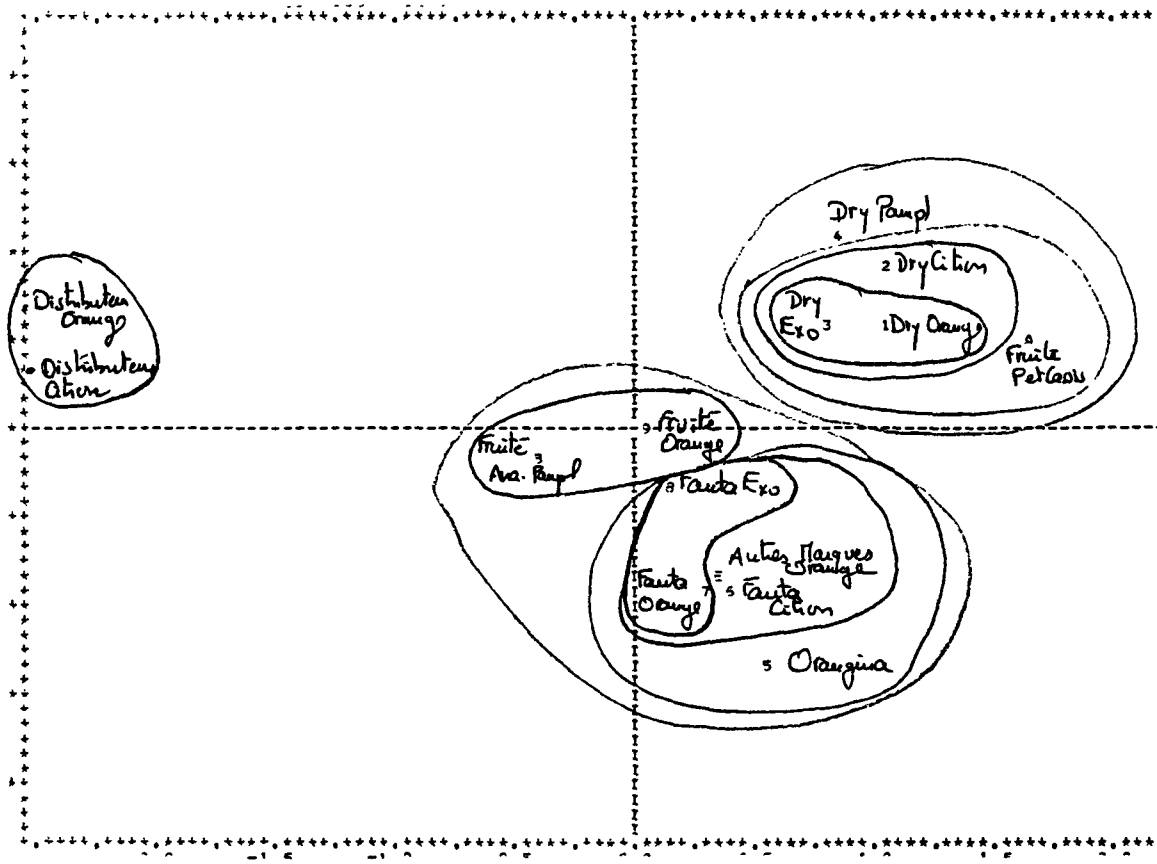


MARCHE PARFUM INDICE SABA

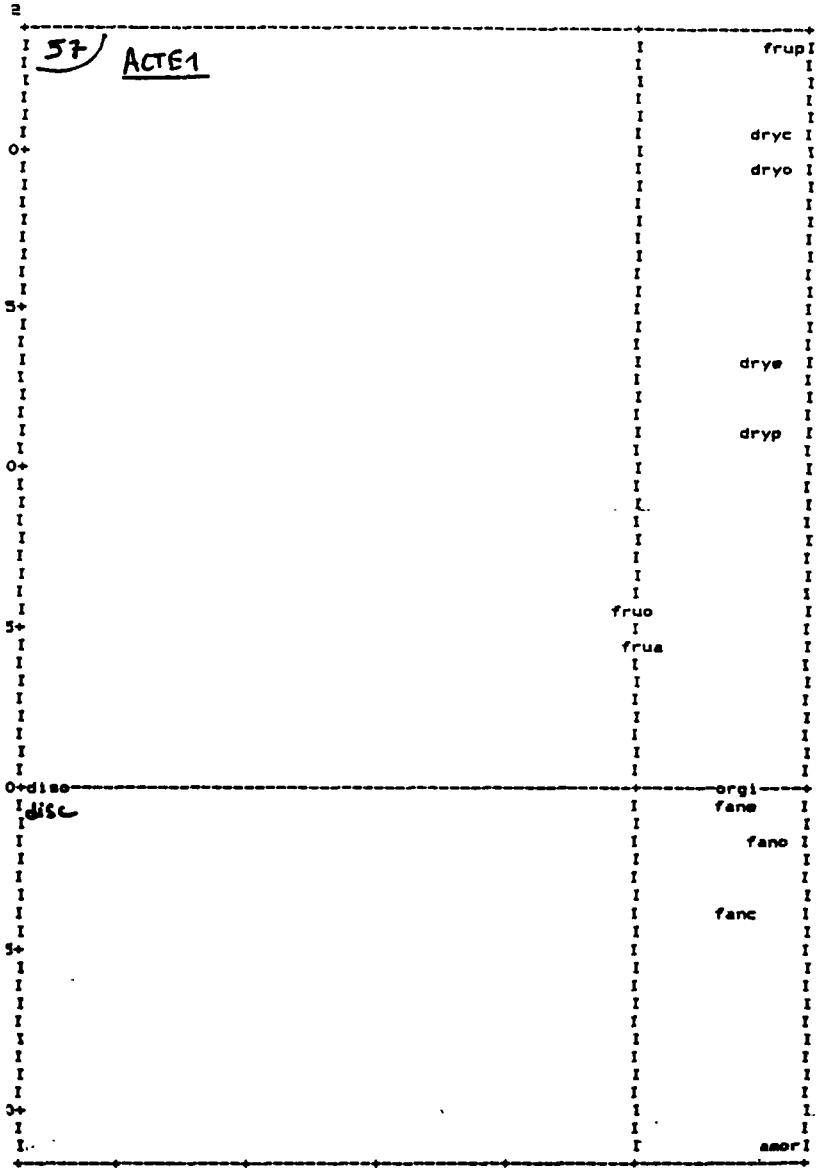


Groupe HEC
BIBLIOTHÈQUE
78350 JOUY-EN-JOSAS

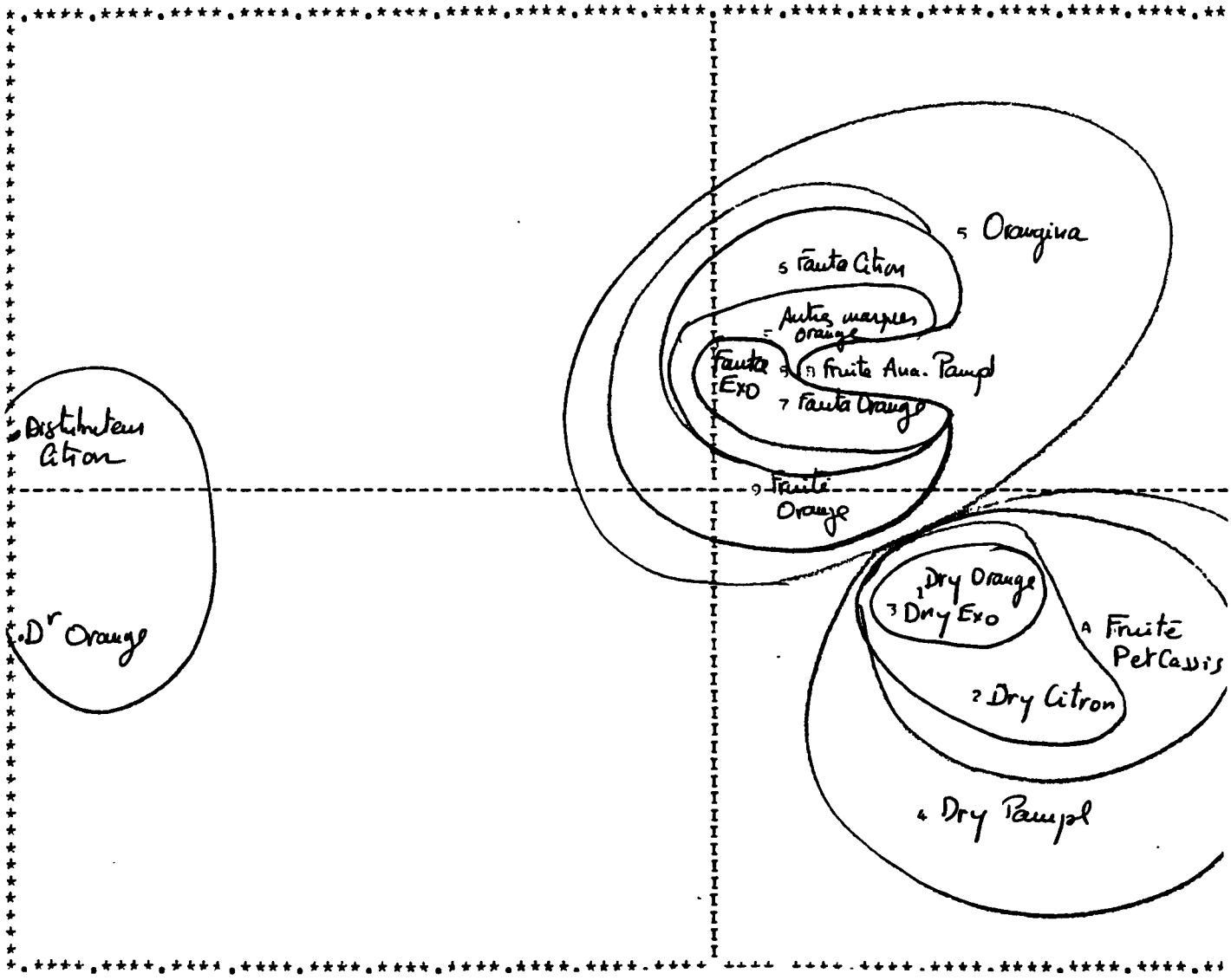
MARCHE PARFUM INDICE TOTT



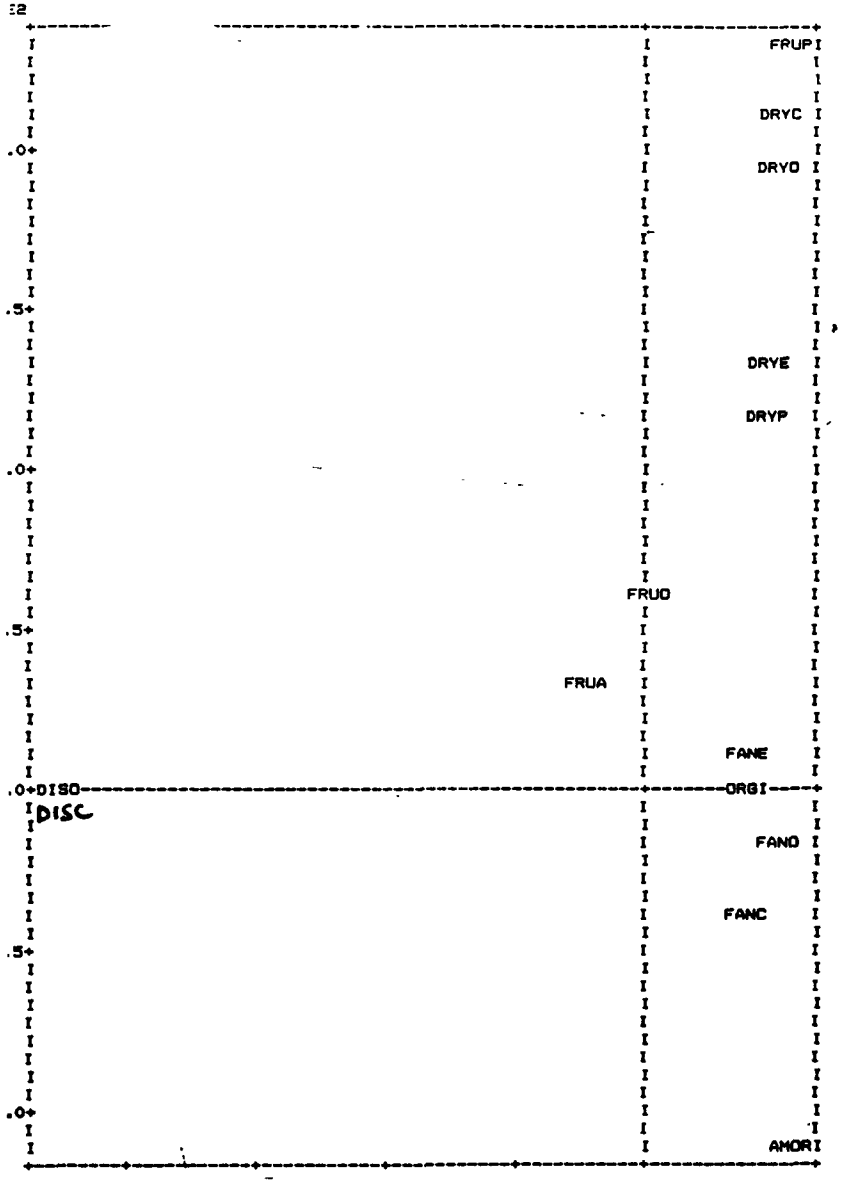
MARCHE PARFUM INDICE ACTE1



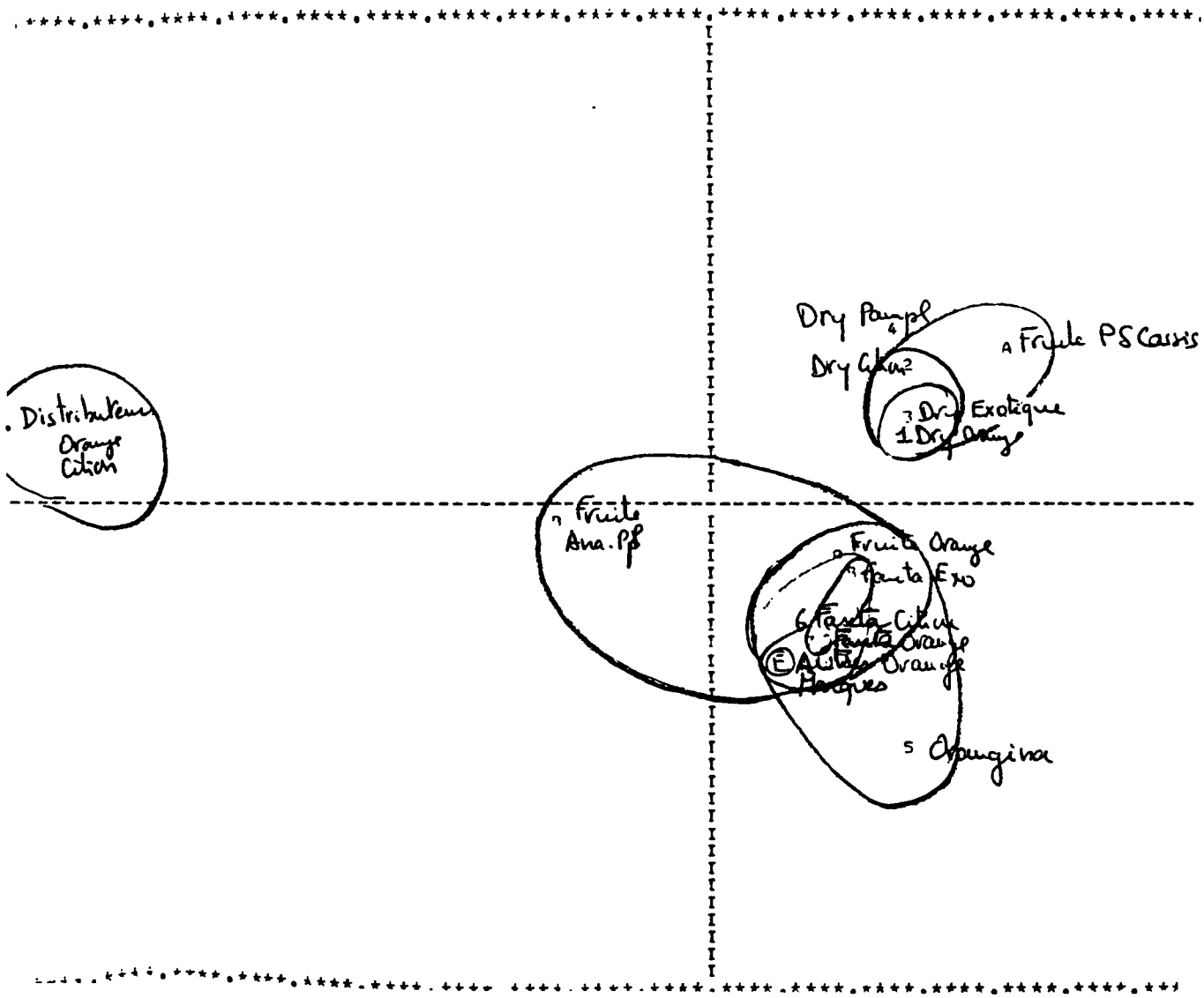
MARCHE PARFUM INDICE TRTO

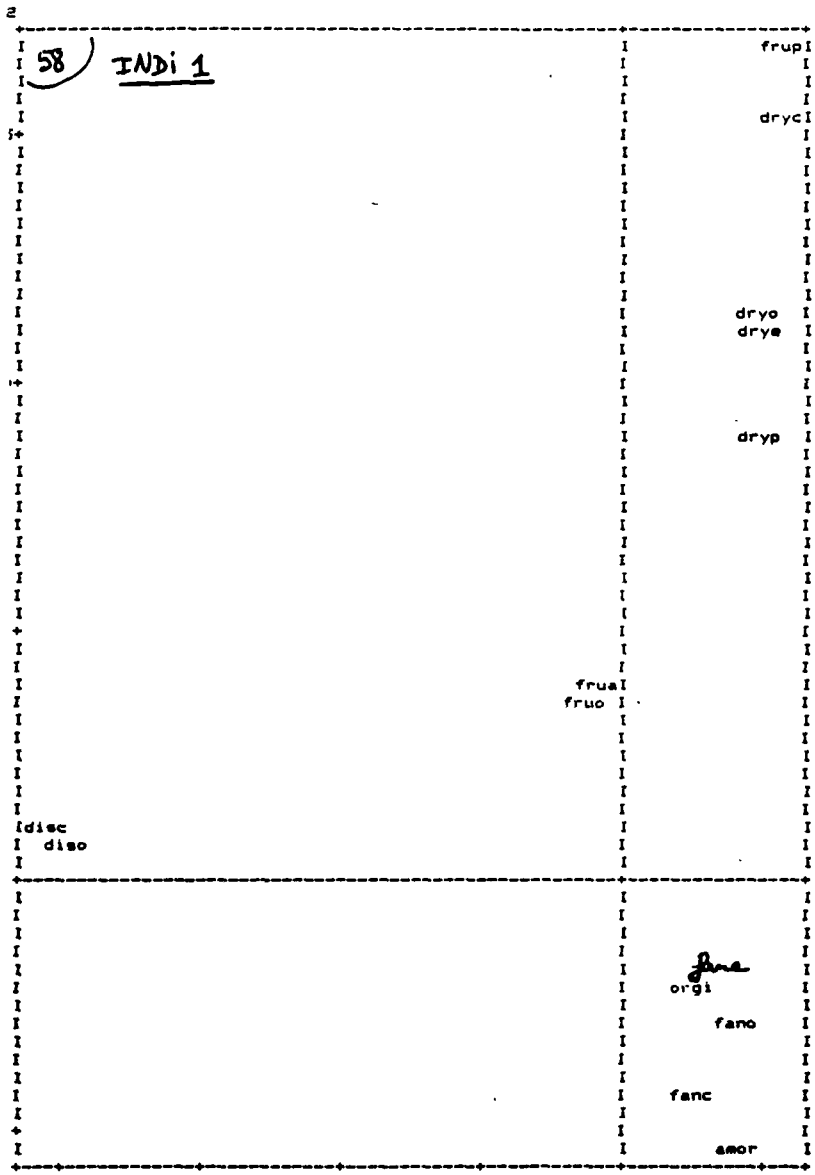


MARCHE PARFUM INDICE ACTE2

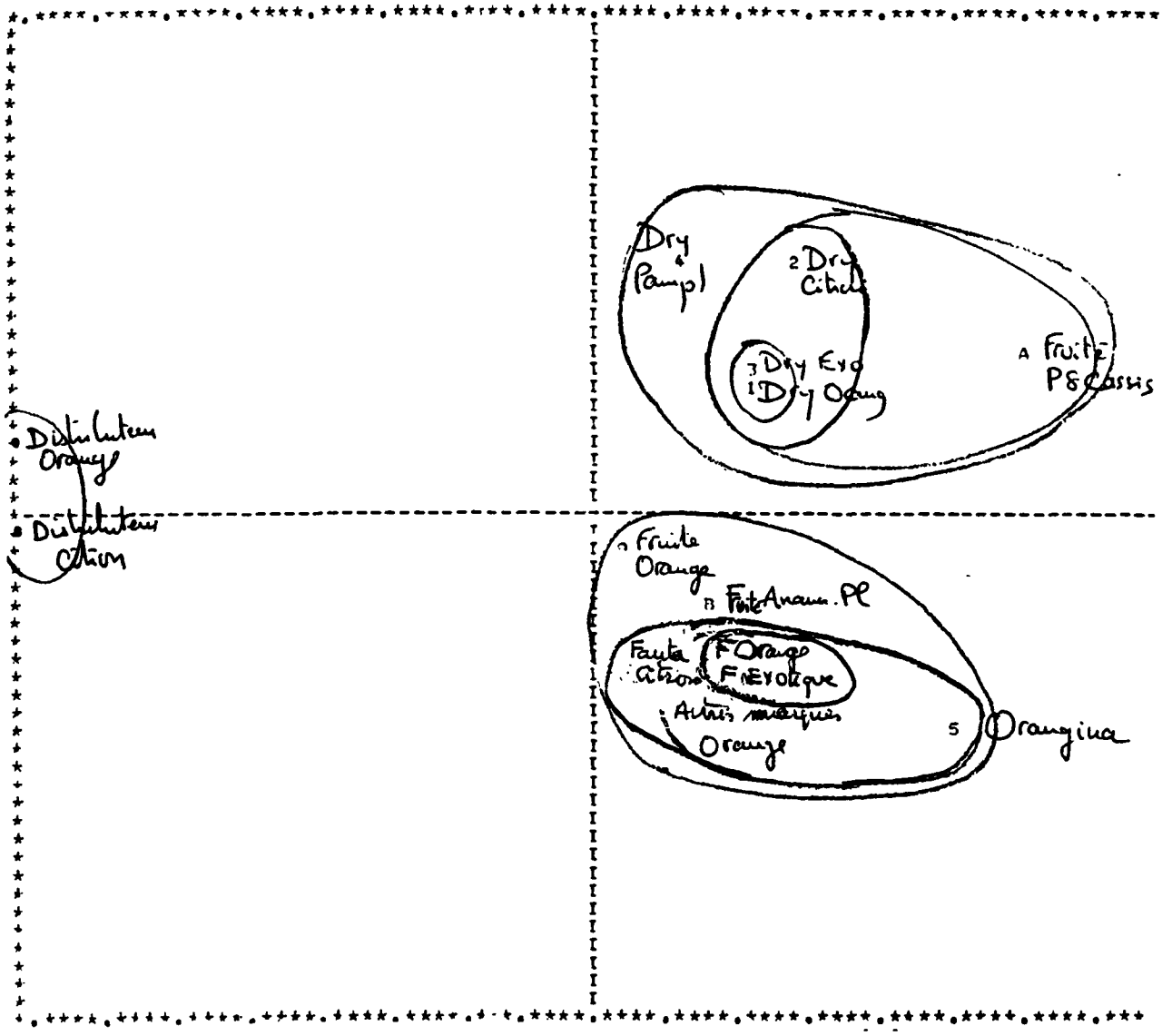


MARCHE PARFUM INDICE ELRO

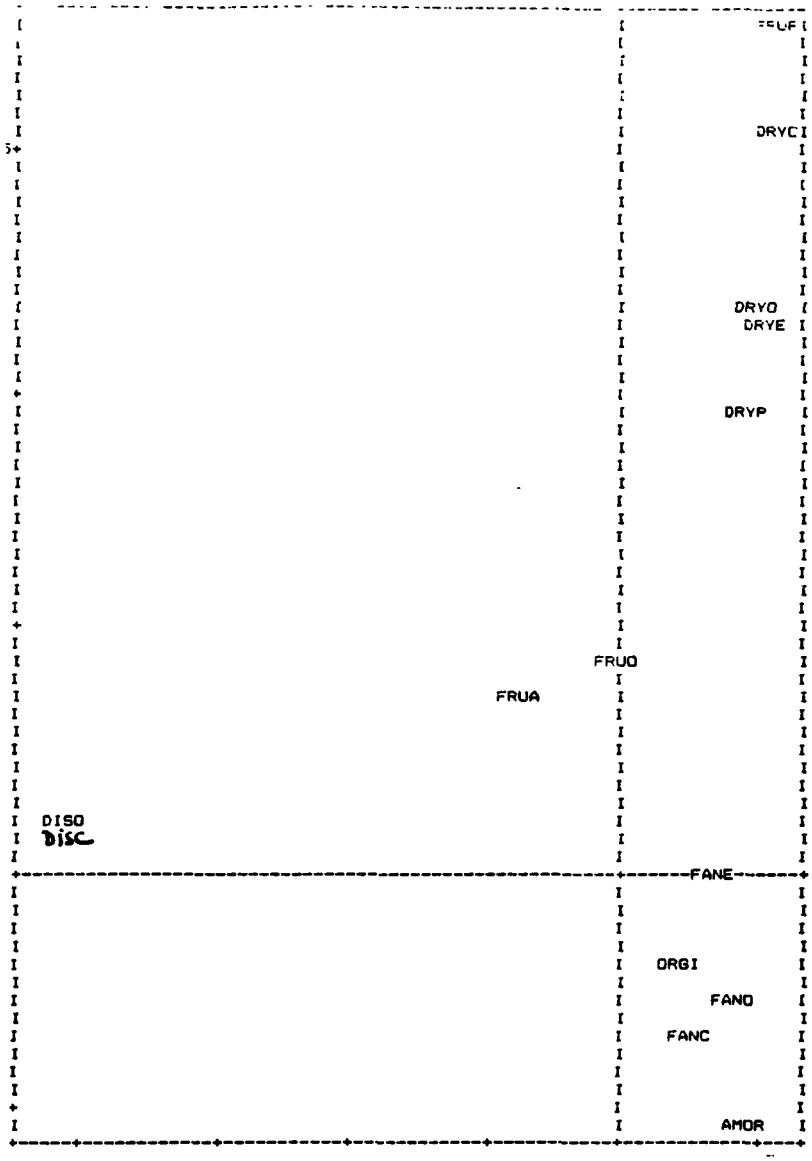



58
INDi 1
frup
dryc
dryo
drye
dryp
frua
fruo
disc
diso
orgi
fano
fanc
amor

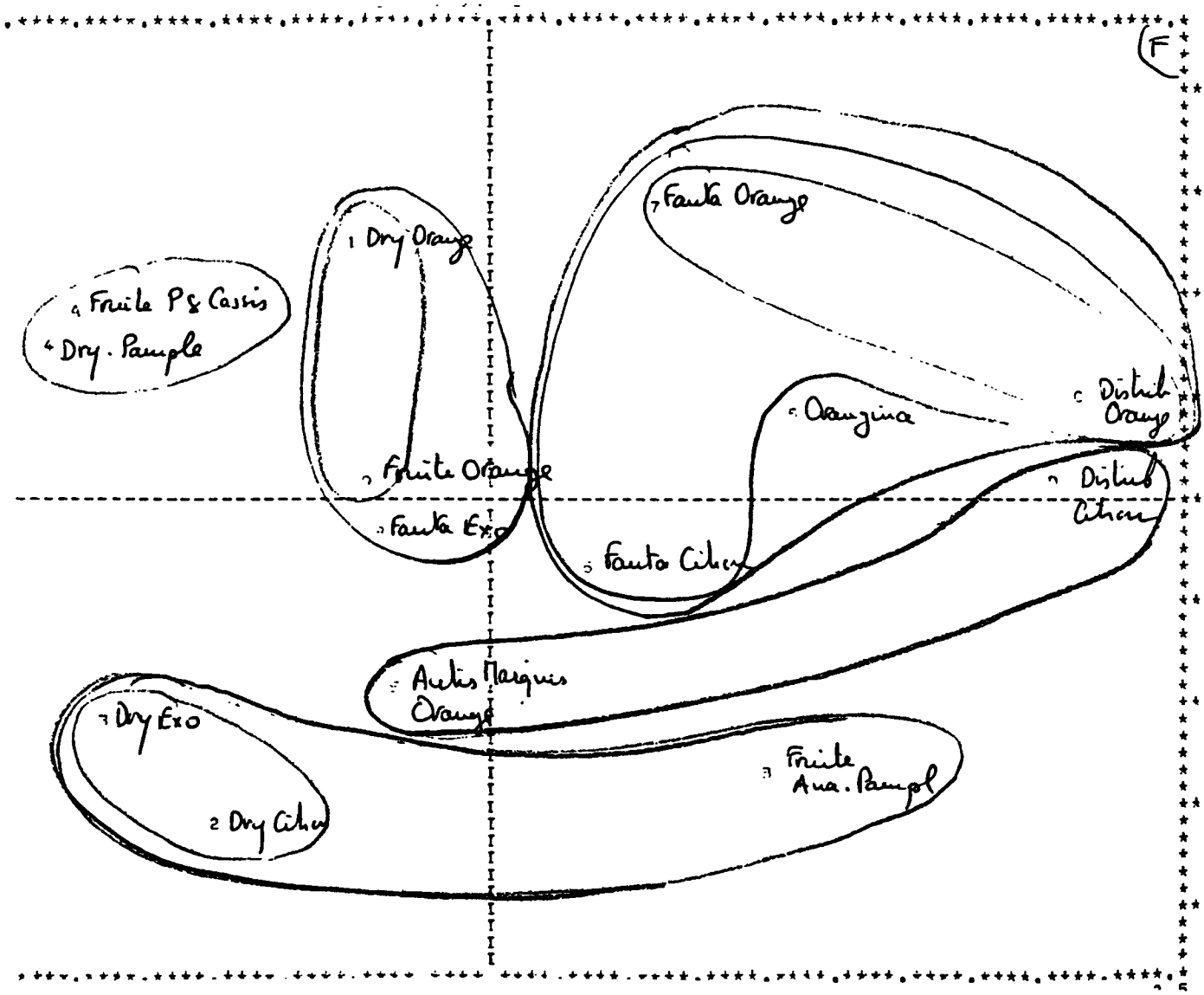
MARCHE PARFUM INDICE TREL



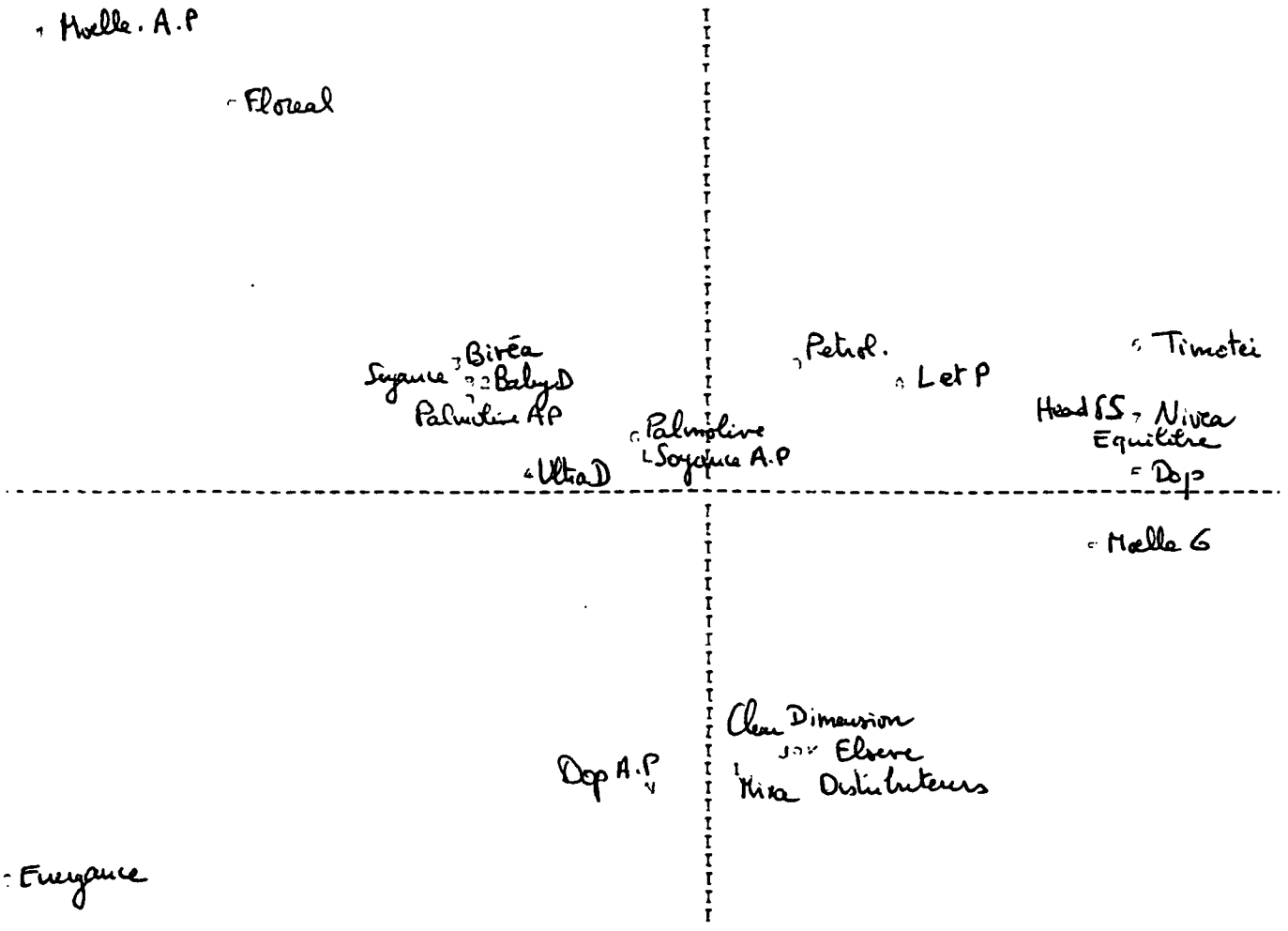
MARCHE PARFUM INDICE INDI2



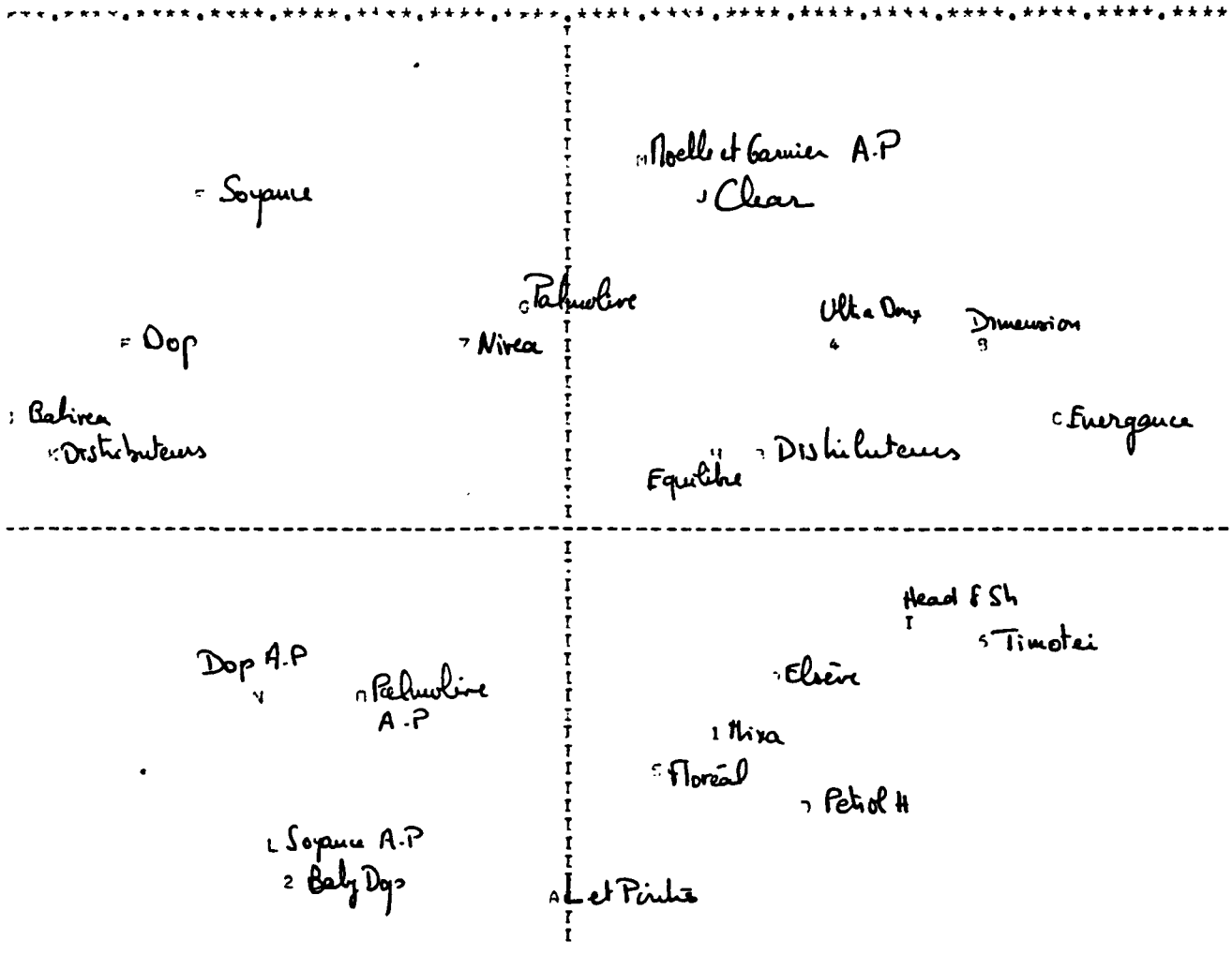
MARCHE PARFUM INDICE MERR



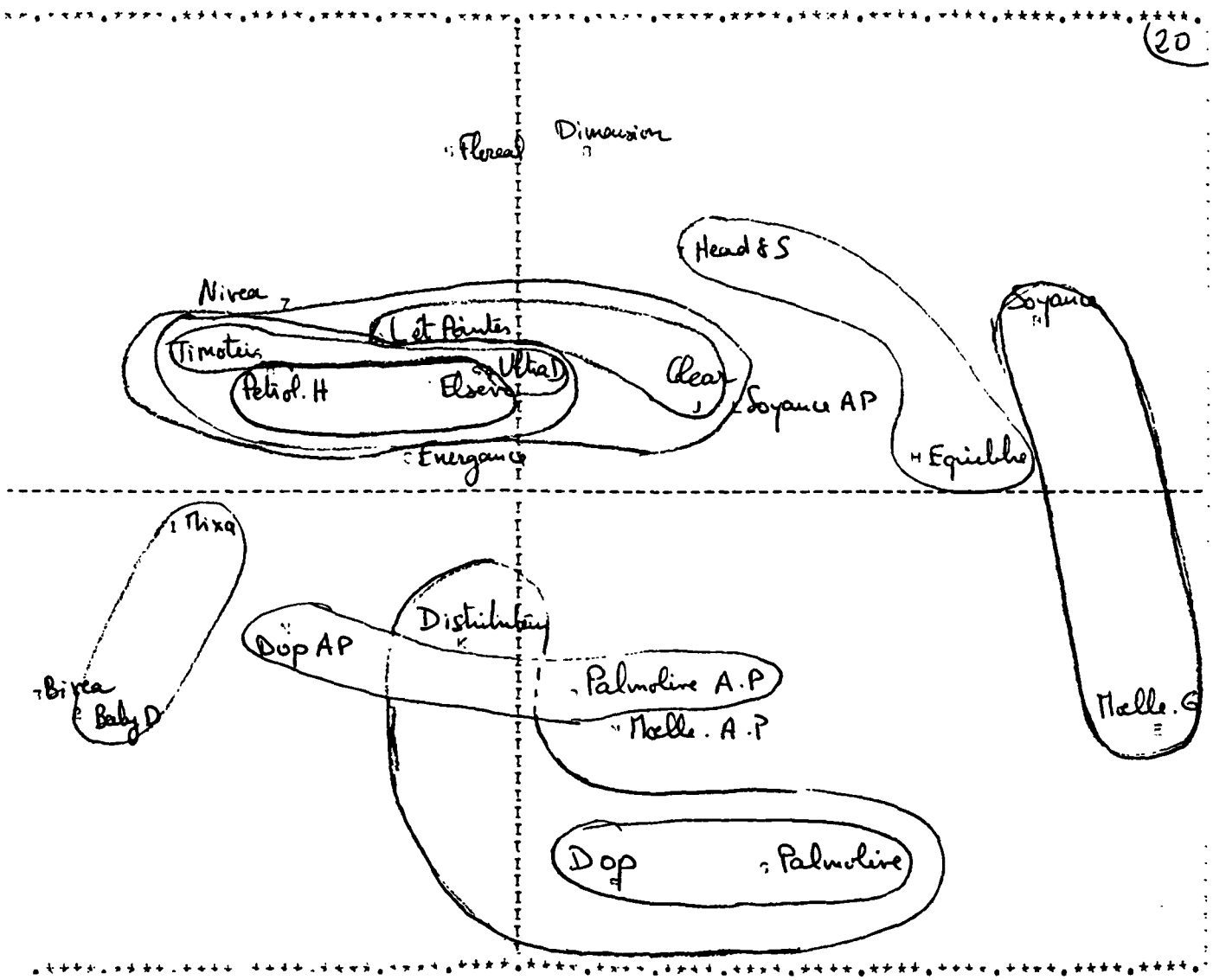
4) MARCHE SHAMP INDICE FRASA



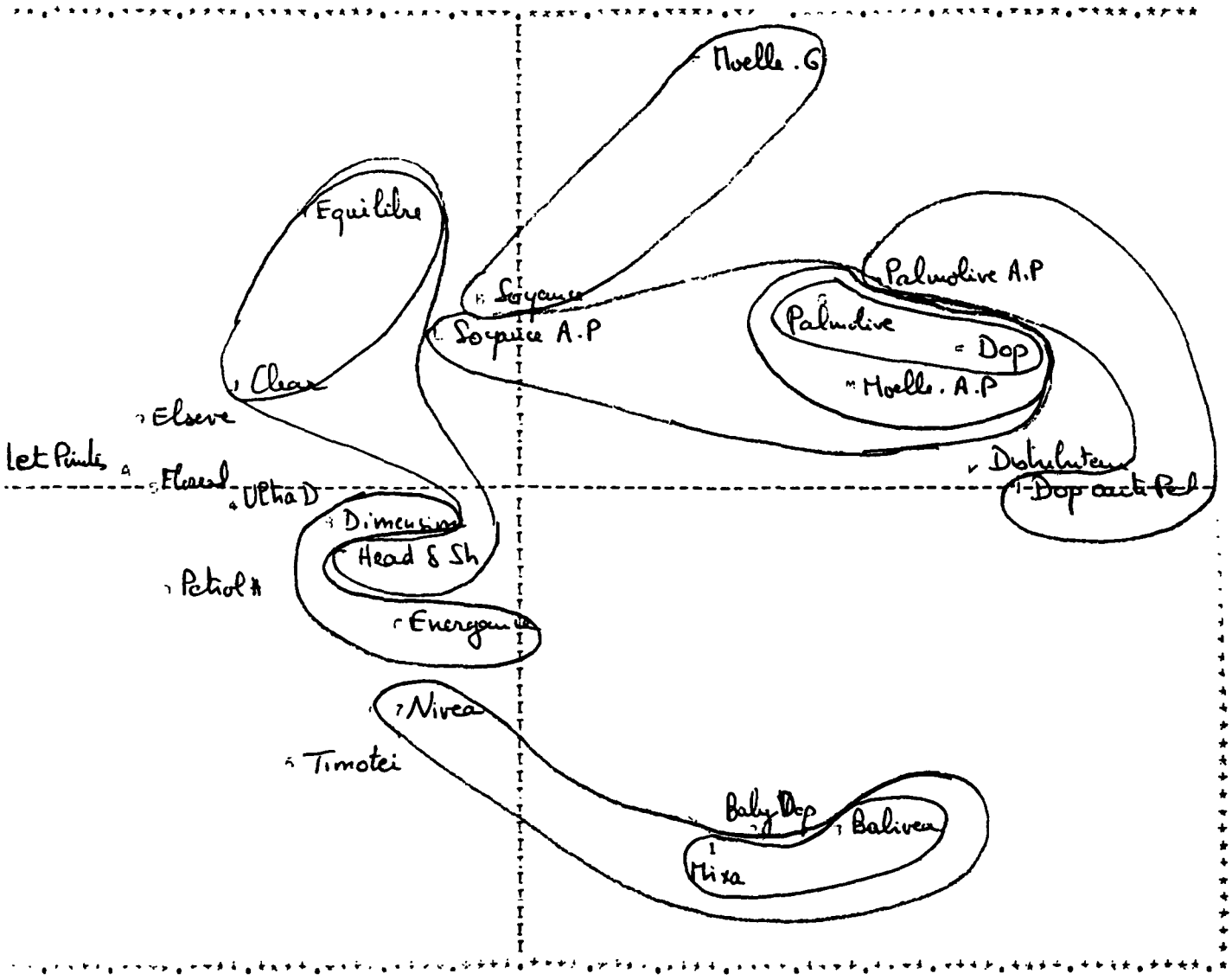
MARCHE SHAMP INDICE FRASB



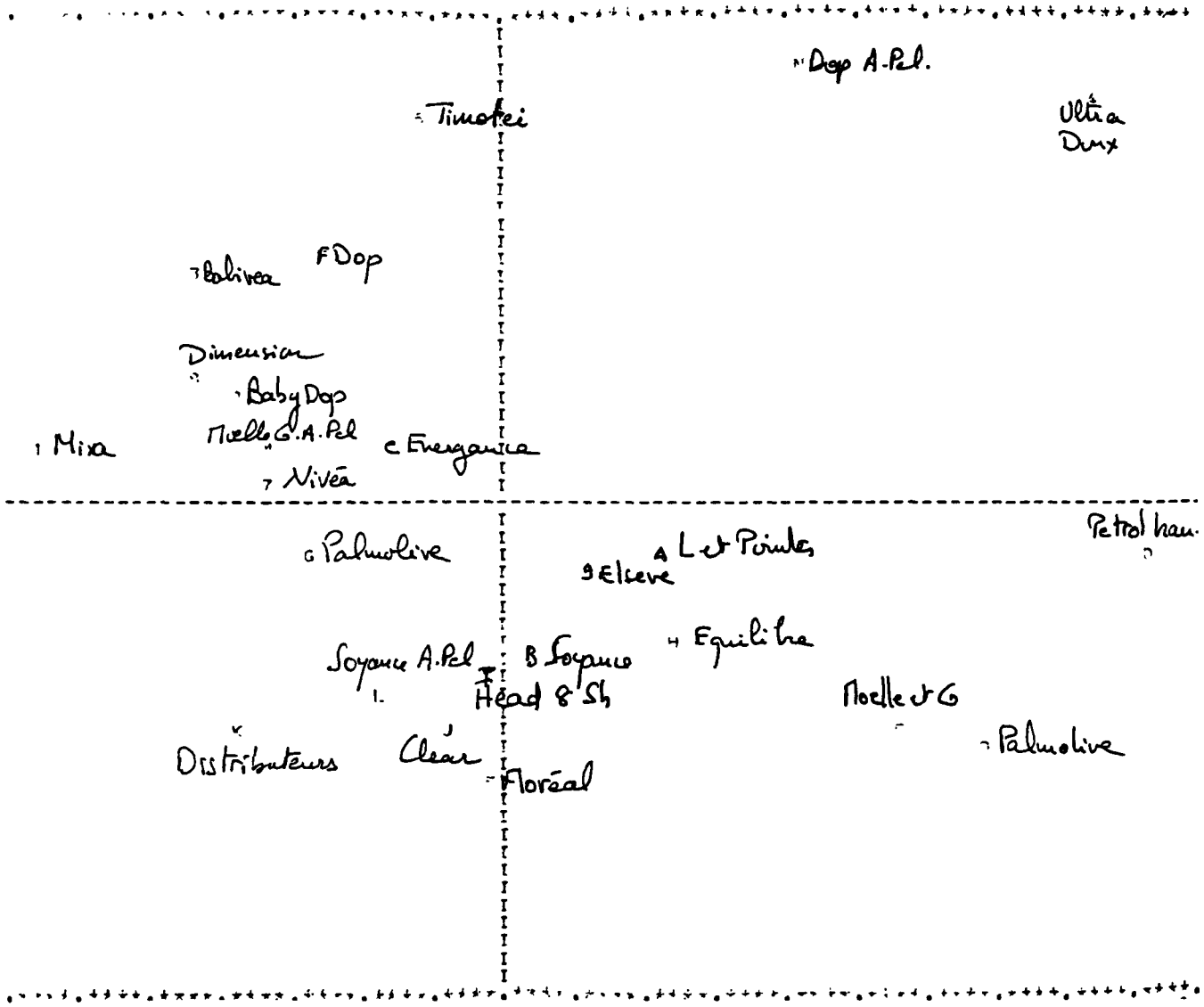
MARCHE SHAMP INDICE LEM



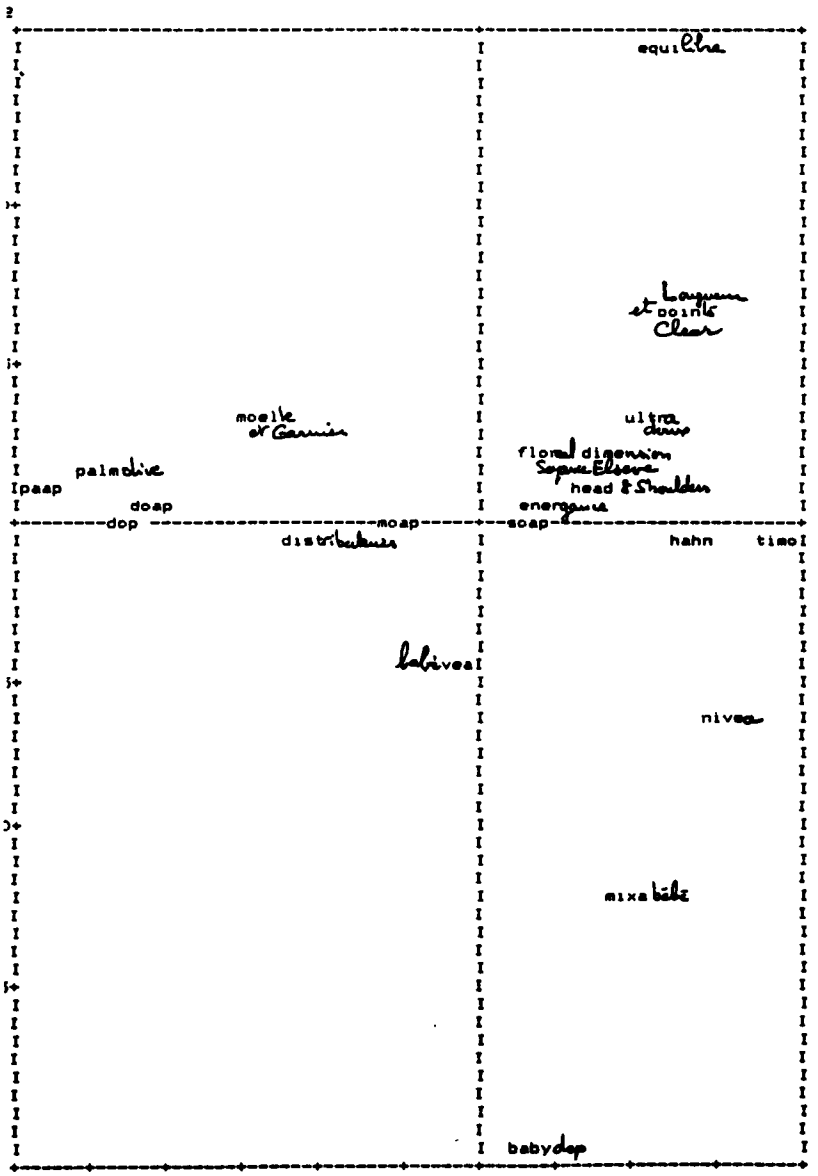
MARCHE SHAMP INDICE RAO



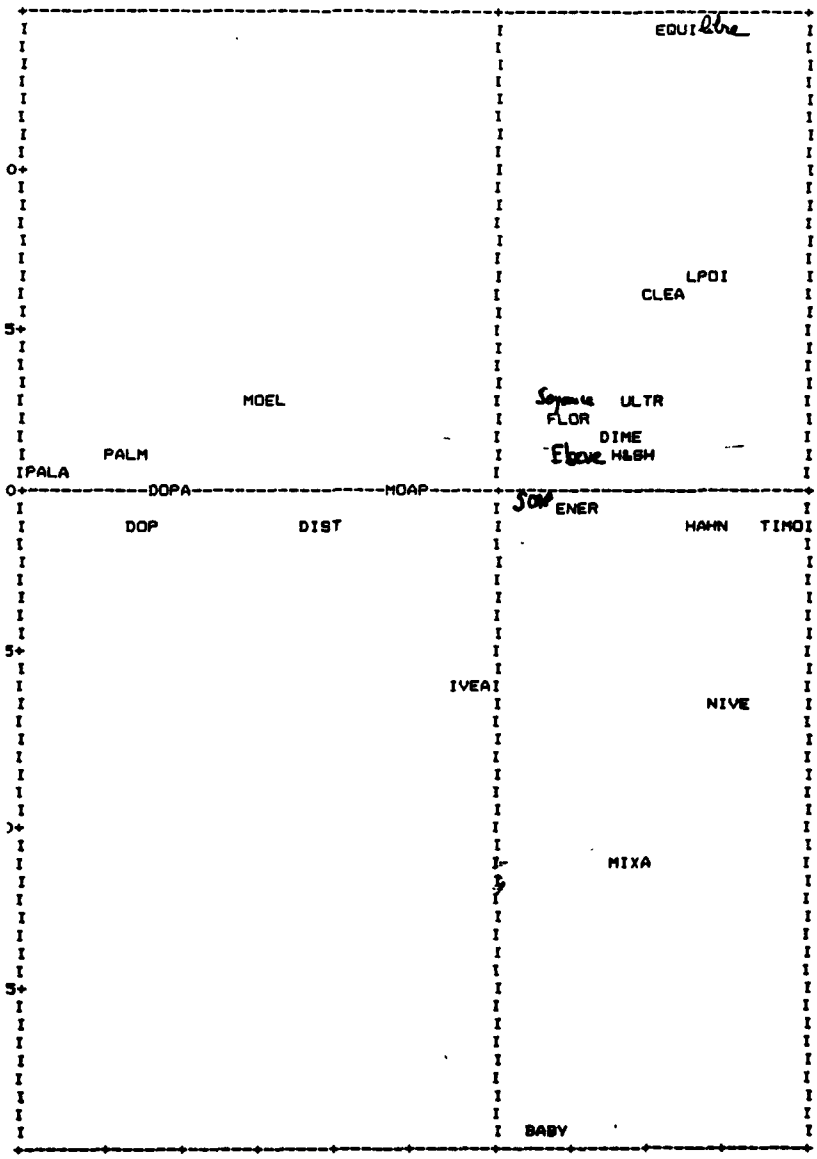
MARCHE SHAMP INDICE SABA



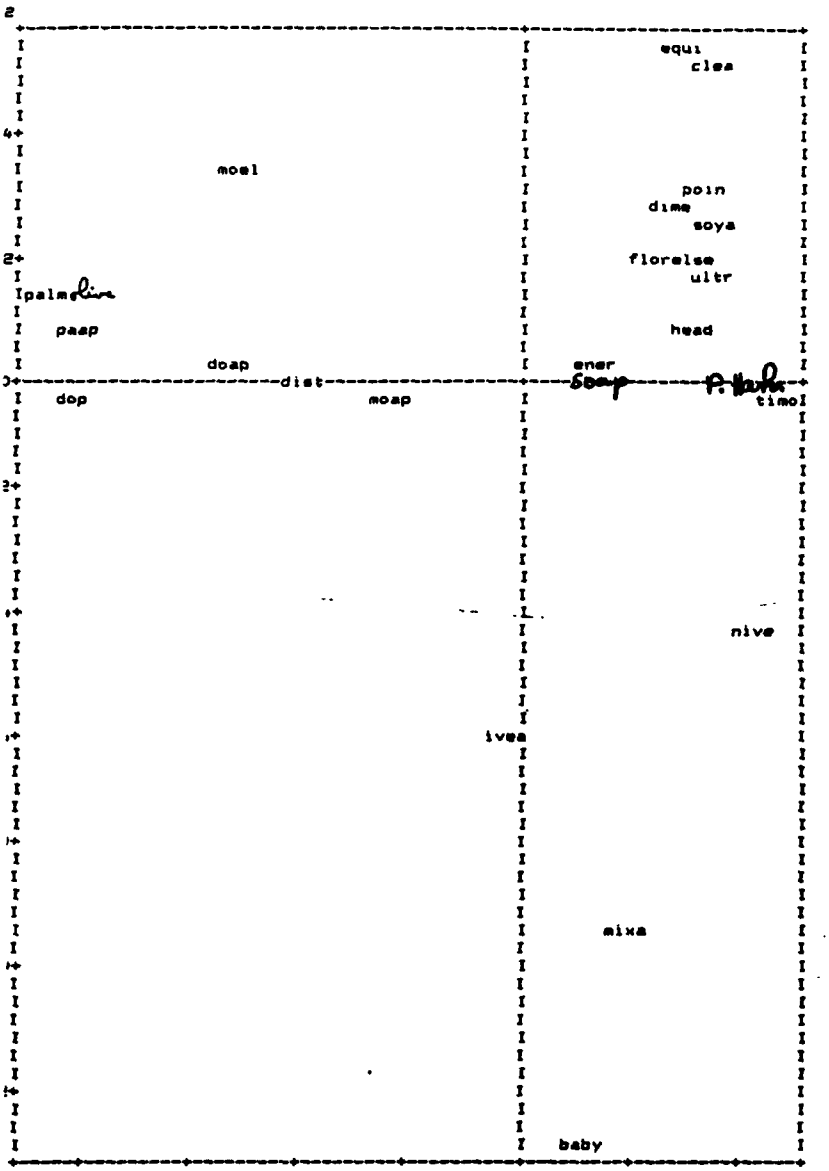
MARCHE SHAMP INDICE ACTE1



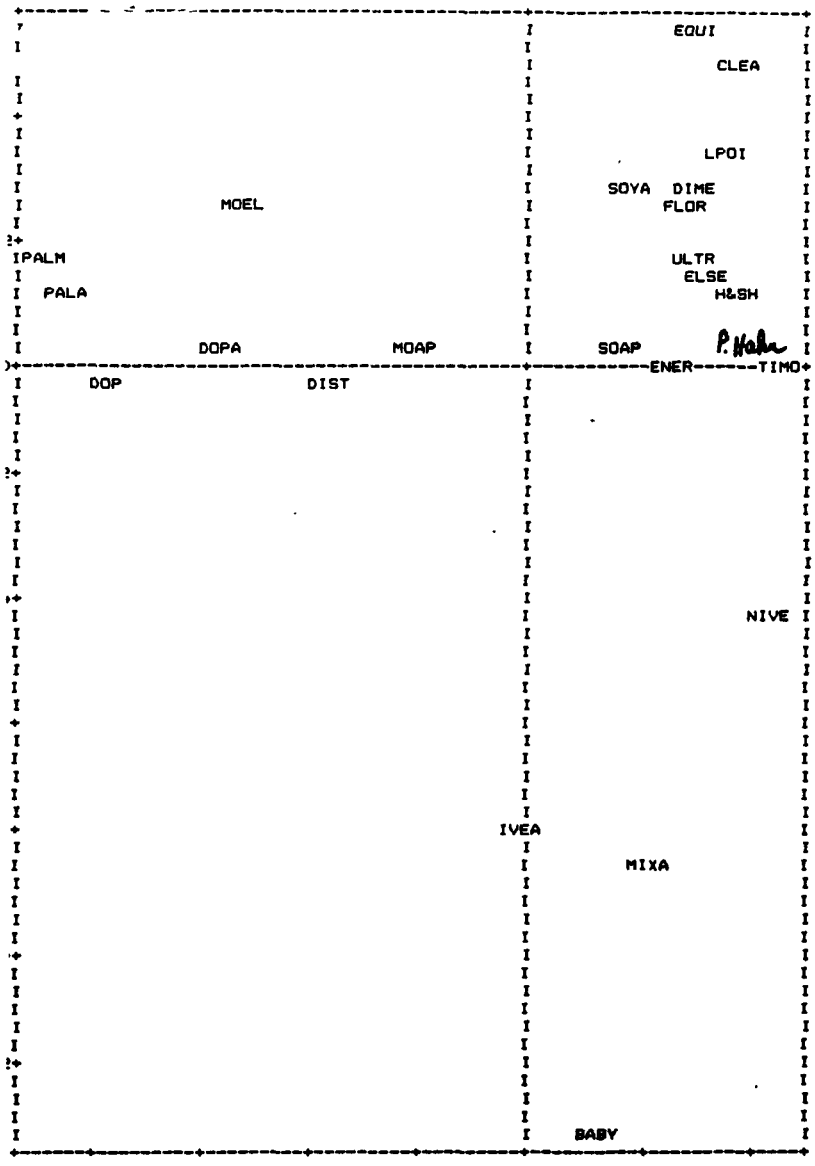
MARCHE SHAMP INDICE ACTE2



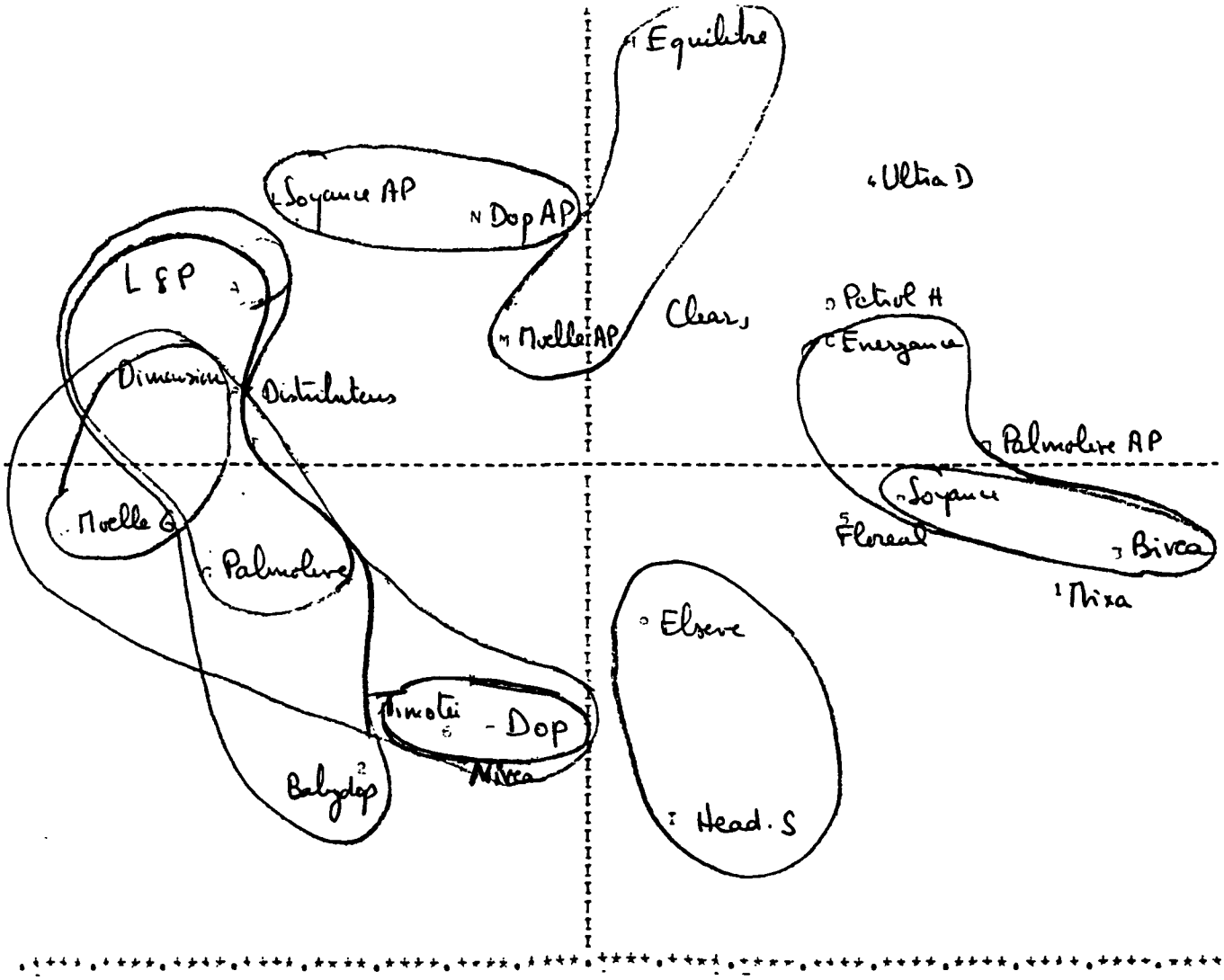
MARCHE SHAMP INDICE INDI1



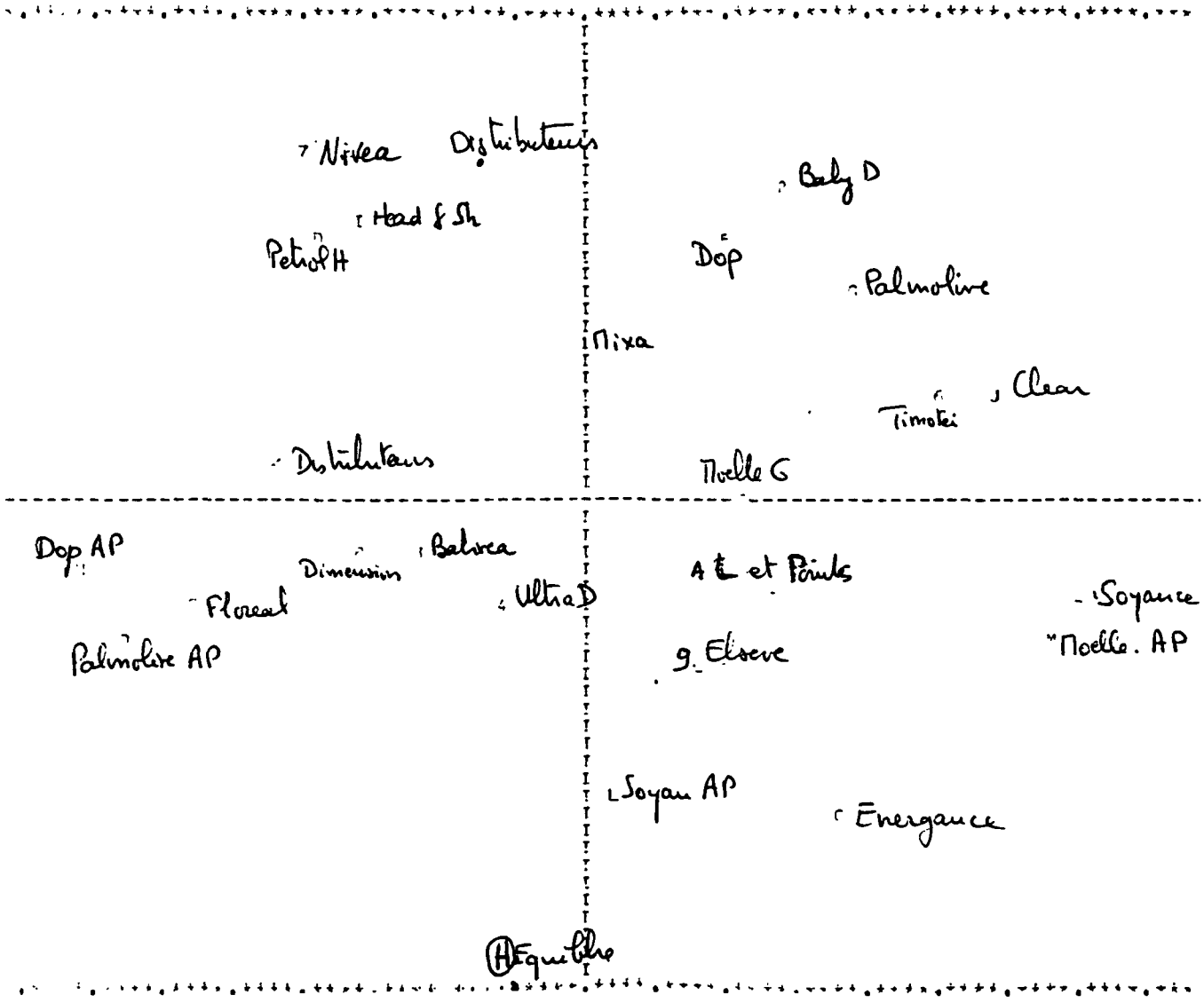
4) MARCHE SHAMP INDICE INDI2



4) MARCHE SHAMP INDICE MERA



MARCHE SHAMP INDICE MERR



ANNEXE 16 ETUDE DES FAMILLES D'INDICES : MATRICES DE CORRELATIONS ENTRE LES INDICES

1) MARCHE HUILE

A) MATRICE A : CORRELATIONS CALCULEES SUR L'UNIVERS GLOBAL

1									
.88	1								
.29	.38	1							
.08	.15	.04	1						
.10	.16	.00	.89	1					
.15	.25	.13	.34	.35	1				
.11	.18	.36	.02	.10	.12	1			
.06	.16	.37	.20	.28	.22	.73	1		
.04	.09	.28	.43	.46	.38	.14	.05	1	
.02	.05	.45	.24	.26	.02	.72	.41	.27	1
.02	.01	.40	.11	.20	.05	.37	.39	.24	.69	1

B) MATRICE B : MOYENNES DES CORRELATIONS INTRA-UNIVERS (U1=U2)

1									
.78	1								
.19	.22	1							
.10	.11	.14	1						
.10	.12	.13	.92	1					
.16	.18	.16	.32	.30	1				
.10	.14	.22	.24	.22	.12	1			
.11	.16	.25	.20	.24	.17	.44	1		
.06	.08	.55	.20	.20	.28	.14	.09	1	
.12	.15	.26	.20	.20	.11	.48	.26	.19	1
.09	.10	.28	.17	.17	.09	.30	.37	.16	.51	1

C) MATRICE C : MOYENNES DES CORRELATIONS INTER-UNIVERS (F1=F2 et U1 différent de U2)

.39									
.42	.41								
.17	.21	.66							
.12	.11	.14	.18						
.13	.12	.14	.19	.22					
.11	.14	.16	.11	.12	.23				
.13	.15	.23	.19	.21	.11	.29			
.14	.16	.24	.19	.24	.14	.25	.31		
.07	.09	.18	.13	.13	.14	.11	.13	.19	
.14	.14	.23	.19	.21	.11	.24	.21	.10	.25
.14	.14	.27	.22	.22	.08	.26	.21	.11	.27	.30

2) MARCHE BOISSON

A) MATRICE A : CORRELATIONS CALCULEES SUR L'UNIVERS GLOBAL

1									
.83	1								
.53	.61	1							
.14	.13	.19	1						
.14	.17	.21	.92	1					
.33	.44	.37	.18	.22	1				
.03	.04	.10	.20	.22	.08	1			
.20	.18	.27	.22	.18	.07	.73	1		
.44	.47	.69	.05	.06	.38	.31	.32	1	
.22	.17	.32	.16	.13	.07	.08	.08	.11	1
.21	.22	.36	.18	.09	.10	.58	.89	.31	.05	1

B) MATRICE B : MOYENNES DES CORRELATIONS INTRA-UNIVERS (U1=U2)

1									
.65	1								
.32	.48	1							
.11	.13	.14	1						
.11	.14	.14	.53	1					
.16	.25	.24	.14	.16	1				
.11	.13	.23	.15	.15	.11	1			
.17	.19	.29	.15	.13	.13	.55	1		
.19	.26	.65	.12	.10	.19	.15	.13	1	
.15	.14	.20	.13	.12	.13	.17	.21	.07	1
.17	.24	.33	.19	.20	.16	.42	.86	.15	.19	1

C) MATRICE C : MOYENNES DES CORRELATIONS INTER-UNIVERS (F1=F2 et U1 différent de U2)

.20									
.19	.21								
.20	.31	.54							
.09	.10	.10	.23						
.07	.09	.10	.19	.22					
.11	.13	.17	.09	.09	.16				
.11	.12	.21	.14	.12	.10	.13			
.13	.16	.24	.11	.12	.12	.20	.29		
.09	.11	.18	.11	.09	.06	.11	.10	.19	
.15	.13	.16	.15	.13	.11	.16	.16	.08	.36	
.14	.19	.30	.12	.11	.14	.22	.28	.09	.17	.29

3) MARCHE PARFUM

A) MATRICE A : CORRELATIONS CALCULEES SUR L'UNIVERS GLOBAL

1									
.53	1								
.39	.61	1							
.07	.02	.11	1						
.06	.20	.12	.02	1					
.43	.19	.42	.38	.14	1				
.42	.20	.42	.34	.12	.98	1			
.25	.24	.68	.12	.13	.36	.34	1		
.37	.20	.43	.36	.12	.97	.95	.37	1	
.33	.18	.43	.31	.11	.94	.96	.35	.97	1

B) MATRICE B : MOYENNES DES CORRELATIONS INTRA-UNIVERS (U1=U2)

1									
.65	1								
.32	.48	1							
.11	.13	.14	1						
.11	.14	.14	.53	1					
.16	.25	.24	.14	.16	1				
.11	.13	.23	.15	.15	.11	1			
.17	.19	.29	.15	.13	.13	.55	1		
.19	.26	.66	.12	.10	.19	.15	.13	1	
.15	.14	.20	.13	.12	.13	.17	.21	.07	1
.17	.24	.33	.19	.20	.16	.42	.86	.15	.19	1

C) MATRICE C : MOYENNES DES CORRELATIONS INTER-UNIVERS (F1=F2 et U1 différent de U2)

.15									
.22	.36								
.15	.37	.76							
.11	.27	.48	.52						
.07	.08	.09	.10	.10					
.13	.19	.28	.29	.10	.36				
.12	.18	.31	.29	.09	.35	.35			
.12	.27	.39	.31	.07	.18	.19	.49		
.13	.20	.30	.31	.10	.35	.37	.19	.38	
.14	.20	.31	.31	.10	.37	.39	.20	.40	.40

4) MARCHE SHAMP

A) MATRICE A : CORRELATIONS CALCULEES SUR L'UNIVERS GLOBAL

1	.	.	.	.	.	.
.94	1	.	.	.	.	.
.40	.47	1	.	.	.	.
.04	.06	.17	1	.	.	.
.04	.07	.19	.94	1	.	.
.20	.24	.31	.11	.12	1	.
.17	.16	.65	.17	.19	.10	1

B) MATRICE B : MOYENNES DES CORRELATIONS INTRA-UNIVERS (U1=U2)

1	.	.	.	.	.	.
.75	1	.	.	.	.	.
.20	.26	1	.	.	.	.
.07	.09	.18	1	.	.	.
.07	.08	.17	.93	1	.	.
.94	.15	.15	.11	.12	1	.
.09	.11	.60	.13	.12	.11	1

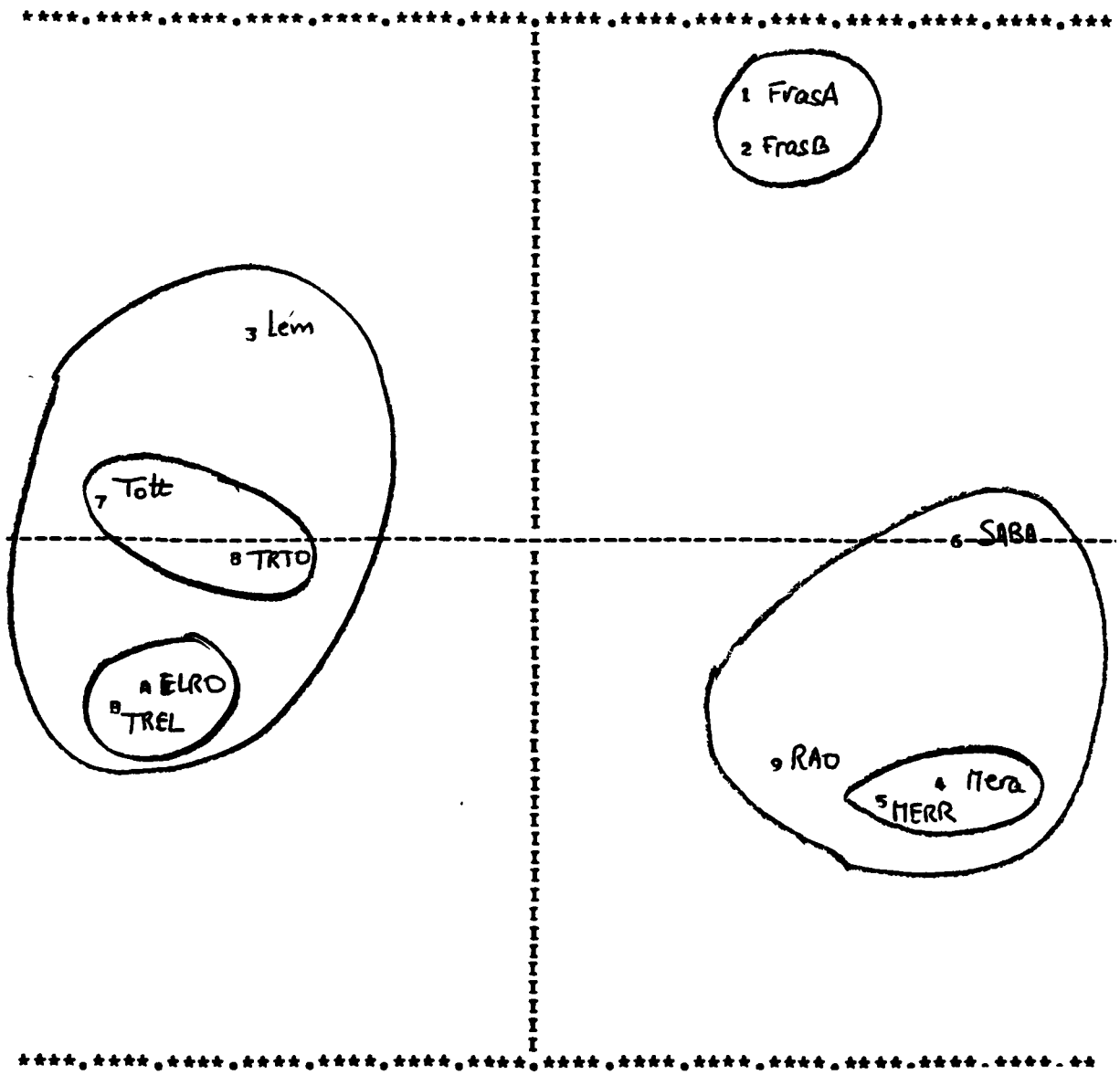
C) MATRICE C : MOYENNES DES CORRELATIONS INTER-UNIVERS (F1=F2 et U1 différent de U2)

.23	.	.	.	.	.	.
.21	.22	.	.	.	.	.
.13	.19	.43	.	.	.	.
.06	.07	.13	.12	.	.	.
.07	.08	.12	.11	.10	.	.
.08	.09	.14	.06	.06	.15	.
.08	.09	.18	.08	.07	.09	.23

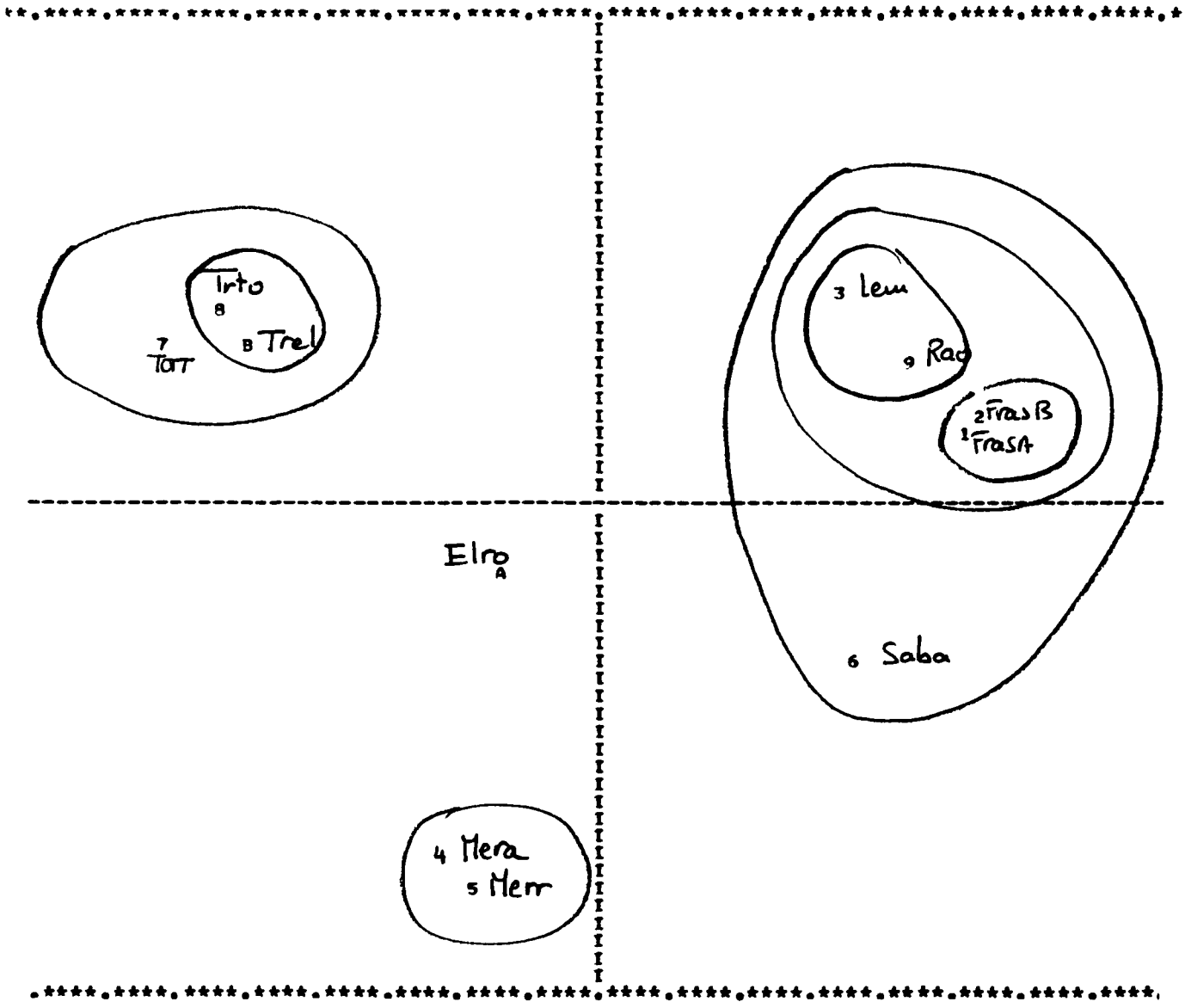
ANNEXE 17 ANALYSE DES FAMILLES D'INDICES
ANALYSE DES SIMILARITES AVEC LE MODELE INDSCALL
(modèle métrique, agrégé)

1) REPRESENTATIONS OBTENUES A PARTIR DE LA MATRICE A
(corrélations calculées sur l'univers "Global")

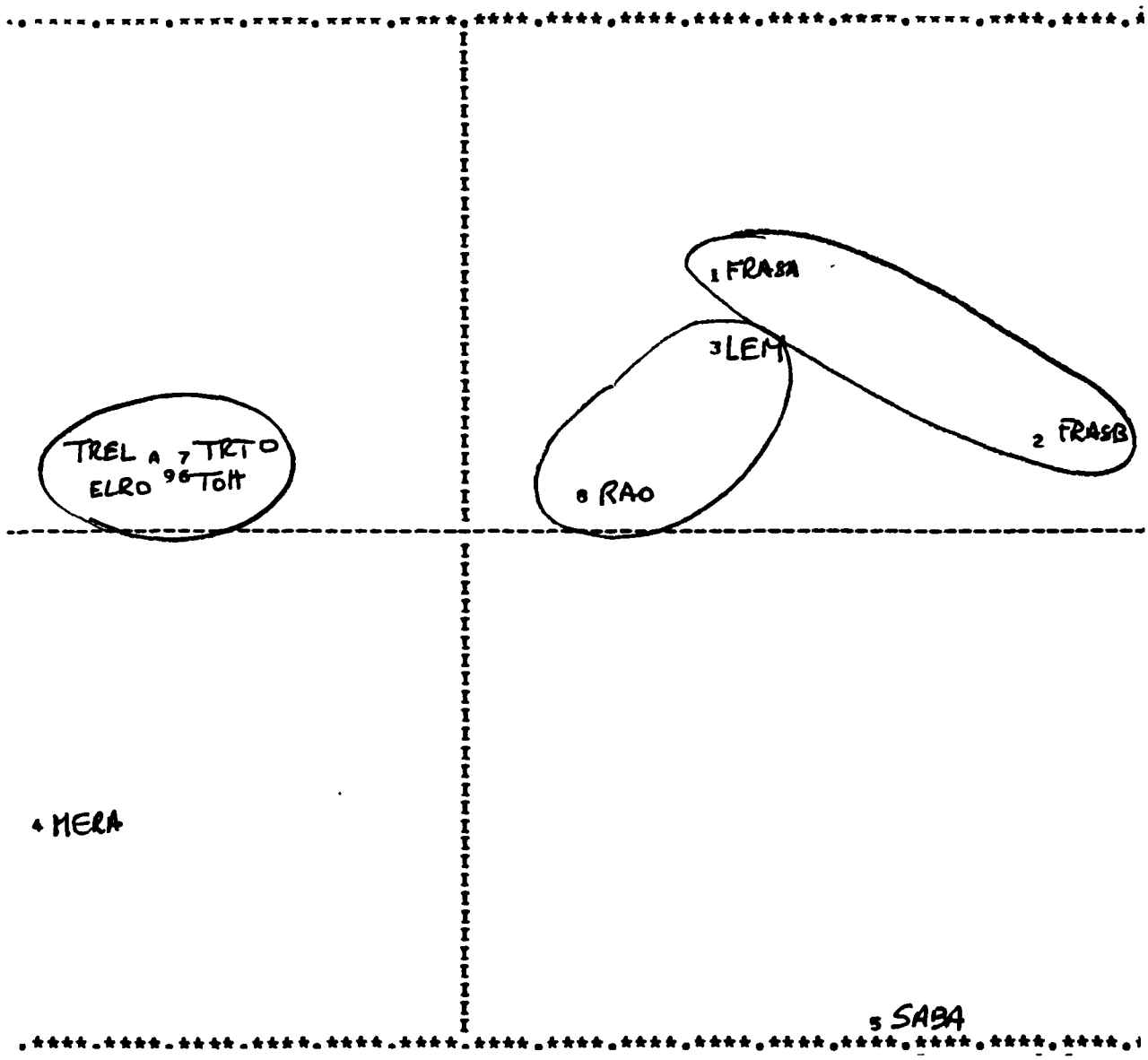
A) MARCHE HUILE



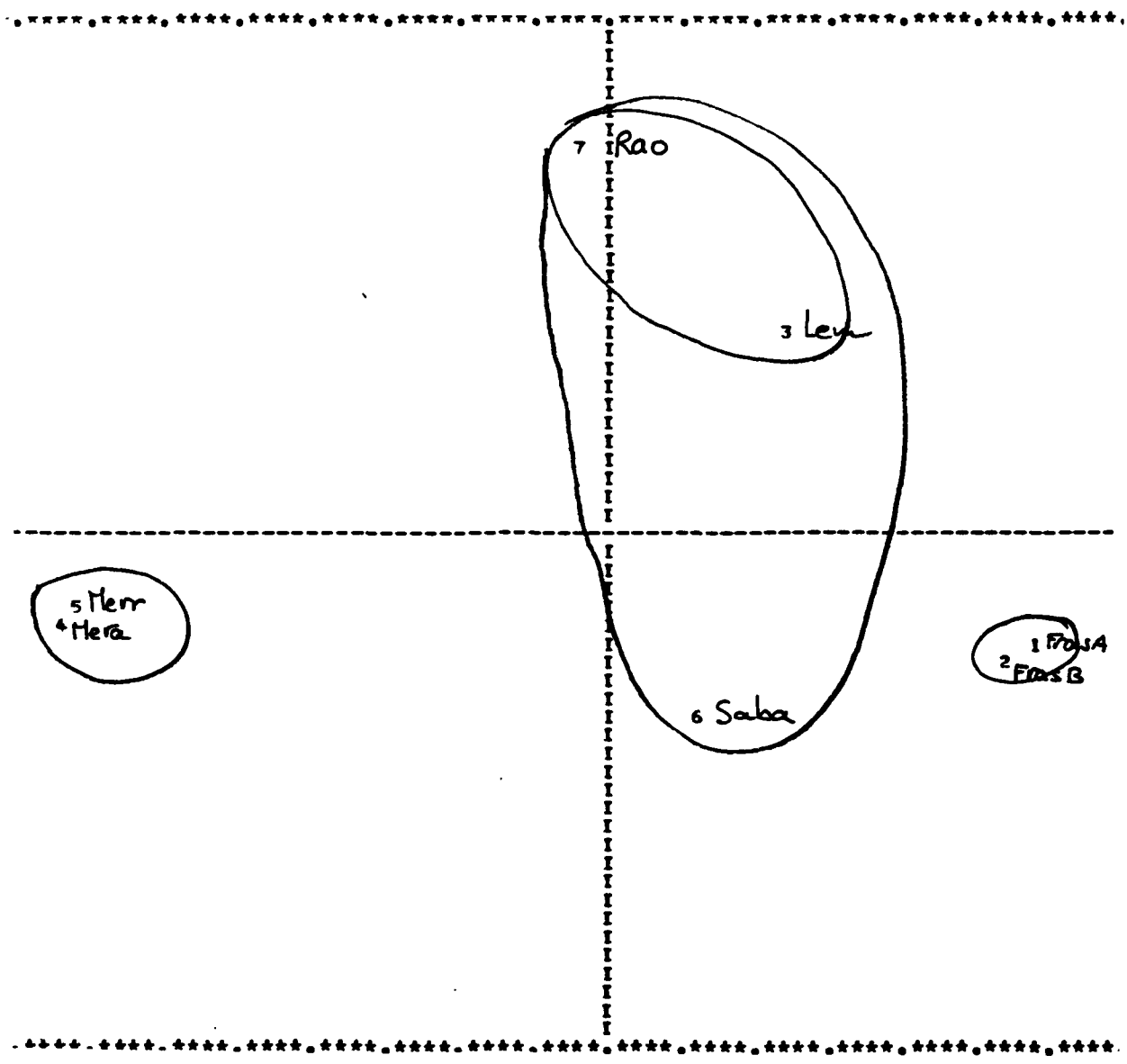
B) MARCHE BOISSON



C) MARCHE PARFUM

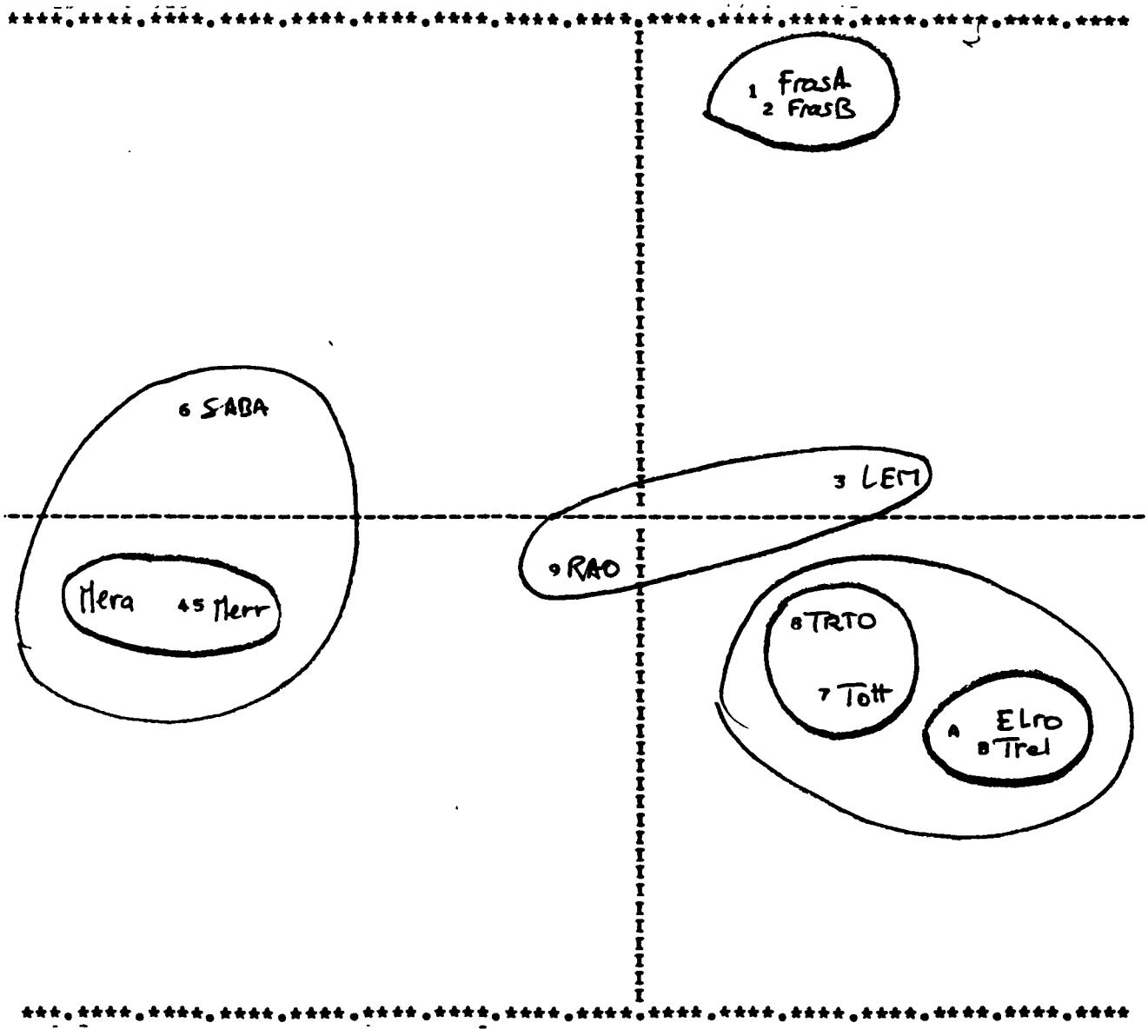


D) MARCHE SHAMP

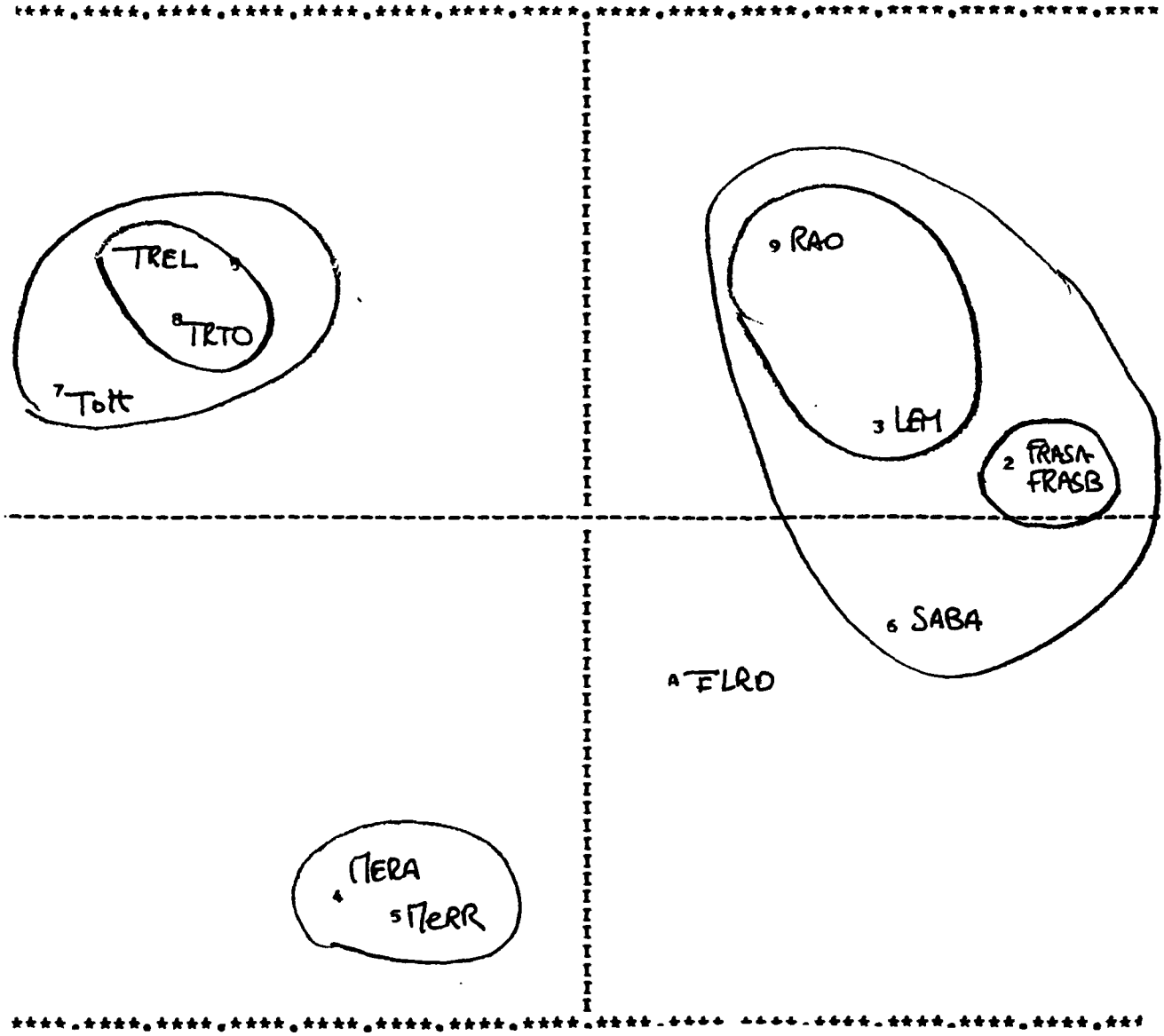


2) REPRESENTATIONS OBTENUES A PARTIR DE LA MATRICE B
(moyenne des corrélations "intra-univers")

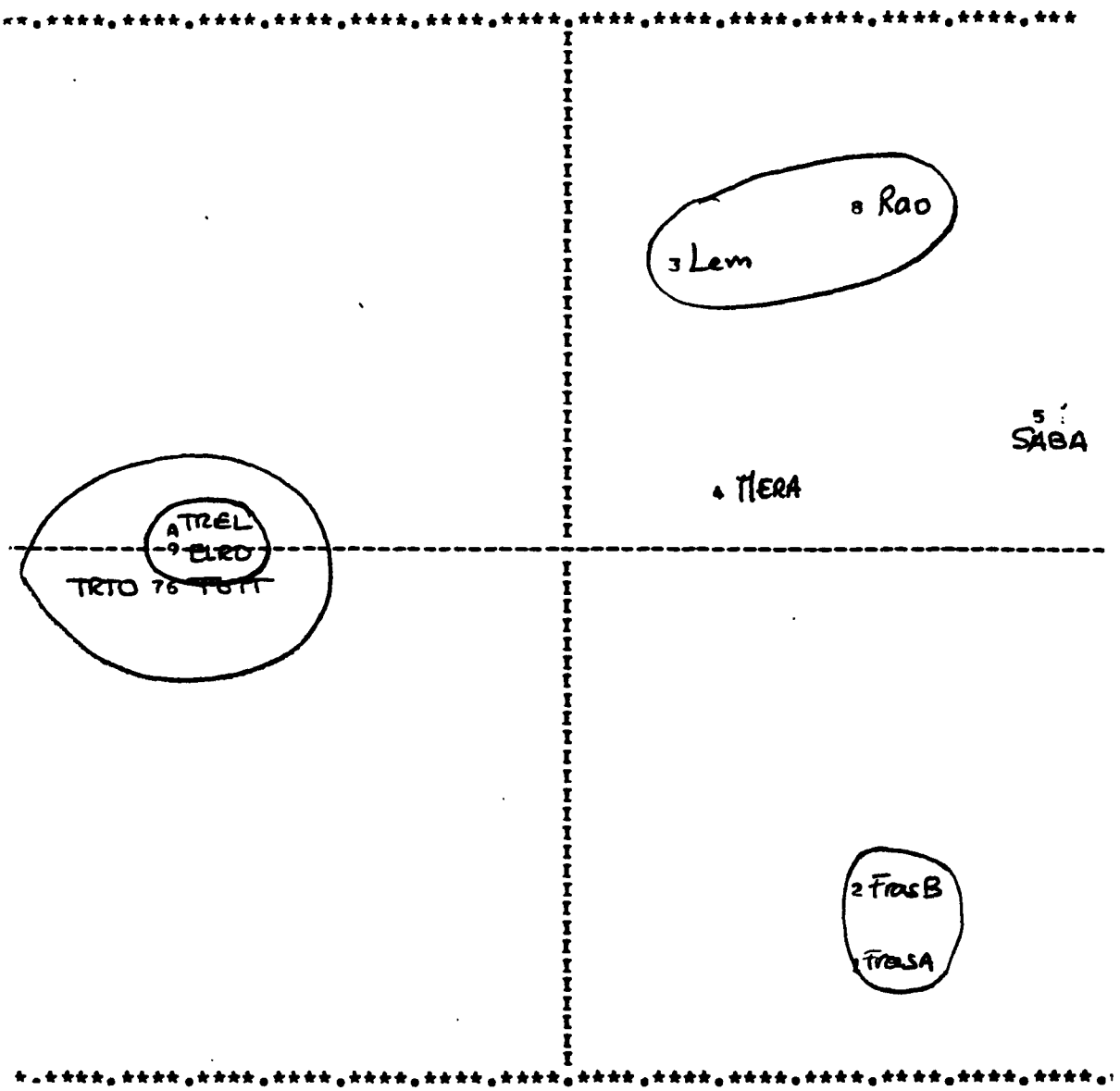
A) MARCHE HUILE



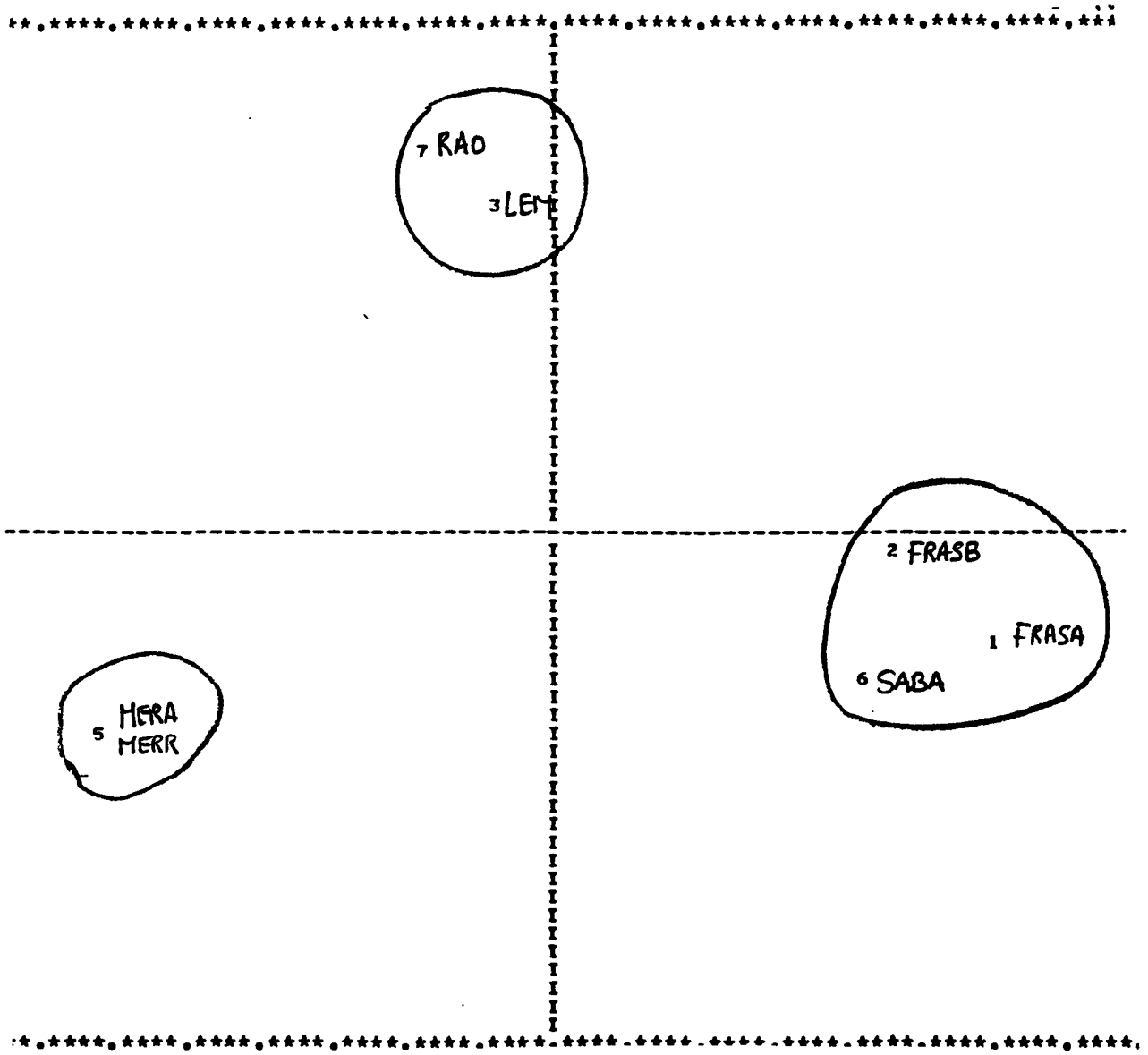
B) MARCHE BOISSON



C) MARCHE PARFUM

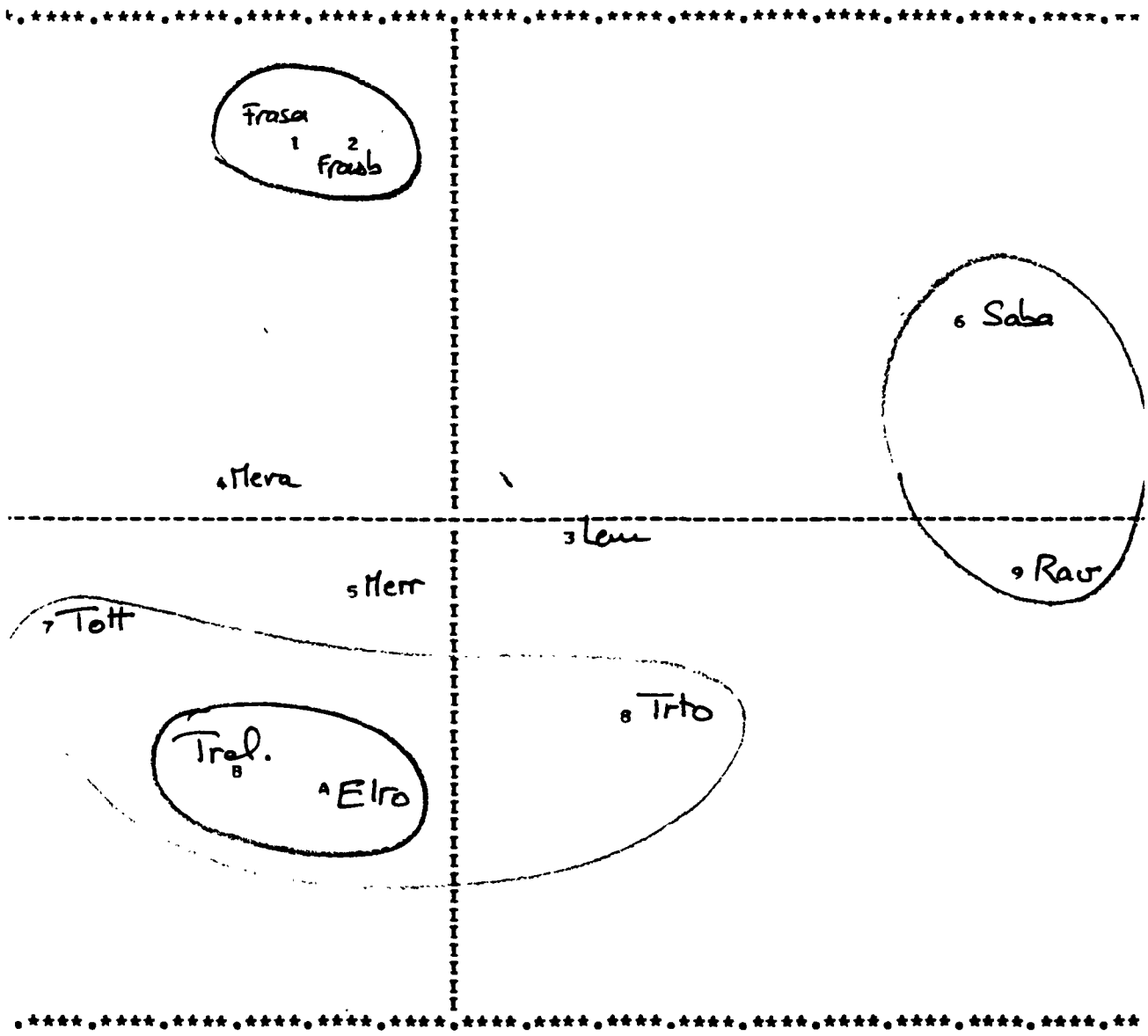


D) MARCHE SHAMP

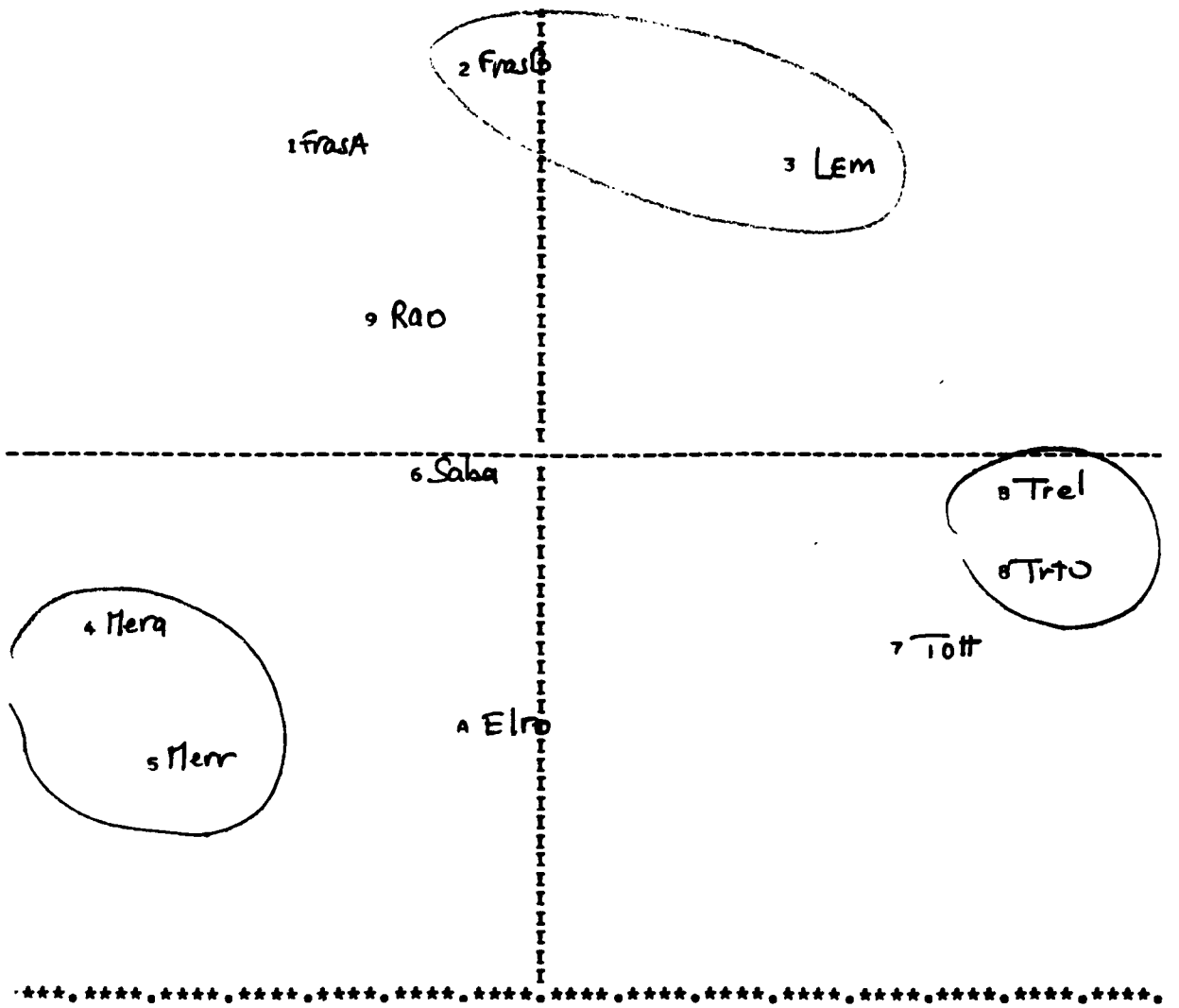


3) REPRESENTATIONS OBTENUES A PARTIR DE LA MATRICE C
(moyenne des corrélations "inter-univers")

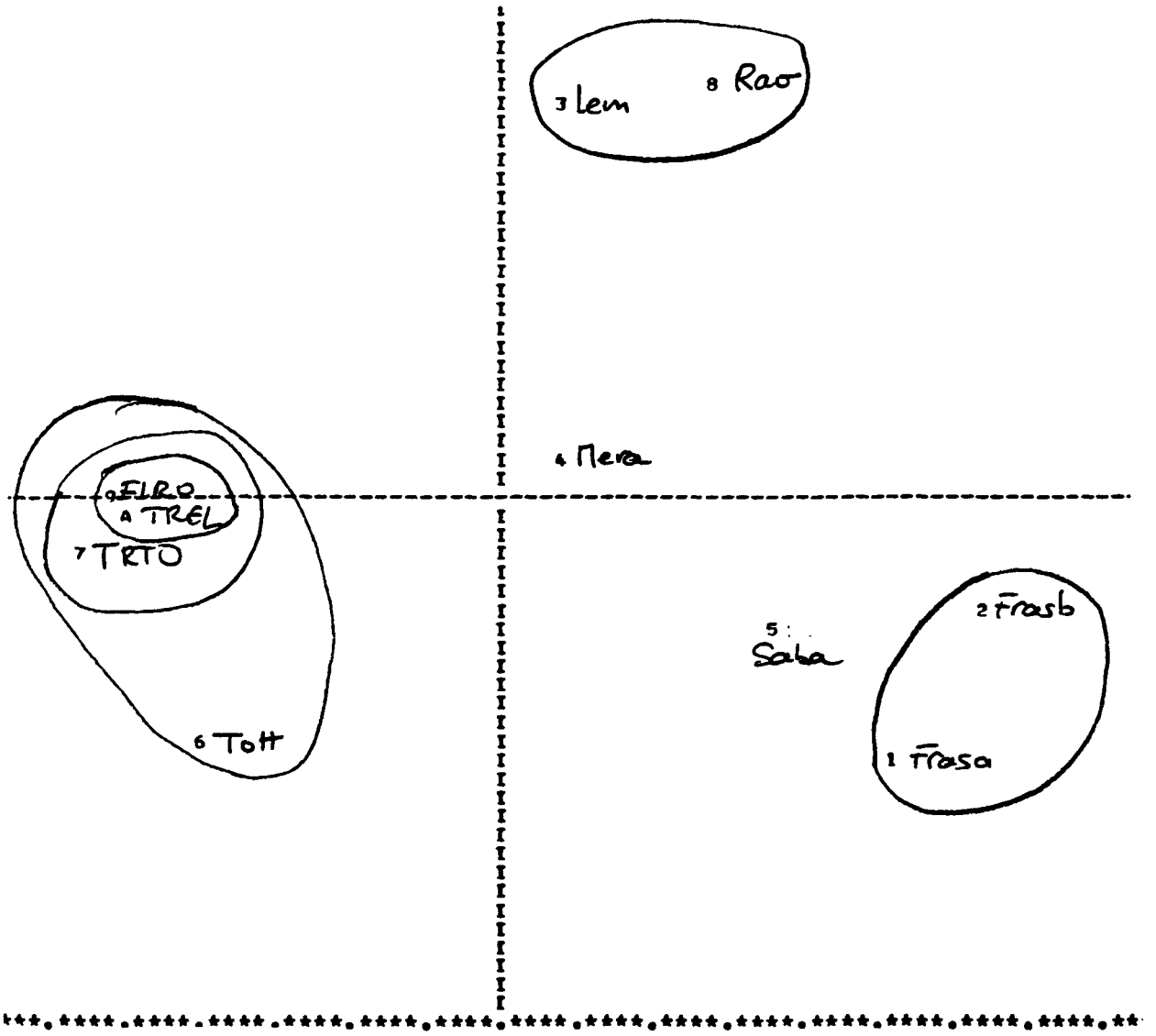
A) MARCHE HUILE



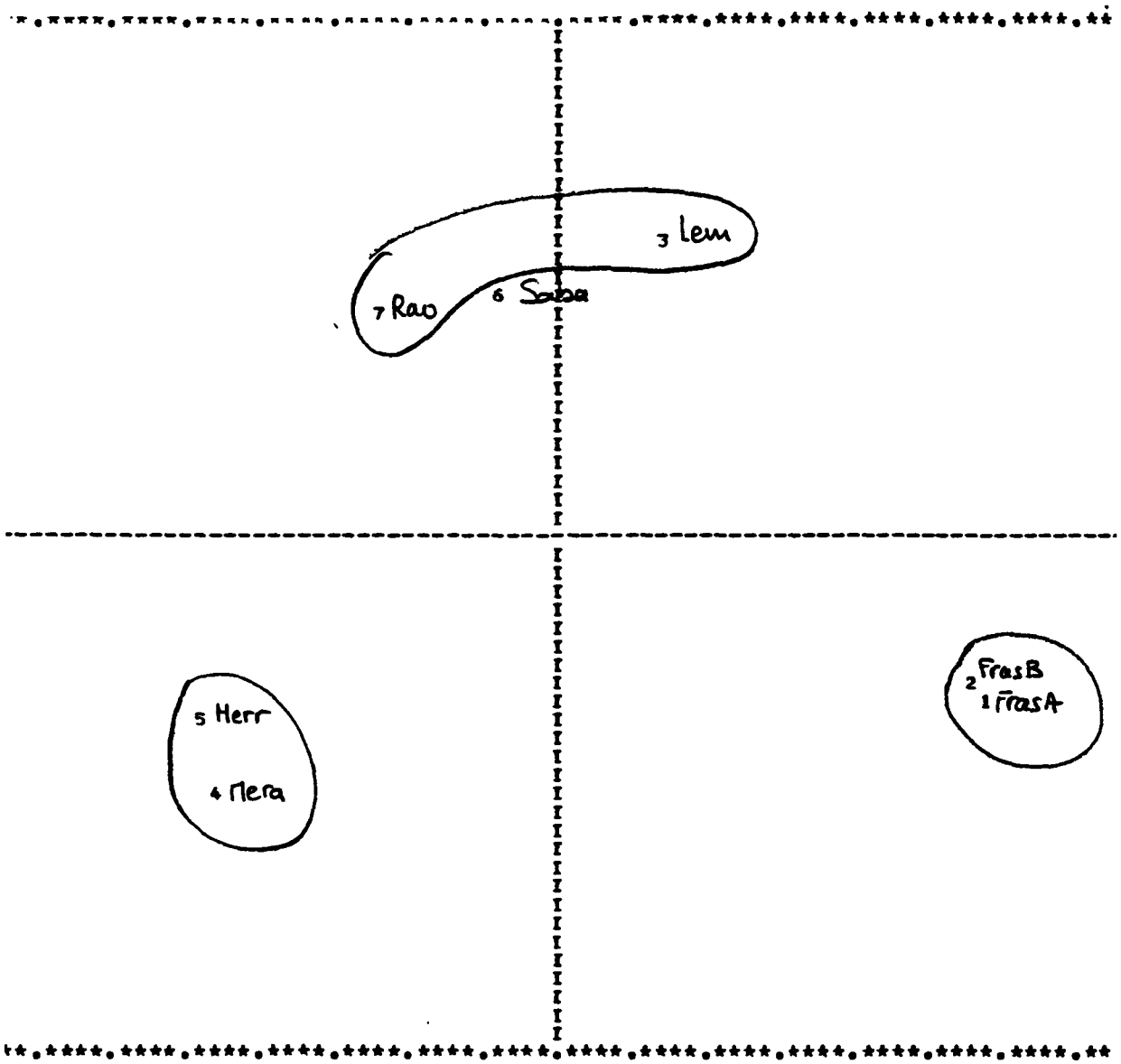
B) MARCHE BOISSON



C) MARCHE PARFUM



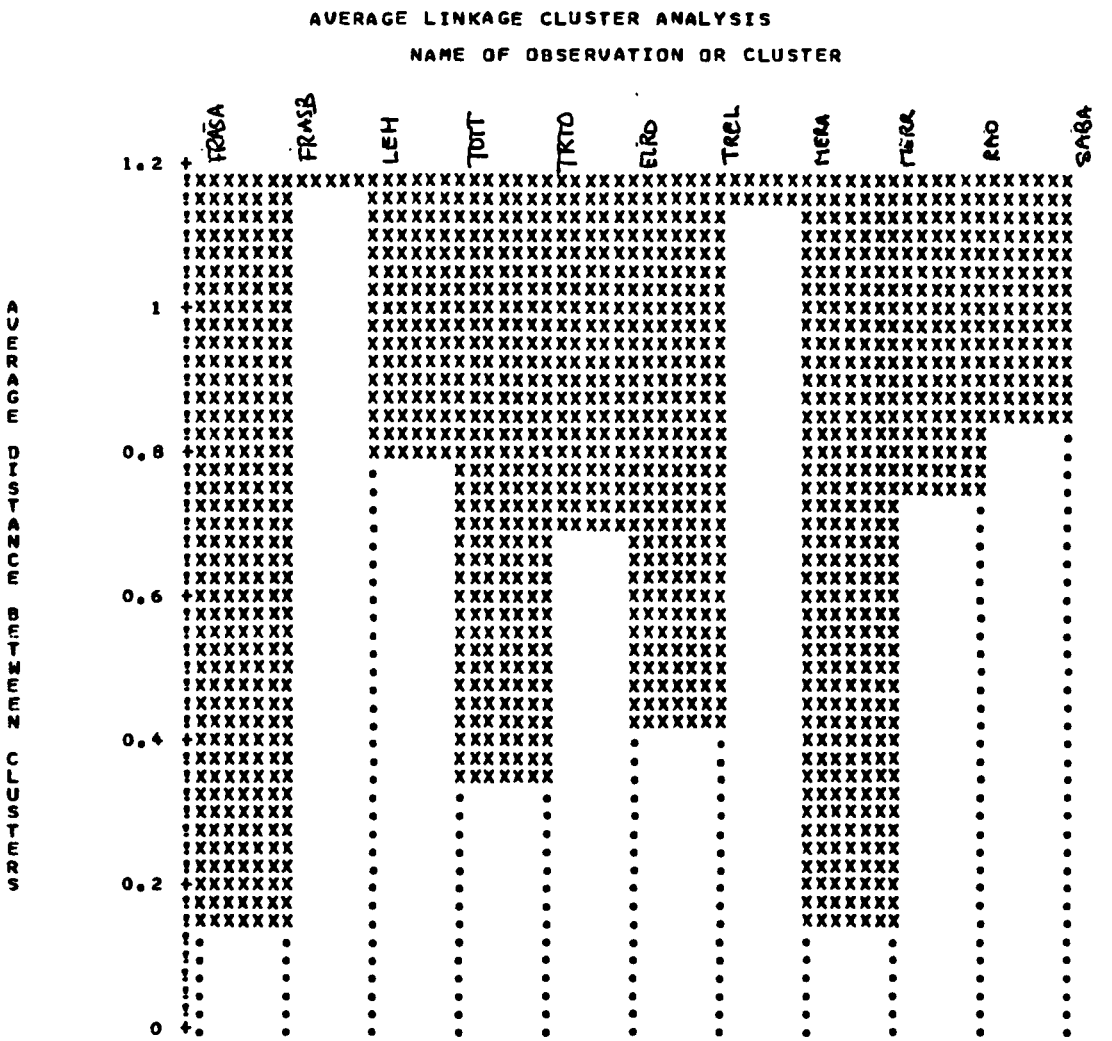
D) MARCHE SHAMP



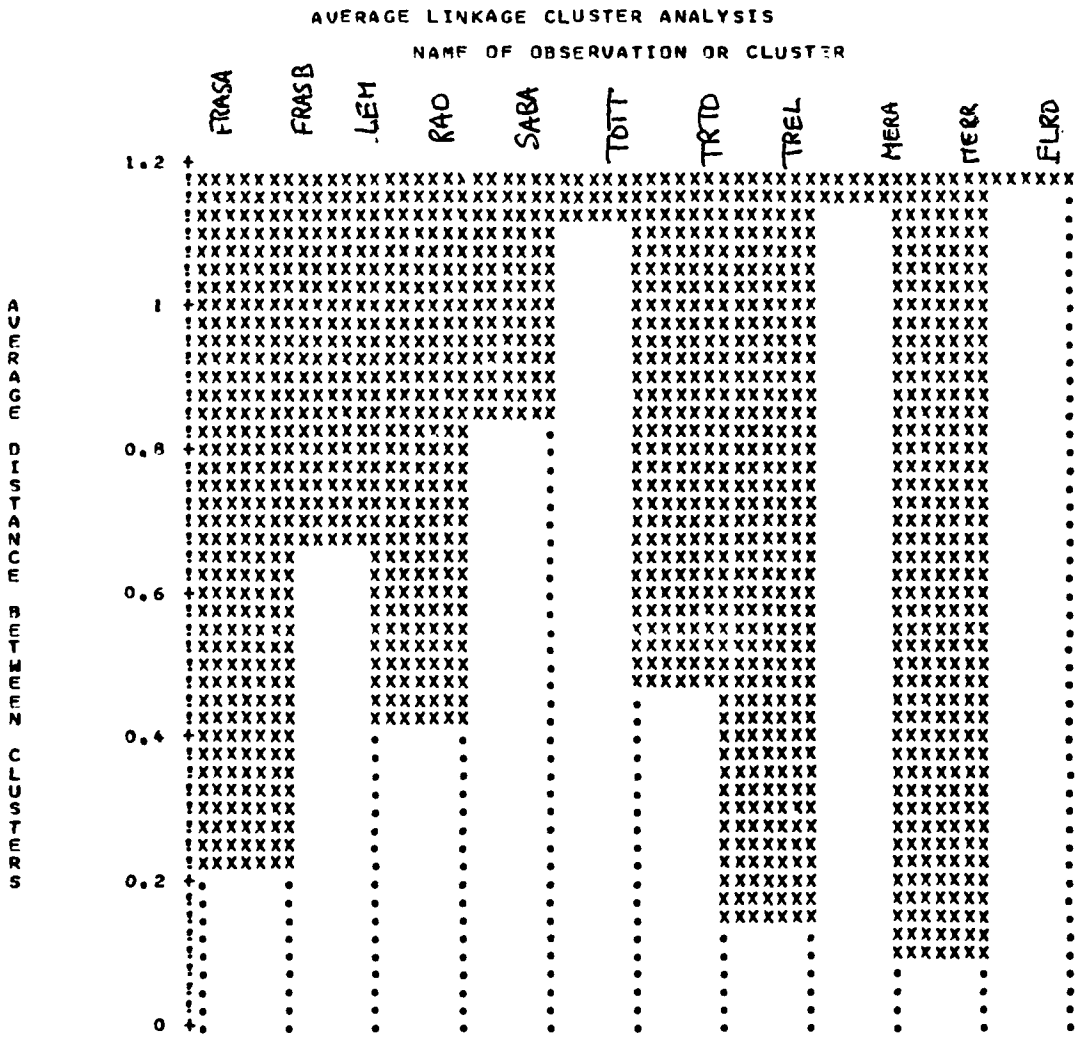
ANNEXE 18 ANALYSE DES FAMILLES D'INDICES
ANALYSE TYPOLOGIQUE HIERARCHIQUE
(distance moyenne)

1) ARBRES HIERARCHIQUE OBTENUS A PARTIR DE LA MATRICE A
(corrélations calculées sur l'univers "Global")

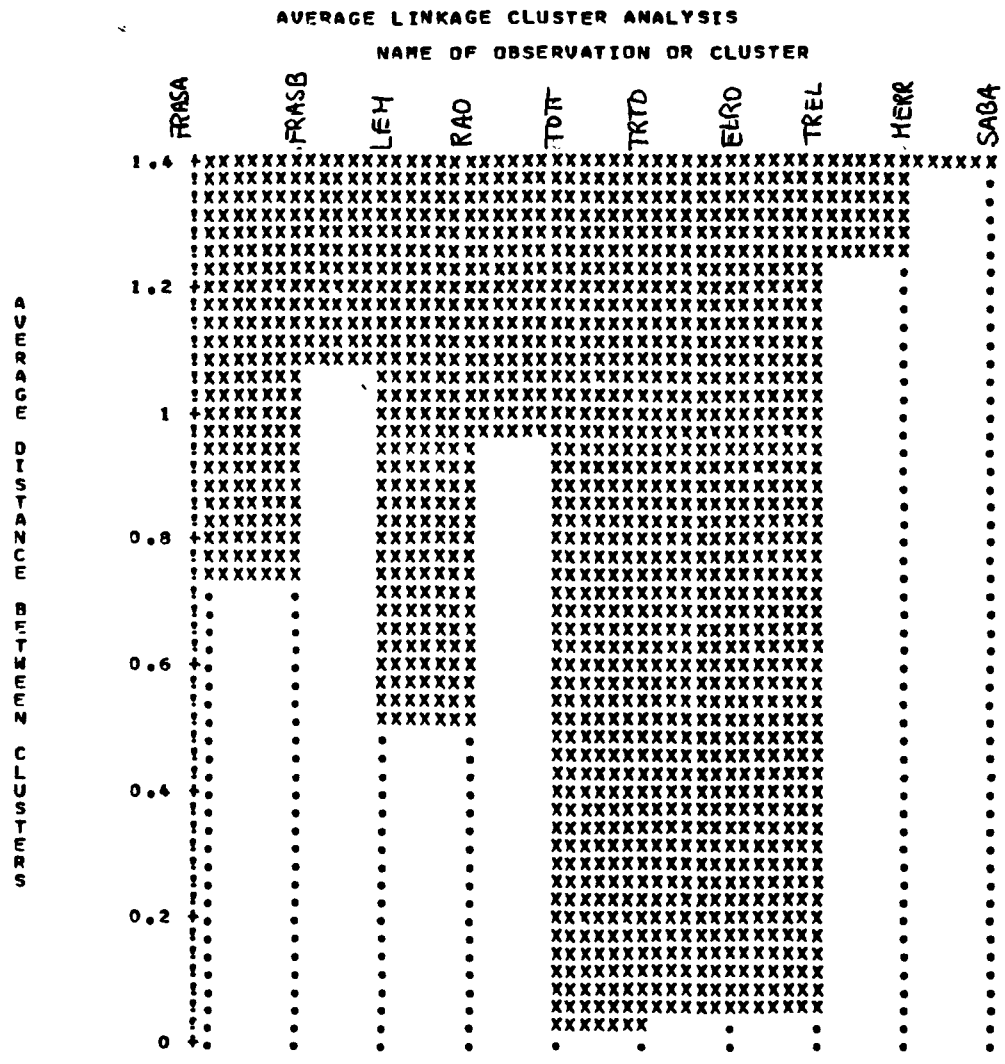
A) MARCHE HUILE



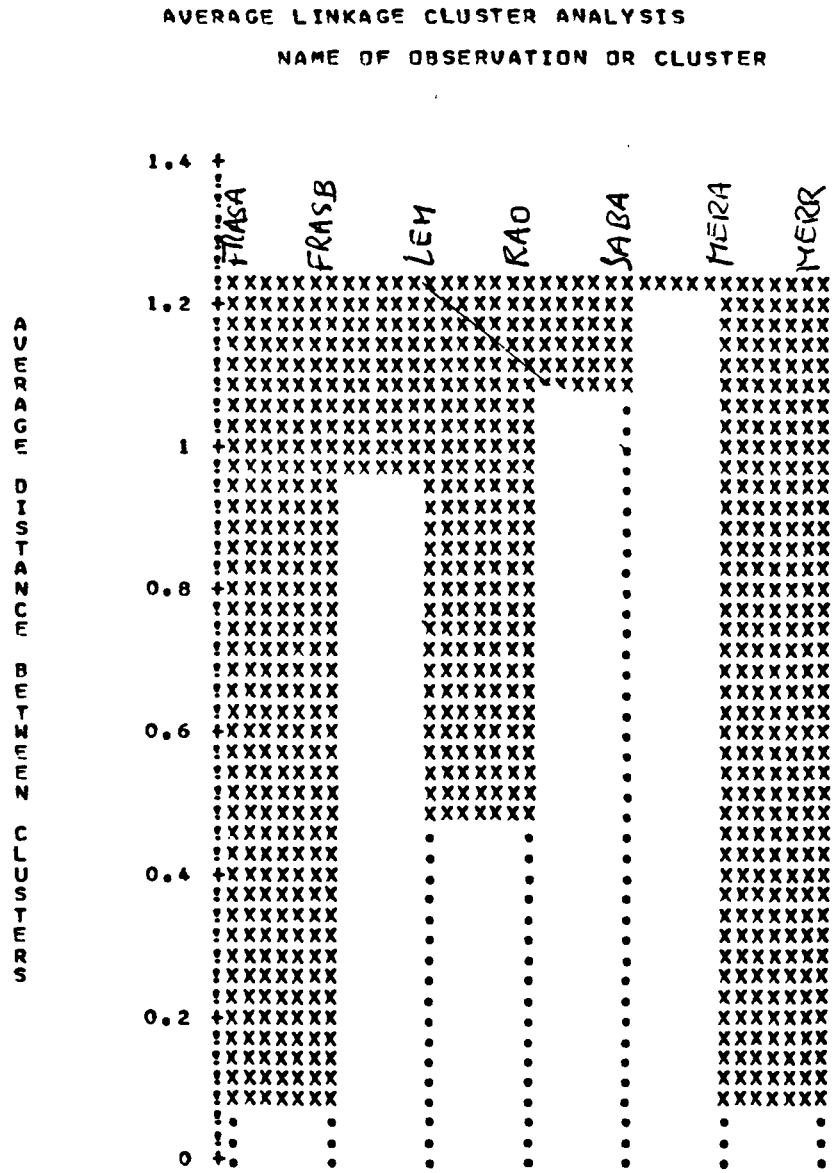
B) MARCHE BOISSON



C) MARCHE PARFUM

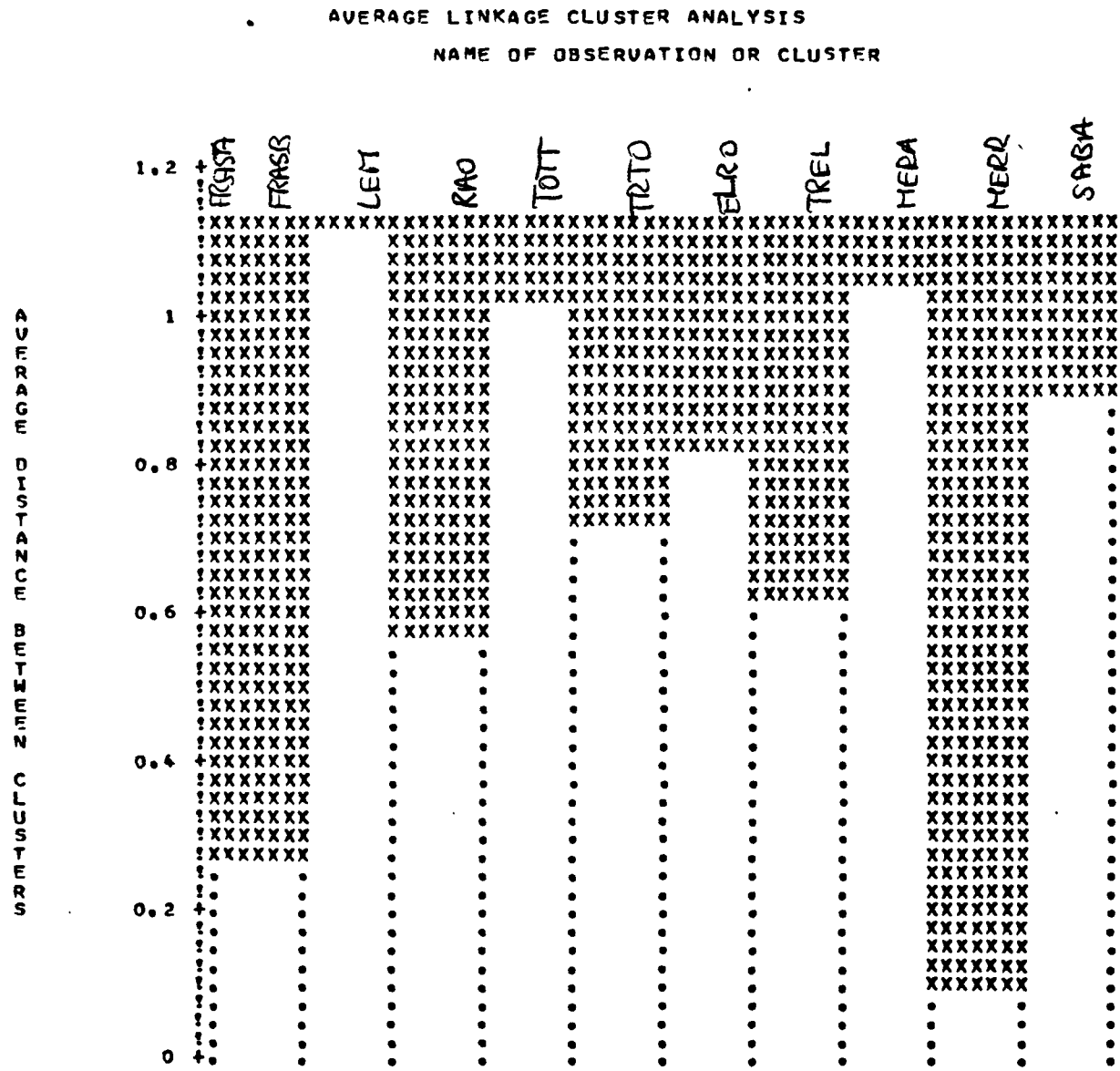


D) MARCHE SHAMP

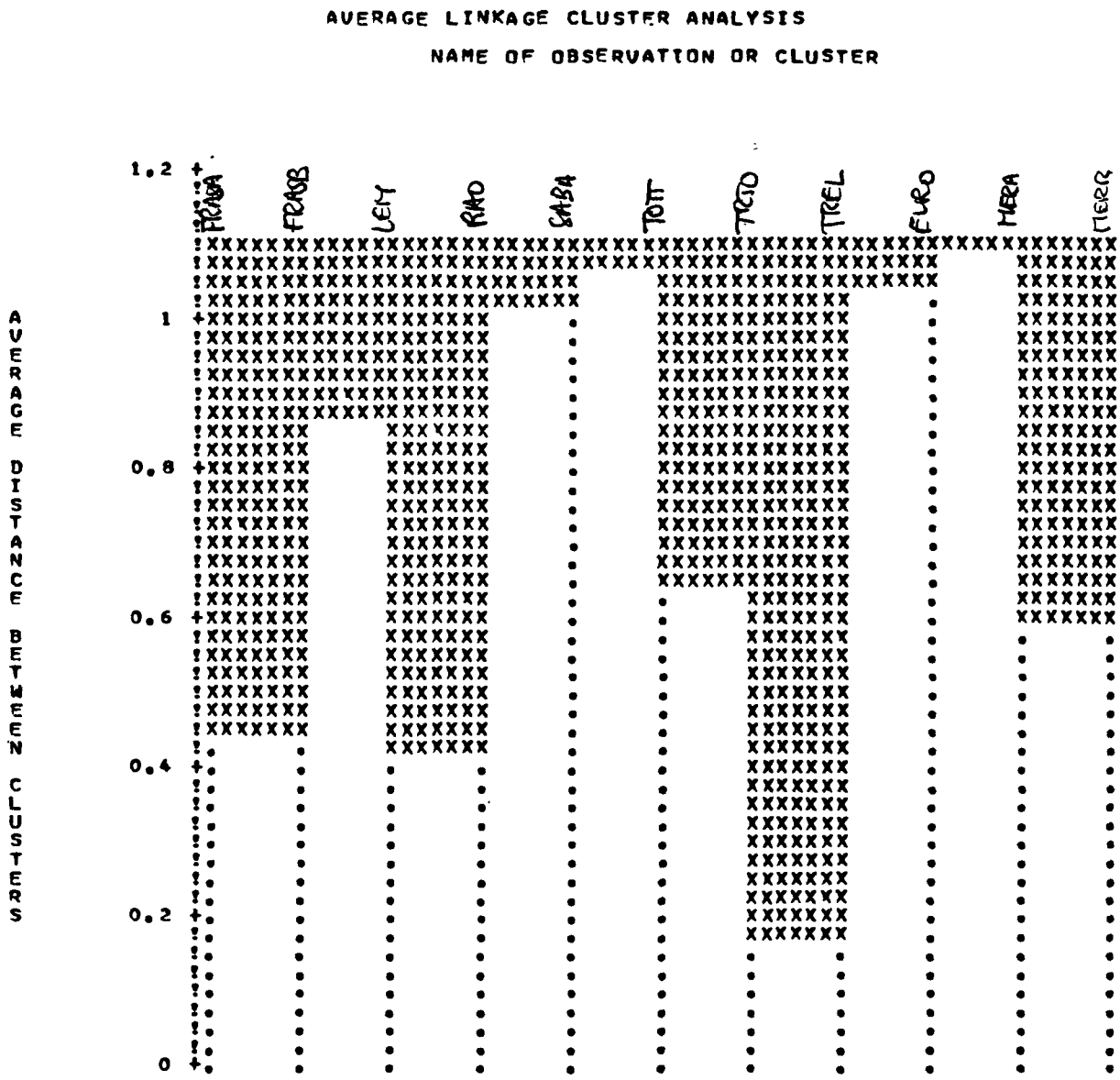


2) ARBRES HIERARCHIQUES OBTENUS A PARTIR DE LA MATRICE B (moyenne des corrélations "intra-univers")

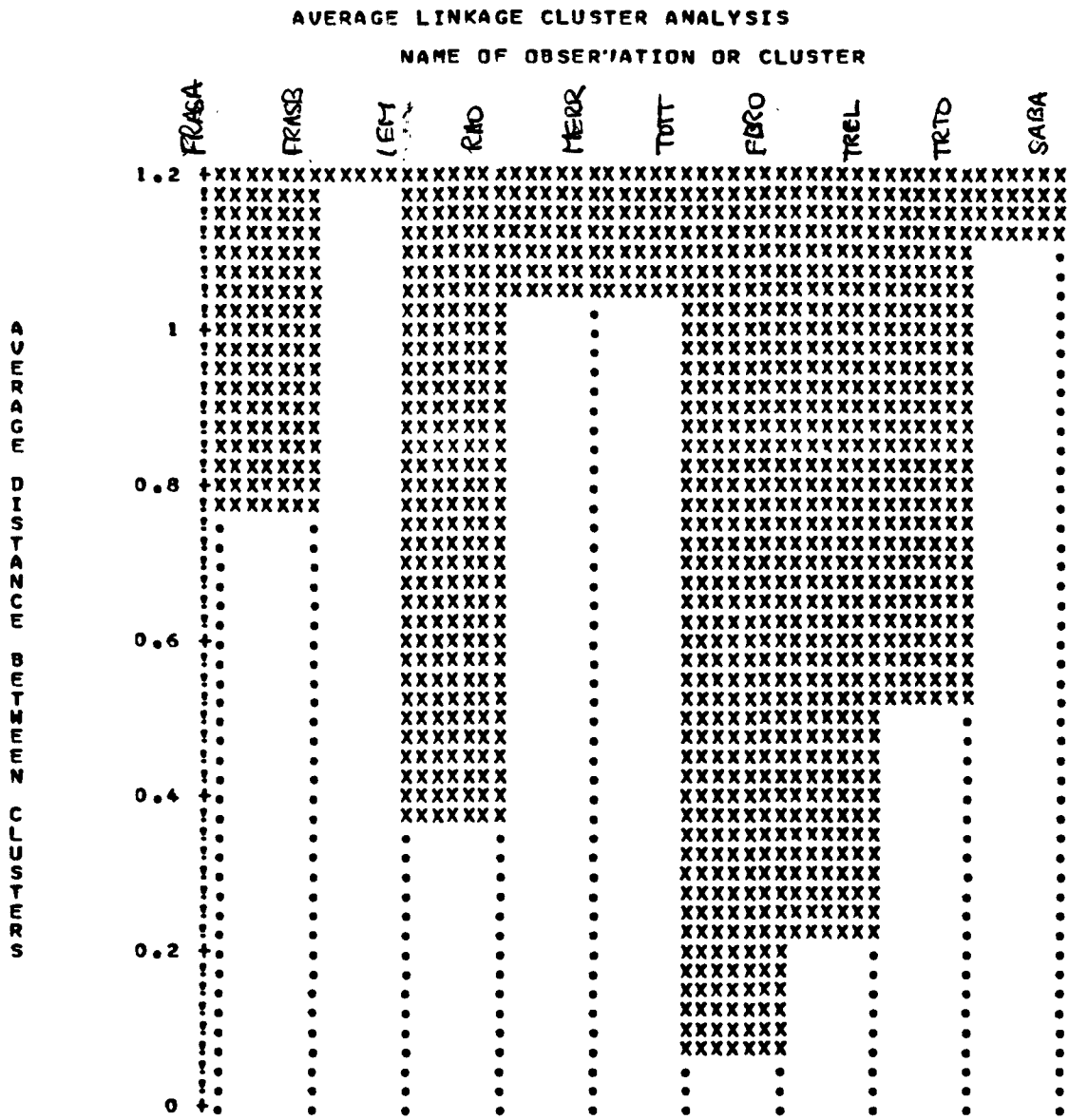
A) MARCHE HUILE



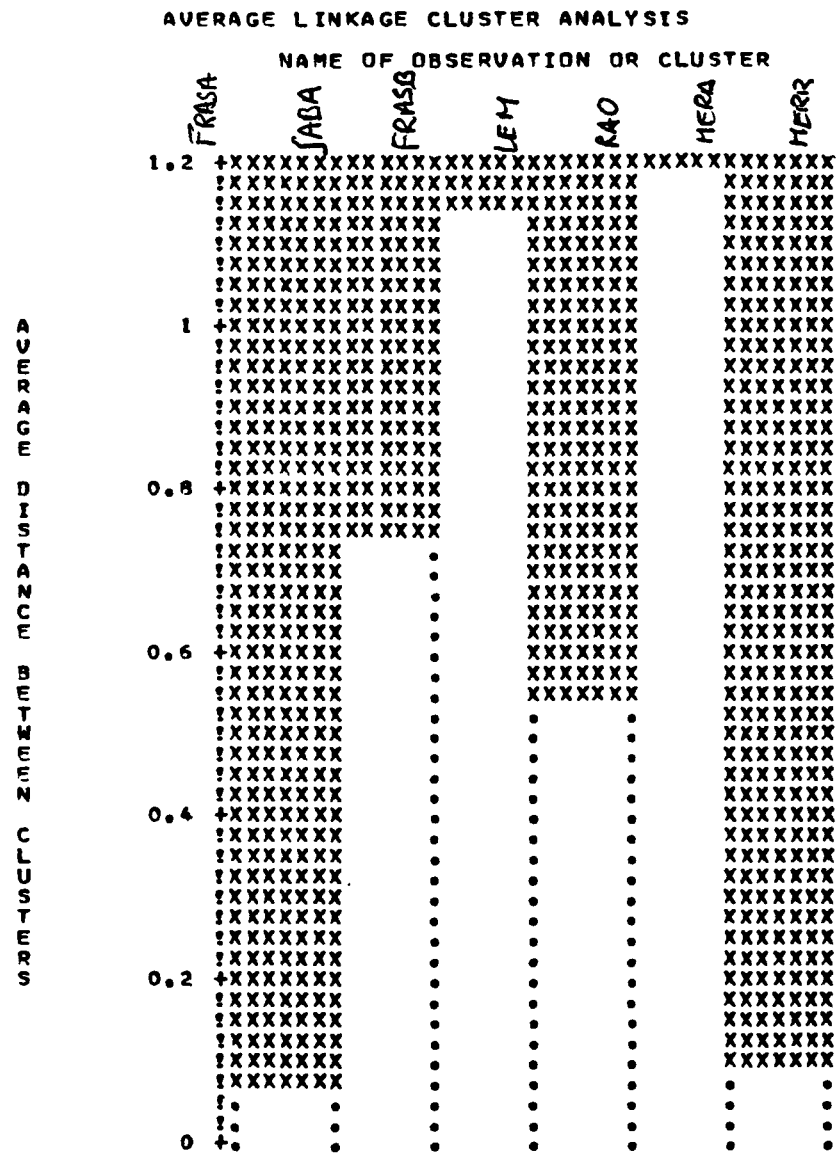
B) MARCHE BOISSON



C) MARCHE PARFUM



D) MARCHE SHAMP



2) ARBRES HIERARCHIQUES OBTENUS A PARTIR DE LA MATRICE C
(moyenne des corrélations "inter-univers")

A) MARCHE HUILE

AVERAGE LINKAGE CLUSTER ANALYSIS

MEAN DISTANCE BETWEEN OBSERVATIONS = 0.832

NUMBER OF CLUSTERS	CLUSTERS JOINED		FREQUENCY OF NEW CLUSTER	NORMALIZED AVERAGE DISTANCE
10	OB1 FRNSA	OB2 FRNSB	2	0.697115
9	OB10 ELRO	OB11 TREL	2	0.877404
8	OB7 TOTT	CL9	3	0.901442
7	OB3 LEM	CL8	4	0.909455
6	OB5 MERA	OB8 TRTO	2	0.913462
5	CL7	CL6	6	0.948017
4	CL5	OB4 MERA	7	0.977564
3	CL10	CL4	9	1.030220
2	OB6 SABA	OB9 TOTT	2	1.033654
1	CL3	CL2	11	1.059696

B) MARCHE BOISSON

AVERAGE LINKAGE CLUSTER ANALYSIS

MEAN DISTANCE BETWEEN OBSERVATIONS = 0.859091

NUMBER OF CLUSTERS	CLUSTERS JOINED		FREQUENCY OF NEW CLUSTER	NORMALIZED AVERAGE DISTANCE
10	OB2 FRASS	OB3 LEM	2	0.803175
9	OB8 TRTO	OB11 TREL	2	0.838095
8	CL10	CL9	4	0.905026
7	OB4 MERA	OB5 MERR	2	0.942857
6	CL8	OB7 TOTT	5	0.945767
5	CL6	OB10 EURO	6	0.982434
4	OB1 FRAS	CL5	7	0.985538
3	CL4	OB6 SABA	8	1.017687
2	CL3	CL7	10	1.038161
1	CL2	OB9 RAO	11	1.045291

C) MARCHE PARFUM

AVERAGE LINKAGE CLUSTER ANALYSIS

MEAN DISTANCE BETWEEN OBSERVATIONS = 0.833778

NUMBER OF CLUSTERS	CLUSTERS JOINED		FREQUENCY OF NEW CLUSTER	NORMALIZED AVERAGE DISTANCE
9	OB9 ELRO	OB10 TREL	2	0.743603
8	OB7 TRIO	CL9	3	0.773587
7	OB6 TOTT	CL8	4	0.787580
6	OB3 LEM	OB8 RAD	2	0.839552
5	OB4 MERR	CL7	5	0.947495
4	CL6	CL5	7	0.997868
3	OB1 FRASA	OB2 FRASB	2	1.007463
2	CL4	OB5 SABA	8	1.070857
1	CL3	CL2	10	1.077176

D) MARCHE SHAMP

AVERAGE LINKAGE CLUSTER ANALYSIS

MEAN DISTANCE BETWEEN OBSERVATIONS = 0.895714

NUMBER OF CLUSTERS	CLUSTERS JOINED		FREQUENCY OF NEW CLUSTER	NORMALIZED AVERAGE DISTANCE
6	OB1 FRASA	OB2 FRASB	2	0.881978
5	OB3 LEM	OB7 RAO	2	0.915470
4	CL6	CL5	4	0.979665
3	OB4 MERA	OB5 MERB	2	0.993620
2	CL4	OB6 SABA	5	1.004785
1	CL2	CL3	7	1.027113

ANNEXE 19 ANALYSE DES FAMILLES D'INDICES ANALYSE FACTORIELLE ANALYSE DE LA MATRICE
(estimation par l'analyse en composantes principales puis rotation des facteurs à l'aide de la méthode Promax)

1) ANALYSES FACTORIELLES A PARTIR DE LA MATRICE A
(corrélations calculées sur l'univers Global)

A) MARCHE HUILE

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	2	6	94 *
FRAB	8	14	95 *
LEM	60 *	1	42
MERA	4	91 *	0
MERR	11	91 *	0
SABA	4	57 *	25
TOTT	85 *	0	11
TRTO	72 *	16	10
RAO	20	66 *	2
ELRO	87 *	17	-7
TREL	74 *	13	-9

FINAL COMMUNALITY ESTIMATES: TOTAL = 7.526870

FRAA	FRAB	LEM	MERA	MERR	SABA	TOTT	TRTO
0.879182	0.929751	0.539687	0.828376	0.847531	0.387976	0.727715	0.553685

RAO	ELRO	TREL
0.471465	0.784627	0.576875

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	22	12
FACTOR2	22	100 *	10
FACTOR3	12	10	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	8	9	94 *
FRAB	15	17	96 *
LEM	62 *	8	45
MERA	14	91 *	6
MERR	21	92 *	6
SABA	11	58 *	29
TOTT	85 *	9	16
TRTO	74 *	23	15
RAO	27	67 *	7
ELRO	87 *	26	-1
TREL	75 *	20	-3

B) MARCHE BOISSON

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	83 *	4	3
FRAB	88 *	3	4
LEM	81 *	20	8
MERA	10	13	96 *
MERR	13	9	96 *
SABA	59 *	-2	18
TOTT	-1	84 *	15
TRTO	13	95 *	9
RAO	70 *	33	-10
ELRO	31	1	20
TREL	19	89 *	1

FINAL COMMUNALITY ESTIMATES: TOTAL = 7.675946

FRAA	FRAB	LEM	MERA	MERR	SABA	TOTT	TRTO
0.697884	0.777737	0.706164	0.942793	0.947645	0.376950	0.729998	0.919484

RAO	ELRO	TREL
0.612088	0.136424	0.828781

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	26	12
FACTOR2	26	100 *	15
FACTOR3	12	15	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	83 *	15	9
FRAB	88 *	15	10
LEM	83 *	31	15
MERA	16	21	97 *
MERR	18	17	97 *
SABA	59 *	7	22
TOTT	10	84 *	21
TRTO	25	96 *	17
RAO	73 *	41	-3
ELRO	32	6	22
TREL	30	91 *	9

C) MARCHE PARFUM

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	23	68 *	-11
FRAB	-6	83 *	14
LEM	22	85 *	6
MERR	50 *	-7	0
SABA	6	10	98 *
TOTT	94 *	28	5
TRTO	93 *	28	3
RAO	23	63 *	9
ELRO	94 *	28	4
TREL	93 *	26	4

FINAL COMMUNALITY ESTIMATES: TOTAL = 7.50768

FRAA	FRAB	LEM	MERR	SABA	TOTT	TRTO
0.519166	0.705065	0.774276	0.256112	0.983536	0.968230	0.954503

6

RAO	ELRO	TREL
0.453574	0.958831	0.934392

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	42	9
FACTOR2	42	100 *	17
FACTOR3	9	17	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	37	70 *	-3
FRAB	14	81 *	22
LEM	41	88 *	15
MERR	47	3	2
SABA	11	17	99 *
TOTT	98 *	46	11
TRTO	98 *	46	9
RAO	37	66 *	16
ELRO	98 *	46	10
TREL	97 *	44	9

D) MARCHE SHAMP

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	0.96005	-0.01574	0.09906
FRAB	0.96896	0.00890	0.13003
LEM	0.37758	0.08673	0.83117
MERA	0.03733	0.97933	0.08445
MERR	0.04022	0.97696	0.10829
SABA	0.33306	0.12457	0.25368
RAO	0.01053	0.08305	0.92235

FINAL COMMUNALITY ESTIMATES: TOTAL = 5.712508

FRAA	FRAB	LEM	MERA	MERR	SABA	RAO
0.931760	0.955867	0.840930	0.967622	0.967799	0.190801	0.857730

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	1.00000	0.05596	0.27345
FACTOR2	0.05596	1.00000	0.20433
FACTOR3	0.27345	0.20433	1.00000

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	0.96459	0.02580	0.23664
FRAB	0.97768	0.05362	0.27098
LEM	0.48253	0.17965	0.88203
MERA	0.05830	0.98360	0.19071
MERR	0.06422	0.98366	0.21434
SABA	0.36430	0.15963	0.31139
RAO	0.13029	0.17270	0.91749

2) ANALYSES FACTORIELLES A PARTIR DE LA MATRICE B
(moyenne des corrélations "intra-univers")

A) MARCHE HUILE

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FRAA	6	6	93 *	6
FRAB	11	7	93 *	8
LEM	28	0	15	80 *
MERA	17	94 *	2	4
MERR	18	94 *	2	3
SABA	0	47 *	18	36
TOTT	74 *	14	6	3
TRTO	66 *	16	10	3
RAO	6	16	-3	88 *
ELRO	76 *	5	5	14
TREL	73 *	1	1	16

FINAL COMMUNALITY ESTIMATES: TOTAL = 7.729146

FRAA	FRAB	LEM	MERA	MERR	SABA	TOTT	TRTO
0.882181	0.884423	0.731207	0.919148	0.915058	0.384371	0.571118	0.468328

RAO	ELRO	TREL
0.811695	0.599352	0.562265

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FACTOR1	100 *	25	17	26
FACTOR2	25	100 *	12	17
FACTOR3	17	12	100 *	17
FACTOR4	26	17	17	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3	FACTOR4
FRAA	14	11	94 *	14
FRAB	19	13	94 *	17
LEM	37	9	23	83 *
MERA	27	96 *	9	15
MERR	28	95 *	10	14
SABA	10	50 *	24	40
TOTT	75 *	25	13	15
TRTO	60 *	25	17	15
RAO	17	22	6	89 *
ELRO	77 *	17	13	26
TREL	74 *	13	8	26

B) MARCHE BOISSON

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	71 *	5	5
FRAB	79 *	9	7
LEM	77 *	26	4
MERA	7	11	85 *
MERR	8	10	86 *
SABA	39	8	25
TOTT	7	72 *	9
TRTO	11	94 *	3
RAO	66 *	8	3
ELRO	19	28	19
TREL	15	88 *	10

FINAL COMMUNALITY ESTIMATES: TOTAL = 6.332618

FRAA	FRAB	LEM	MERA	MERR	SABA	TOTT	TRTO
0.502235	0.644746	0.665613	0.738794	0.747631	0.222949	0.525792	0.890613

RAO	ELRO	TREL
0.445323	0.147761	0.801163

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	29	17
FACTOR2	29	100 *	19
FACTOR3	17	19	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	71 *	16	11
FRAB	80 *	21	14
LEM	80 *	37	13
MERA	16	19	86 *
MERR	16	18	86 *
SABA	42	16	29
TOTT	19	72 *	16
TRTO	25	94 *	13
RAO	67 *	18	9
ELRO	24	32	23
TREL	29	89 *	20

C) MARCHE PARFUM

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	8	4	84 *
FRAB	9	13	82 *
LEM	18	85 *	0
MERR	25	36	8
SABA	7	46 *	14
TOTT	97 *	11	7
TRTO	75 *	14	11
RAO	3	91 *	-2
ELRO	91 *	14	5
TREL	90 *	16	5

FINAL COMMUNALITY ESTIMATES: TOTAL = 6.67203

FRAA	FRAB	LEM	MERR	SABA	TOTT	TRTO
0.718202	0.690524	0.758289	0.199417	0.234031	0.951710	0.589334

RAO	ELRO	TREL
0.832904	0.855724	0.841901

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	26	18
FACTOR2	26	100 *	10
FACTOR3	18	10	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	15	7	85 *
FRAB	17	16	83 *
LEM	30	87 *	7
MERR	31	39	12
SABA	14	47	18
TOTT	98 *	23	17
TRTO	77 *	23	19
RAO	16	91 *	4
ELRO	92 *	25	15
TREL	92 *	27	14

D) MARCHE SHAMP

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	1.03300	0.01653	0.05134
FRAB	0.68213	0.01561	0.20429
LEM	0.16277	0.10414	0.87335
MERA	0.04893	0.97660	0.08812
MERR	0.05088	0.97810	0.07455
SABA	0.85122	0.08741	-0.01103
RAO	0.03088	0.05100	0.88940

FINAL COMMUNALITY ESTIMATES: TOTAL = 5.832986

FRAA	FRAB	LEM	MERA	MERR	SABA	RAO
1.069996	0.507277	0.800074	0.963905	0.964823	0.732331	0.794580

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	1.00000	0.10208	0.16796
FACTOR2	0.10208	1.00000	0.16267
FACTOR3	0.16796	0.16267	1.00000

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	1.03364	0.07301	0.14659
FRAB	0.69535	0.06631	0.26630
LEM	0.23145	0.18075	0.89009
MERA	0.10000	0.98173	0.17025
MERR	0.10101	0.98225	0.15707
SABA	0.85118	0.12942	0.07376
RAO	0.09882	0.12241	0.88971

3) ANLYSES FACTORIELLES A PARTIR DE LA MATRICE C
(moyenne des corrélations "inter-univers")

A) MARCHE HUILE

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	12	82 *	2
FRAB	11	82 *	10
LEM	43 *	24	30
MERA	50 *	1	14
MERR	56 *	3	13
SABA	5	15	71 *
TOTT	61 *	8	3
TRTO	53 *	11	20
RAO	13	-5	70 *
ELRO	61 *	10	-1
TREL	66 *	8	-5

FINAL COMMUNALITY ESTIMATES: TOTAL = 4.921257

FRAA	FRAB	LEM	MERA	MERR	SABA	TOTT	TRTO
0.692621	0.695611	0.333377	0.269224	0.303593	0.523029	0.381225	0.330370

RAO	ELRO	TREL
0.566481	0.381648	0.444078

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	28	26
FACTOR2	28	100 *	18
FACTOR3	26	18	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	23	83 *	10
FRAB	23	83 *	18
LEM	49 *	32	38
MERA	51 *	10	20
MERR	54 *	12	20
SABA	15	21	72 *
TOTT	62 *	18	12
TRTO	56 *	20	28
RAO	21	3	74 *
ELRO	61 *	19	8
TREL	65 *	18	5

B) MARCHE BOISSON

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	6	59 *	5
FRAB	17	66 *	-3
LEM	41 *	59 *	-3
MERA	8	8	72 *
MERR	12	1	69 *
SABA	22	30	10
TOTT	57 *	3	23
TRTO	66 *	11	5
RAO	-17	53 *	33
ELRO	35	13	35
TREL	68 *	19	0

FINAL COMMUNALITY ESTIMATES: TOTAL = 4.527494

FRAA	FRAB	LEM	MERA	MERR	SABA	TOTT	TRTO
0.356822	0.462147	0.513557	0.526623	0.493082	0.151275	0.377087	0.456075

RAO	ELRO	TREL
0.422231	0.263152	0.505442

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	30	11
FACTOR2	30	100 *	16
FACTOR3	11	16	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	16	60 *	9
FRAB	27	67 *	2
LEM	49 *	63 *	3
MERA	13	15	72 *
MERR	16	9	70 *
SABA	27	34	13
TOTT	58 *	13	25
TRTO	68 *	20	9
RAO	-7	53 *	36
ELRO	39	21	38
TREL	71 *	28	5

C) MARCHE PARFUM

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	5	-5	77 *
FRAB	5	14	70 *
LEM	27	72 *	5
MERR	45 *	18	5
SABA	19	6	31
TOTT	68 *	3	14
TRTO	69 *	8	4
RAO	3	84 *	10
ELRO	69 *	8	12
TREL	71 *	9	10

FINAL COMMUNALITY ESTIMATES: TOTAL = 4.75714

FRAA	FRAB	LEM	MERR	SABA	TOTT	TRTO
0.598339	0.510765	0.590737	0.233488	0.138611	0.476702	0.478956

RAO	ELRO	TREL
0.712754	0.498751	0.518044

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	100 *	26	20
FACTOR2	26	100 *	14
FACTOR3	20	14	100 *

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	14	1	77 *
FRAB	15	20	71 *
LEM	36	75 *	10
MERR	47 *	25	10
SABA	24	11	33
TOTT	69 *	14	19
TRTO	69 *	18	10
RAO	14	84 *	14
ELRO	71 *	19	17
TREL	72 *	19	16

D) MARCHE SHAMP

ROTATED FACTOR PATTERN

	FACTOR1	FACTOR2	FACTOR3
FRAA	0.77854	0.01236	0.01524
FRAB	0.74243	0.11884	0.05437
LEM	0.28197	0.53378	0.25643
MERA	-0.02596	0.11476	0.71851
MERR	0.09799	-0.01814	0.73383
SABA	0.04202	0.67914	-0.09243
RAO	-0.01733	0.68169	0.08863

FINAL COMMUNALITY ESTIMATES: TOTAL = 3.627910

FRAA	FRAB	LEM	MERA	MERR	SABA	RAO
0.606514	0.566288	0.430186	0.530095	0.548435	0.471536	0.472856

INTER-FACTOR CORRELATIONS

	FACTOR1	FACTOR2	FACTOR3
FACTOR1	1.00000	0.19480	0.16051
FACTOR2	0.19480	1.00000	0.17608
FACTOR3	0.16051	0.17608	1.00000

FACTOR STRUCTURE (CORRELATIONS)

	FACTOR1	FACTOR2	FACTOR3
FRAA	0.77510	0.08665	0.07785
FRAB	0.75240	0.19214	0.12313
LEM	0.35022	0.57729	0.32350
MERA	0.03848	0.17086	0.72146
MERR	0.14986	0.05192	0.73490
SABA	0.10020	0.67010	-0.02925
RAO	0.05494	0.68203	0.14606

ANNEXE 20 EFFET DE LA MANIPULATION DES FACTEURS SUR LES INDICES

α MARCHE HUILE

ANALYSE DE VARIANCE ACORR = f(INDICE)
F = 7.5 Prob = 0.0001

INDICE	MOYENNE bi..	ECART TYPE	T-TEST à 5%
FRASA	.40	.24	B
FRASB	.41	.26	B
LEM	.66	.25	A
MERA	.18	.20	C
MERR	.22	.17	C
SABA	.23	.15	C
TOTT	.29	.17	BC
TRTO	.31	.19	BC
RAO	.19	.12	C
ELRO	.25	.17	C
TREL	.30	.21	BC
MOYE (b...)	.31	.23	

β MARCHE DES BOISSON

ANALYSE DE VARIANCE ACORR = f(INDICE)
F = 8.3 Prob = 0.0001

INDICE	MOYENNE bi..	ECART TYPE	T-TEST à 5%
FRASA	.20	.18	CD
FRASB	.21	.16	CD
LEM	.54	.13	A
MERA	.23	.23	CD
MERR	.22	.11	CD
SABA	.16	.09	D
TOTT	.13	.13	BC
TRTO	.29	.18	CD
RAO	.19	.13	CD
ELRO	.36	.11	B
TREL	.29	.21	BC
MOYE (b...)	.26	.19	

Γ MARCHE DES PARFUM

ANALYSE DE VARIANCE ACORR = f(INDICE)
F = 7.5 Prob = 0.0001

INDICE	MOYENNE bi..	ECART TYPE	T-TEST à 5%
FRASA	.14	.12	DE
FRASB	.20	.20	CD
LEM	.58	.19	A
MERR	.29	.09	BC
SABA	.10	.07	E
TOTT	.32	.22	BC
TRTO	.33	.29	B
RAO	.26	.12	BCD
ELRO	.36	.25	B
TREL	.38	.28	B
MOYE	(b...).30	.23	

π MARCHE DES SHAMPOOING

ANALYSE DE VARIANCE ACORR = f(INDICE)
F = 4.5 Prob = 0.0007

INDICE	MOYENNE bi..	ECART TYPE	T-TEST à 10%
FRASA	.23	.22	B
FRASB	.22	.25	B
LEM	.43	.08	A
MERA	.12	.08	B
MERR	.10	.08	B
SABA	.15	.17	B
RAO	.23	.17	B
MOYE	(b...) .21	.19	

ANNEXE 21 EFFETS DES FACTEURS SUR CHAQUE INDICE

α MARCHE HUILE

INDICE	FISHER	PROB	REV	CSP	MAG	TAY	CONS	DAT	bi..
FRASA	2.70	.09							
bif.			.50	.14	.51	.58	.25	-	.40
t-test			AB	C	AB	B	BC		
FRASB	0.60	.67							
bif.			.54	.28	.45	.52	.28	-	.41
t-test			A	A	A	A	A		
LEM	3.82	.03							
bif.			.73	.62	.27	.79	.81	.84	.66
t-test			A	A	B	A	A	A	
MERA	0.56	.70							
bif.			.08	.22	.32	.12	.18	-	.18
t-test			A	A	A	A	A		
MERR	0.33	.88							
bif.			.17	.25	.34	.19	.17	.17	.22
t-test			A	A	A	A	A	A	
SABA	4.43	.022							
bif.			.30	.33	.08	.23	.12	.50	.23
t-test			BC	B	D	BCD	CD	A	
TOTT	1.62	.24							
bif.			.29	.27	.24	.33	.21	.70	.29
t-test			B	B	B	B	B	A	
TRTO	1.31	.33							
bif.			.46	.43	.23	.18	.32	.15	.31
t-test			A	A	A	A	A	A	
RAO	5.83	.009							
bif.			.26	.20	.17	.13	.06	.49	.18
t-test			B	B	BC	BC	C	A	
ELRO	0.47	.79							
bif.			.22	.19	.20	.35	.23	.43	.24
t-test			A	A	A	A	A	A	
TREL	0.44	.81							
bif.			.30	.41	.18	.29	.36	.12	.30
t-test			A	A	A	A	A	A	

β MARCHE BOISSON

INDICE	FISHER	PROB	REV	CSP	MAG	TAY	CONS	DAT	bi..
FRASA	1.01	.46							
bif.			.08	.30	.11	.28	.17	.38	.20
t-test			C	A	BC	AB	B	A	
FRASB	2.42	.11							
bif.			.26	.12	.20	.30	.13	.43	.21
t-test			B	B	B	AB	B	A	
LEM	3.01	.07							
bif.			.66	.55	.37	.63	.40	.78	.54
t-test			A	ABC	C	AB	BC	A	
MERA	1.17	.39							
bif.			.16	.16	.20	.47	.23	.03	.23
t-test			AB	AB	AB	A	AB	B	
MERR	0.52	.76							
bif.			.30	.18	.20	.21	.23	.19	.22
t-test			A	A	A	A	A	A	
SABA	1.92	.18							
bif.			.18	.22	.07	.14	.12	.32	.16
t-test			AB	AB	B	B	B	A	
TOTT	3.31	.05							
bif.			.18	.11	.19	.06	.04	.38	.13
t-test			B	B	B	B	B	A	
TRTO	5.23	.013							
bif.			.33	.36	.11	.40	.30	.14	.29
t-test			A	A	B	A	A	B	
RAO	2	.15							
bif.			.23	.20	.11	.22	.11	.40	.19
t-test			B	B	B	B	B	A	
ELRO	3.29	0.5							
bif.			.51	.24	.24	.65	.12	.45	.36
t-test			AB	BC	BC	A	C	AB	
TREL	2.6	0.09							
bif.			.32	.23	.11	.50	.32	.13	.29
t-test			AB	B	B	A	AB	B	

Γ PARFUM

INDICE	FISHER	PROB	REV	CSP	MAG	TAY	CONS	DAT	bi..
FRASA	0.40	.83							
bif.			.21	.16	.17	.08	.11	.08	.14
t-test			A	A	A	A	A	A	
FRASB	0.13	.53							
bif.			.12	.26	.23	.32	.12	.02	.20
t-test			A	A	A	A	A	A	
LEM	3.87	.03							
bif.			.72	.50	.36	.74	.51	.77	.58
t-test			AB	BC	BC	AB	BC	A	
MERR	0.5	.74							
bif.			.26	.31	.32	.23	.35	.28	.29
t-test			AB	AB	AB	B	A	B	
SABA	2.73	.08							
bif.			.15	.16	.06	.11	.04	.01	.10
t-test			AB	A	BCD	ABC	CD	D	
TOTT	1.01	.46							
bif.			.37	.27	.29	.53	.15	.31	.32
t-test			A	A	A	A	A	A	
TRTO	1.18	.38							
bif.			.41	.30	.11	.28	.36	.87	.33
t-test			A	B	B	B	B	B	
RAO	0.69	.64							
bif.			.29	.17	.21	.34	.29	.29	.26
t-test			A	A	A	A	A	A	
ELRO	3.29	.05							
bif.			.51	.24	.24	.65	.12	.45	.36
t-test			AB	BC	BC	A	C	ABC	
TREL	4.53	.011							
bif.			.51	.24	.24	.65	.12	.85	.38
t-test			AB	BC	BC	A	C	A	

π MARCHE SHAMP

INDICE	FISHER	PROB	REV CSP	MAG	TAY	CONS	DAT	bi..
FRASA	0586	0.65						
bif.				.18	.13	.36	.31	.23
t-test				A	A	A	A	
FRASB	0.25	0.86						
bif.				.23	.13	.33	.17	.22
t-test				A	A	A	A	
LEM	2.23	0.19						
bif.				.35	.43	.47	.52	.43
t-test				B	AB	AB	A	
MERA	3	0.11						
bif.				.20	.09	.09	.01	.12
t-test				A	AB	AB	B	
MERR	1.93	.23						
bif.				.17	.08	.08	.01	.10
t-test				A	AB	AB	B	
SABA	0.26	0.85						
bif.				.11	.23	.10	.18	.15
t-test				A	A	A	A	
RAO	0.19	0.90						
bif.				.19	.23	.30	.17	.23
t-test				A	B	A	A	

ANNEXE 22 VALEURS DES CORRELATION INTER FACTEUR PAR INDICE (bifl)

α MARCHE HUILE

		FRASA	FRASB	LEM	MERA	MERR	SABA	TOTT	TRTO	RAO	ELRO	TREL
REV	1	.61	.56	.74	.14	.26	.36	.52	.49	.31	.01	.23
	2	.34	.42	.74	.08	.20	.32	.33	.41	.34	.47	.61
	3	.56	.64	.71	.03	.05	.22	.02	.49	.14	.17	.05
CSP	1	-.01	-.18	.64	.26	.19	.29	.35	.18	.22	.37	.59
	2	-.04	-.20	.65	.25	.18	.42	.33	.51	.26	.09	.32
	3	.38	.46	.56	.15	.37	.27	.13	.59	.12	.11	.32
MAG	1	.73	.89	.70	.85	.81	.15	.36	.42	.25	.37	.51
	2	.49	.25	.04	.10	.11	-.03	.32	-.19	.17	.03	.00
	3	.32	.20	.06	.02	.10	-.05	.04	.09	.09	.21	.04
TAY	1	.66	.45	.93	.03	.18	.16	.35	.15	.14	.28	.22
	2	.60	.36	.74	.05	.14	.43	.37	.10	.17	.38	-.47
	3	.47	.74	.69	.27	.26	.11	.26	.28	-.09	.38	-.17
CON	1	.58	.75	.89	.21	.20	.19	.32	.17	.04	.50	.52
	2	.14	.08	.82	.22	.12	.07	.11	-.16	.13	.00	-.07
	3	.04	.02	.73	.11	.19	-.12	.20	.63	.01	-.19	.49
DAT	1	-	-	.84	-	.17	.50	.70	.15	.49	.43	-.12

β MARCHE BOISSONS

		FRASA	FRASB	LEM	MERA	MERR	SABA	TOTT	TRTO	RAO	ELRO	TREL
REV	1	.02	.20	.68	.18	.29	.08	.19	.34	.26	.42	.47
	2	.07	.29	.67	-.23	.23	.13	.26	.43	.29	.64	.47
	3	.15	.29	.64	.08	.38	.34	.08	.22	.14	.48	.02
CSP	1	-.08	.28	.64	.08	.23	.26	.08	.32	.30	.00	.16
	2	.42	.06	.54	.35	.18	.27	.14	.44	.17	.45	.26
	3	-.41	.01	.48	.05	.14	.12	.10	.32	.12	-.26	.26
MAG	1	.11	.23	.32	.17	.21	.05	-.10	-.19	.15	.41	-.10
	2	.14	.07	.52	.34	.32	.09	-.37	.09	.16	-.31	-.19
	3	.07	.30	.28	.08	.06	.06	.11	.05	.01	.01	-.04
TAY	1	.66	.35	.63	.15	.13	.05	.08	.28	.09	.45	.57
	2	.14	.29	.64	.86	.25	.20	.11	.47	.30	.62	.50
	3	.04	.25	.63	.39	.22	.18	.00	.45	.28	.88	.43
CON	1	.20	.00	.68	.39	.39	.07	-.04	.34	.22	.28	.31
	2	.17	-.19	.35	.22	.22	-.15	-.03	.30	.08	.06	.48
	3	.14	.19	.17	.08	.08	-.15	.05	.26	.02	.01	.18
DAT	1	.38	.43	.78	.03	.19	.32	.38	.14	.40	.45	.13

Γ MARCHE PARFUMS

		FRASA	FRASB	LEM	MERR	SABA	TOTT	TRTO	RAO	ELRO	TREL
	1	.08	.02	.71	.11	.22	.27	.04	.35	.42	.42
REV	2	-.47	.08	.66	.27	.10	.43	.77	.25	.64	.64
	3	.09	.26	.78	.39	.13	.41	.42	.26	.48	.48
	1	.16	.42	.54	.25	.22	.04	.14	.24	.00	.00
CSP	2	.19	-.20	.36	.35	.12	.47	.34	.10	.45	.47
	3	.12	.16	.59	.34	.13	-.29	.42	.16	-.26	-.26
	1	.35	.53	.43	.37	.08	.49	-.11	.14	.41	.41
MAG	2	.15	.10	.19	.27	.01	-.31	-.16	.08	-.31	-.31
	3	.00	.06	.47	.33	.08	-.08	.05	.41	.01	.01
	1	-.12	.74	.71	.20	.09	.21	.08	.19	.45	.45
TAY	2	-.09	.20	.80	.23	.18	.55	.05	.50	.62	.62
	3	.02	.03	.70	.27	.06	.84	.72	.32	.88	.88
	1	.23	.03	.75	.50	.10	.38	.72	.37	.28	.28
CON	2	.05	.18	.36	.19	-.02	.05	-.12	.13	-.06	-.06
	3	.04	.14	.42	.35	.01	.03	.25	.38	.01	.01
DAT	1	.08	.02	.77	.28	.01	.31	.87	.29	.45	.85

π MARCHE SHAMPOOINGS

		FRASA	FRASB	LEM	MERA	MERR	SABA	RAO
	1	.11	.39	.77	.22	.21	.14	.70
MAG	2	.25	.30	.82	.18	.14	.08	.73
	3	.45	.69	.78	.17	.14	.17	.56
	1	.25	.26	.80	.45	.45	.02	-
TAY	2	.088	.13	.80	.16	.16	.63	.63
	3	-.05	.53	.85	.23	.23	.07	-
	1	.27	.28	.80	.33	.35	.11	.70
CON	2	.37	.36	.81	.16	.26	.17	.76
	3	.83	.84	.86	.25	.35	.06	.75
DAT	1	.54	.49	.91	.42	.49	.19	.69

a r

Groupe HEC
BIBLIOTHÈQUE