

Université Paris Ouest Nanterre La Défense

Ecole doctorale Economie, organisations, société –
(ED 396)

THÈSE DE DOCTORAT EN SCIENCES DE GESTION

**Facteurs de contingence dans la mesure de satisfaction
clients**

Présentée et soutenue publiquement le **10 septembre 2013**

par

Zouaoui Sid-Ahmed Amine Bennabi

JURY

Paul-Valentin Ngobo, *Professeur à l'université d'Orléans*

Olivier de la Villarmois, *Professeur à l'IAE de Paris*

Jean-Louis Moulin, *Professeur, à l'Université de la Méditerranée*

Christian Barbaray, *Directeur Général' INIT*

Christophe Benavent *Professeur, à l'université de Paris Ouest*

Rapporteur

Rapporteur

Directeur de thèse

Remerciement

A l'heure où je clôture ce travail de thèse, le moment est venu d'adresser mes remerciements à tous ceux qui ont contribué, de près ou de loin, à l'aboutissement de ce long travail.

Je tiens à remercier dans un premier temps, toute l'équipe INIT, présidée par Mr Barbaray, qui m'a accueilli et accompagné durant cette thèse.

Je tiens à remercier tout particulièrement et à témoigner toute ma reconnaissance à mon directeur de thèse le Professeur Christophe Benavent pour l'expérience enrichissante et pleine d'intérêt qu'il m'a fait vivre durant cette thèse.

Je tiens également à exprimer ma profonde gratitude au Professeur Paul-Valentin Ngobo, au Professeur Olivier de la Villarmois et au Professeur Jean-Louis Moulin, pour avoir accepté d'évaluer ce travail doctoral et me faire l'honneur d'être membre du jury de soutenance.

Le mot de la fin sera pour mes Parents, et à toutes les personnes qui me sont chères.

Tables des matières détaillées

Tables des matières détaillées.....	2
Introduction.....	6
<i>Section 1 – La satisfaction comme indicateur de performance.....</i>	<i>16</i>
I – Mesure de satisfaction et calcul d’indicateurs de satisfaction.....	17
II – Objectif des mesures barométriques de la satisfaction.....	18
III – La mesure barométrique de satisfaction comme outils de pilotage.....	19
IV – Les motivations de la mesure barométrique.....	20
V – La satisfaction clients comme élément de contrôle de gestion.....	21
VI – Tableau de Bord.....	22
VII – La satisfaction et exigence des instruments de pilotage.....	24
<i>Section 2 – Le cadre de recherche : les effets de contingence.....</i>	<i>27</i>
I – Les effets de contingence comme causes des problématiques.....	29
➤ Les baromètres de satisfaction comme instruments de recherche.....	32
<i>Section 3 – Structure de la thèse.....</i>	<i>33</i>
Chapitre 1 : La satisfaction clients et sa mesure barométrique.....	36
<i>Section 1 – Le concept de satisfaction.....</i>	<i>36</i>
I – Les types de satisfaction.....	37
II – La satisfaction et les concepts voisins.....	39
<i>Section 2 – Modèle de la formation de la satisfaction : ses déterminants et conséquences.....</i>	<i>41</i>
I – Le modèle cognitif de base.....	42
II – Le modèle cognitif amélioré.....	42
III – Autres antécédents cognitifs de la satisfaction.....	43
IV – La satisfaction et l’insatisfaction.....	47
V – Conséquences de la satisfaction.....	48
<i>Section 3 – La mesure de satisfaction dans le cadre barométrique.....</i>	<i>49</i>
I – L’usage des baromètres de satisfaction.....	51
II – Les recherches basées sur les baromètres de satisfaction.....	52
<i>Section 4 – La satisfaction clients dans le contrôle de gestion.....</i>	<i>57</i>
Chapitre 2 : Facteurs de contingence face à l’évaluation des indicateurs.....	65

<i>Section 1 – Présence des facteurs de contingence le constat empirique</i>	66
I – Différences systématiques : l'exemple de l'âge et le sexe des individus...	67
II – Comparaison longitudinale : l'invariance des baromètres.....	72
<i>Section 2 – La différence entre unité sociodémographiques et géographique</i>	74
<i>Section 3 – La dynamique de la satisfaction dans la littérature</i>	76
<i>Section 4 – Vers une définition des facteurs de contingence</i>	82
I – Mesure barométrique comme agrégat	82
II – Comparer des agrégats : sources des différences et stabilité des indicateurs	85
III – Indicateurs de satisfaction face aux facteurs de contingence.....	88
<i>Section 5 – Propositions de test de représentativité et de sensibilité des indicateurs de satisfaction.....</i>	95
Chapitre 3 : Méthodologie de test des propositions.....	101
<i>Section 1 – Les données d'analyse : sources et structures.....</i>	103
I – Baromètre B1 : concernant une enseigne de prêt-à-porter :.....	104
II – Baromètre B2 : Concernant une grande enseigne de Bricolage	108
<i>Section 2 – La méta-analyse une méthode de gain en puissance statistique..</i>	113
I – La méta-analyse dans le marketing.....	116
II – Méthodologie d'une méta-analyse classique.....	118
III – Limite de la méta-analyse.....	121
<i>Section 3 – Existence, importance et sens de l'influence d'un l'effet.....</i>	123
IV – Importance de l'effet	123
V – Statistique descriptive	124
VI – Indicateurs de l'importance des effets et valeurs repères	125
<i>Section 4 – L'analyse de comparaison bayésienne</i>	128
I – Inférence fiducio-bayésienne pour l'étape inductive sur l'importance de l'effet	130
<i>Section 5 – Procédure de test des propositions : la méta-analyse par les procédés de l'analyse de comparaison.....</i>	133
I – Procédure fiducio-bayésienne conjuguée à de la méta-analyse par niveau supra-individus.....	133
II – Apport des procédures fiducio-bayésienne aux méta-analyses classiques	135
Chapitre 4 : Résultats	139

<i>Section 1 – Procédure et résultats du test de la proposition P1</i>	141
I – Indice pour juger de pouvoir explicatif d'un ensemble de variables.....	143
II – Organisation des données pour le test de P1	144
III – Effets des facteurs de contingence sur la variation du coefficient de corrélation multiple.....	149
IV – En résumé	155
<i>Section 2 – Procédure et résultats du test de la proposition P2</i>	156
I – Indice pour juger de la représentativité de l'indicateur lors de la comparaison 157	
II – Organisation des données pour le test de P2	158
III – Effets des facteurs de contingence sur la fluctuation de la variance intra magasin.....	162
IV – En résumé	166
<i>Section 3 – Procédure et résultats du test de la proposition P3</i>	167
I – Indice pour juger de la sensibilité dans le temps des indicateurs.....	168
II – Organisation des données pour le test de P3	169
III – Sensibilité dans le temps et facteurs de contingence	173
IV – Saisonnalité dans le temps et facteurs de contingence.....	175
V – En résumé	177
<i>Section 4 – Existence d'une structure hiérarchique</i>	177
I – Prise en compte du niveau Magasin	180
II – Prise en compte du niveau Région	183
<i>Section 5 – Conclusion du chapitre</i>	185
Chapitre 5 : L'indicateur de satisfaction face à son principal concurrent, le cas du Net Promotor Score	187
I – La satisfaction et la recommandation.....	189
II – La dynamique des deux dimensions de la croissance future	190
III – Le taux de satisfaits (TS) et Net Promotor Score (NPS) comme indices concurrents.....	191
IV – Satisfaction et recommandation quelle liaison	192
V – Méthodologie	194
VI – Résultats.....	196
VII – Interprétation de la dépendance des indicateurs	197
VIII – Conclusion	198

Discussion générale.....	200
<i>Section 1 – La modélisation multiniveaux une approche naturelle pour les données barométriques</i>	<i>201</i>
I – L’analyse multiniveaux en Marketing.....	203
II – L’intérêt de l’analyse multiniveaux	204
III – La modélisation multiniveaux pour les données barométrique.	205
<i>Section 2 – Vers une recherche ciblée sur la mesure barométrique</i>	<i>206</i>
I – La question de l’échelle	207
II – Échelles multi-items et mono-items.....	209
III – Organisation des données barométriques comme des modèles formatifs	211
Conclusion générale.....	213
Bibliographie	217
Liste des figures et tables.....	231

Introduction

L'intégration de la satisfaction clients dans le système de gestion des entreprises a pour origine la transformation de la notion de performance, qui évolue d'une pensée uniquement interne à l'entreprise, essentiellement axée sur les coûts engagés, vers une considération externe, tournée vers le marché. Cette adaptation est due à la saturation des marchés qui a fait que le succès futur des entreprises dépend aussi de la capacité à satisfaire ses clients. Face à une concurrence accrue, la valeur du produit/service commercialisé se voit liée au jugement des clients. Une valeur qui se matérialise donc par le prix que les clients sont prêts à payer.

La performance de l'entreprise sous-entend que, quel que soit le contexte, une entreprise est performante, au sens où elle est rentable, si elle satisfait les objectifs qu'elle se fixe¹. Avec ce nouveau mode de pensée alliant les deux composantes de la performance, interne et externe, la satisfaction clients comme concept théorique capable de répondre aux exigences de la performance « moderne » s'est imposée naturellement.

Des auteurs comme Kaplan et Norton ont fait de la satisfaction clients un des éléments du tableau de bord prospectif (Balanced Score Card ou BSC²). L'intérêt croissant porté à l'importance de la satisfaction a été confirmé par les travaux mettant en relation la satisfaction et la performance financière (Anderson, Fornell, Lehman, 1994; O'Sullivan, Hutchinson, O'Connell, 2009). La satisfaction fait désormais l'objet de mesure par baromètres³, cela permet de lui attribuer le statut d'indicateur de performance :

La mesure de satisfaction barométrique inclut plusieurs étapes qui vont de la conception des questionnaires (formulation des questions et choix de l'échelle), sélection de l'échantillon représentatif des clients, le mode de recueil des avis (face à face, par téléphone, par courriers papiers ou par courriers électroniques) et le calcul statistique visant à résumer l'enquête par des indicateurs qui servent à contrôler et suivre la satisfaction clients. Les résultats de ces mesures, qui sont présentés sous forme d'indicateurs, servent à évaluer la performance issue de l'exigence des clients. Les

¹ Citation de MJ Lebas(1995) dans Ferrah 2002

² Le BSC pour Balanced Scorecard, est constitué d'un ensemble d'indicateurs financiers et non financiers (satisfaction clients) nécessaires pour le pilotage de la performance.

³ Les baromètres sont des enquêtes de satisfaction basées sur des questionnaires uniques et qui sont répétées sur plusieurs vagues (ex : trimestre ou année). Ils visent à suivre l'avis clients concernant le produit /et ou le service de l'entreprise.

baromètres constituent donc un des outils privilégiés pour connaître le jugement client, mais il existe d'autres moyens pour apprécier les avis des clients, comme le traitement de réclamations, les enquêtes mystères. Le but recherché par tous ces procédés est double : d'une part ils visent à contrôler l'axe clients et d'autre part à produire des bons indicateurs (au sens du contrôle de gestion) pour piloter l'activité. Cette famille d'indicateurs, dont la satisfaction fait partie, est nommée : indicateur de performance. Un indicateur de performance est défini comme « un évènement observé, prélevé, mesuré, déterminé par un calcul, permettant d'identifier qualitativement ou quantitativement une amélioration positive (ou négative) du comportement d'un procédé » Périgord. M et Fournier J.P. (1993). Cette définition, même si elle reste relative à un cadre industriel, apporte l'information qu'un indicateur est le résultat d'un calcul, et l'objectif recherché est de pouvoir inférer une amélioration ou une détérioration d'un procédé. L'usage de l'indicateur doit donc être accompagné d'un souci de fiabilité, répond-il bien à ces objectifs ? Être opérationnel et permette de piloter l'activité. L'indicateur de satisfaction comme instrument de pilotage n'échappant pas à cette exigence de fiabilité, doit s'intégrer au fonctionnement du contrôle de gestion

Depuis longtemps la satisfaction clients est au cœur de la politique marketing. Sa mesure doit être l'objet d'un meilleur suivi vu son caractère opérationnel. Une satisfaction qui induit la fidélité et qui assure la pérennité de l'entreprise devient de moins en moins évidente lorsque les résultats des baromètres sont présentés. En pratique, les managers ont des difficultés à gérer efficacement leurs activités. Les points sur lesquels ils s'interrogent touchent la sensibilité des indicateurs de satisfaction dans le temps et lors des comparaisons entre unités de gestion, par exemple entre plusieurs magasins. Parce que le besoin de se comparer est essentiel pour dégager ses supériorités et identifier ses faiblesses, il est devenu important de statuer sur la fiabilité de leurs pratiques et de résoudre les problèmes méthodologiques de la mesure de satisfaction dans le cadre barométrique.

L'indicateur de satisfaction se distingue de la famille des indicateurs de performance par la nature de sa construction, il est basé sur le concept de satisfaction clients. La manière dont il est produit, résultats de sondage, dépend des avis clients. En reprenant par exemple le cas de la réclamation clients ou des délais de livraison, l'indicateur issu de ces attributs est un résultat déterministe, presque physique, contrairement à l'indicateur de satisfaction qui est lui entaché d'erreur aléatoire. Une erreur qui est liée à

l'échantillonnage, à la formulation des questions, au choix de l'échelle de mesure (sémantique ou numérique) et au traitement statistique des résultats. Même si il y a un contrôle de l'aléa dans l'évaluation de l'indicateur de satisfaction, l'erreur ne peut pas être éliminée complètement.

L'autre point qui distingue l'indicateur de satisfaction reste le concept de la satisfaction clients lui-même. Un concept qui est le fruit d'une construction psychologique possédant un caractère latent, difficile à évaluer. Ainsi deux caractéristiques de taille font que l'indicateur de satisfaction hérite de problématiques supplémentaires qui le rendent atypique parmi les indicateurs de performance.

L'utilisation des indicateurs dans la pratique de contrôle des performances, nous a conduits à engager une voie de recherche axée sur l'investigation des fiabilités des indicateurs produits par les baromètres de satisfaction. Deux points essentiels se sont dégagés dans nos réflexions : la notion de l'usage de la satisfaction comme indicateur, et celui du concept de satisfaction. Précisons que l'indicateur de satisfaction est un résultat agrégé relatif à un ensemble d'individus et a pour but de mesurer un niveau de satisfaction d'un segment⁴. Le concept de satisfaction constitue le **support théorique**, aujourd'hui compris et accepté par tous, qui a fait émerger l'intérêt de la satisfaction. De son côté, l'indicateur est **l'opérationnalisation du concept**, qui a pour objectif de servir le pilotage de la performance des entreprises. Bien sûr, nous ne pouvons pas penser l'indicateur de satisfaction sans le concept, toutefois nous insistons sur le fait qu'ils se situent sur deux plans différents. Le concept de satisfaction est théorique et l'indicateur est relatif à l'usage de la satisfaction.

De ce fait, une distinction claire doit se faire entre les deux objets : l'indicateur de satisfaction et la satisfaction comme concept. Notre recherche, en s'intéressant à la problématique des baromètres, s'attache à l'usage de la satisfaction pour piloter l'activité de l'entreprise par la performance commerciale (marketing) et non au concept psychologique de la satisfaction. Les conclusions induites à partir des indicateurs étant attachées à un groupe (ex : clients d'un magasin, clients d'une entreprise etc.). Nous avons voulu savoir si les résultats des indicateurs issus des baromètres représentent fidèlement l'objet qu'ils doivent mesurer face à l'existence de facteurs exogènes relatifs à

⁴ C'est la satisfaction relative à un ensemble de consommateurs qui partage la caractéristique d'être clients d'une entreprise, d'un magasin etc.

l'environnement et aux caractéristiques des individus. Ces facteurs que nous avons qualifiés de facteurs de contingence, sont les caractéristiques naturelles des groupes comme l'âge ou la région. Ils seront confrontés, tout le long de cette thèse, à l'indicateur de satisfaction afin **d'effectuer un diagnostic de ce dernier dans la pratique barométrique. Il s'agit de savoir s'il remplit bien son rôle d'indicateur de performance en représentant fidèlement le niveau de satisfaction.** Ce travail a pour ambition d'apporter des éléments de réponses aux problématiques rencontrées dans le quotidien des managers et permettre ainsi de bien situer l'indicateur de satisfaction dans la pratique barométrique. L'objet de la thèse s'axe sur l'étude de l'indicateur de satisfaction barométrique comme outils de pilotage et non sur l'étude de la satisfaction clients d'un point de vue de la psychologie de l'individu.

➤ **La pratique des mesures barométriques de la satisfaction clients**

Depuis une quinzaine d'années les enquêtes de satisfaction sont devenues monnaie courante pour des entreprises qui font appel à des instituts de sondage pour mesurer la satisfaction de leur clientèle et pour comprendre leurs besoins. Les plans d'action marketing visant à améliorer la relation clients et à garantir un bon échange entreprises/clients sont axés sur les résultats de ces enquêtes. L'accroissement de ces enquêtes de satisfaction (dans le temps et entre entreprises) a donné lieu à des baromètres de satisfaction dont l'utilité première est de produire un ensemble d'indicateurs de résultat et de suivis, destinés aux managers. Les indicateurs de résultats servent à juger l'accomplissement final d'une action : réalisation d'un objectif. Ils sont utilisés pour informer les niveaux hiérarchiques supérieurs sur la performance réalisée. Les indicateurs de suivi permettent d'observer la satisfaction clients dans le temps ou par rapport à la concurrence. Ils permettent aussi de récolter des informations afin d'entamer des actions correctives avant l'obtention des résultats. Ces deux types d'indicateurs sont étroitement liés. Ils visent essentiellement à orienter les actions pour le contrôle de l'activité.

Ces indicateurs de satisfaction clients sont aujourd'hui une partie intégrante des tableaux de bord prospectifs (comme le BSC) nécessaires au pilotage de l'activité. D'autres moyens sont aussi utilisés pour connaître les avis clients nécessaire pour piloter l'activité. Ces techniques regroupées sous le nom de pratiques d'écoute clients visent l'étude des comportements. Parmi ces outils nous trouvons les enquêtes mystères, analyses du sentiment, suivis des taux de réclamation ou encore des études de panel. Dans

notre recherche nous nous intéressons aux baromètres de satisfaction, car ils sont répandus chez les entreprises pour examiner les avis clients.

Toutes les entreprises modernes possèdent des baromètres de satisfaction clients qui leurs sont propres. Elles en font l'usage en basant leur management interne et leur communication externe sur les résultats obtenus. Bien plus qu'une simple mesure de satisfaction clients, les conclusions des baromètres représentées sous forme d'indicateurs, permettent aux commanditaires de connaître leurs positions auprès de leurs clients et d'apporter les améliorations nécessaires. Ces mêmes résultats sont aussi utilisés comme base de comparaison pour juger de la performance d'une firme donnée en la comparant à une autre entreprise proposant le même produit ou service. Ainsi les baromètres ont un avantage **endogène** (à chaque entreprise), et **exogène** (entre les entreprises d'un secteur) permettant le suivi de leurs activités et le contrôle de leurs marchés.

Le double usage, endogène et exogène, des résultats de mesures a produit deux sortes de baromètres :

- les baromètres internes : chaque entreprise en possède un spécifique à son activité et à ses besoins. Souvent ces baromètres contiennent des informations sur la satisfaction générale des clients et sur les satisfactions vis à vis de chaque attribut du produit et/ou service de la firme (ex : satisfaction envers l'accueil, prix du produit etc.).
- Les baromètres externes : ils sont souvent qualifiés de baromètres nationaux, car sont relatifs à un ensemble d'entreprises d'un même pays. Le but de ce baromètre est d'avoir un indice synthétique de la satisfaction clients, capable d'évaluer des performances économiques à un niveau national. Le plus populaire de ces baromètres, est celui de Fornell (1992)⁵ dans son article de présentation du baromètre suédois (SCDB). Ses travaux qui concernaient au départ la Suède et les États-Unis, ont vite été étendus à d'autres pays comme la Norvège puis à toute l'Europe.

L'utilité première des baromètres est donc de fournir des résultats sous forme d'un taux de satisfaction ou d'une moyenne qui sont employés comme des quantités résumant une information. Ils ont pour but de suivre et de piloter l'activité par unité de gestion

⁵ Fornell, Claes, (1992), A National Customer Satisfaction Barometer: The Swedish Experience, *Journal of Marketing*, 1-21

interne à l'entreprise⁶. La mise en œuvre de ces baromètres passe par plusieurs stades, qui vont de la sélection de l'échantillon de sondage jusqu'aux analyses statistiques visant à évaluer les indicateurs (moyenne ou taux). Curieusement les résultats de ces baromètres sont peu étudiés dans leurs fonctions de contrôle, malgré une littérature abondante qui s'est fort axée sur la satisfaction clients. Les résultats des baromètres ont un objectif principal qui est celui de permettre de faire des comparaisons dans le temps et entre unités de gestion⁷. Notre positionnement dans cette recherche est de savoir si les indicateurs de satisfaction permettent de telles comparaisons et à quelles conditions. C'est justement une des préoccupations qui a amené ce travail de thèse. Par leurs usages, les indicateurs de satisfaction issus des baromètres prennent une place importante⁸. Ainsi le fait de s'assurer de leur bon fonctionnement devient un sujet notable.

➤ **Problématique de comparaison : contrôle de gestion opérationnel**

Les pratiques de comparaison dans la satisfaction clients, visent à lier la supériorité d'une unité (temporelle ou spatiale) aux faits des actions Marketing. Le tout en basant ce raisonnement sur le seul résultat des indicateurs de satisfaction. Sur ce point de vue, l'indicateur doit être valide et ne doit en aucun cas être perturbé par des éléments autres que ceux qu'il doit mesurer : c'est-à-dire la satisfaction des clients de l'entreprise.

Le point soulevé dans la comparaison, qui touche la capacité des indicateurs de satisfaction à remplir leurs rôles de comparaison, est une question qui est relative à l'usage et spécifique aux baromètres. Là où la recherche marketing de la satisfaction clients s'intéresse à mettre en évidence des processus psychologiques qui déterminent ou émanent de la satisfaction, la mesure barométrique quant à elle s'intéresse à fournir des indicateurs (mesure agrégée) issus du sondage d'une population de consommateurs. Les résultats barométriques visent à être opérationnels et la recherche académique vise à modéliser le fonctionnement. Cette divergence en terme d'objectif, a empêché l'émergence d'un intérêt aux problématiques de la mesure barométrique dans leurs fonctions de comparaison. Dans les baromètres, les indicateurs supposés être représentatifs et sensibles servent à pointer toutes différences dans le niveau de satisfaction clients

⁶ Nous pensons notamment à l'unité relative à une activité qui fait interagir clients/entreprise.

⁷ Nous entendons par unité de gestion, chaque point d'organisation dans lequel le contrôle s'applique. Un magasin par exemple, doit être surveillé et constitue donc une unité de gestion.

⁸ Il n'est pas rare que des entreprises rémunèrent leurs employés suivant les résultats de la satisfaction clients, d'où l'intérêt de produire des mesures fiables afin d'éviter des erreurs de décision.

Un exemple marquant de ce genre de phénomène touchant les indicateurs lorsqu'on s'intéresse à comparer des unités entre elles et dans le temps, reste celui du baromètre national ACSI (American Customer Satisfaction Index – Cf. figure 1 ci-dessous)



Figure 1 L'évolution des indices ACSI

Nous observons sur la figure 1 une apparente stabilité de l'indicateur ACSI pourtant prévu dans sa construction pour varier de 0 à 100. Ici nous constatons que sa variation est très limitée. Un certain parallélisme entre les courbes est aussi présent. Ceci est synonyme d'une stabilité dans le temps avec un écart constant entre les secteurs. Nous remarquons ainsi que la sensibilité des indicateurs ACSI se manifeste plus dans une vision transversale (entre secteurs) que dans les portées temporelles. Ces résultats appuient l'existence de deux problématiques de comparaison : une longitudinale et l'autre transversale.

Pour les problématiques de comparaison transversale, nous pensons notamment aux différences constatées quotidiennement lors du traitement des enquêtes de satisfaction. Il s'agit de différences systématiques entre niveaux de satisfaction constatées suivant certains facteurs comme l'âge, le sexe ou encore le secteur (voir Figure 1). Certaines observations deviennent à la longue inquiétantes par leur caractère récurrent. Les sources responsables de ces problèmes jouent soit des rôles de modérateurs en surévaluant/ sous-évaluant les valeurs de l'indicateur, soit des rôles de médiateurs entre la mesure individuelle et la mesure agrégée (indicateur). Devant ces constats, les indicateurs

de satisfaction paraissent entachés d'incertitudes, évalués avec erreur, ce qui peut fausser la prise de décision.

Pour les problématiques de comparaison longitudinale, elles sont relatives aux questions développées précédemment, c'est à dire la stabilité dans le temps des indicateurs de satisfaction qui inhibe le processus de comparaison temporelle. L'indicateur représenté par le taux de satisfaction ou les moyennes, subit des manques de crédibilité, dans la mesure où leur stabilité met en défaut l'utilité de la satisfaction clients comme un levier d'action. Devant ce problème, son intérêt est mis en avant et son caractère prédictif du profit futur est contesté.

Comme cela va être détaillé plus loin, nous avons aussi remarqué qu'en pratique cette stabilité dans le temps reste sensible au niveau du calcul⁹. Plus l'agrégation est haute, plus l'indicateur semble être stable. En revenant aux indices ACSI, le détail, par secteur de ces derniers, nous montre aussi une stabilité (Figure 2), au niveau de chaque entreprise. Toutefois cette stabilité n'est pas du même ordre. Elle est marquante pour certains secteurs (secteur des : Brasseries « Breweries ») mais un peu moins pour d'autres (ex : aviation « Airlines »). Ce constat appuie le fait que la stabilité des indicateurs dépend du niveau de calcul, d'où l'intérêt d'examiner la sensibilité de l'indicateur de satisfaction face à ce constat.

⁹ Le niveau de calcul renvoie, à l'agrégation des données. Un indicateur calculé par magasin ou par région ou encore par pays, ne présente pas la même sensibilité. Pour le cas des données des baromètres ACSI les courbes présentent plus de différences entre secteurs que dans le temps. Cette propriété reste aussi valable quand on étudie l'évolution dans le temps des scores de satisfaction par magasins (des différences marquées entre magasins mais une certaine stabilité dans le temps au niveau de chaque magasin).

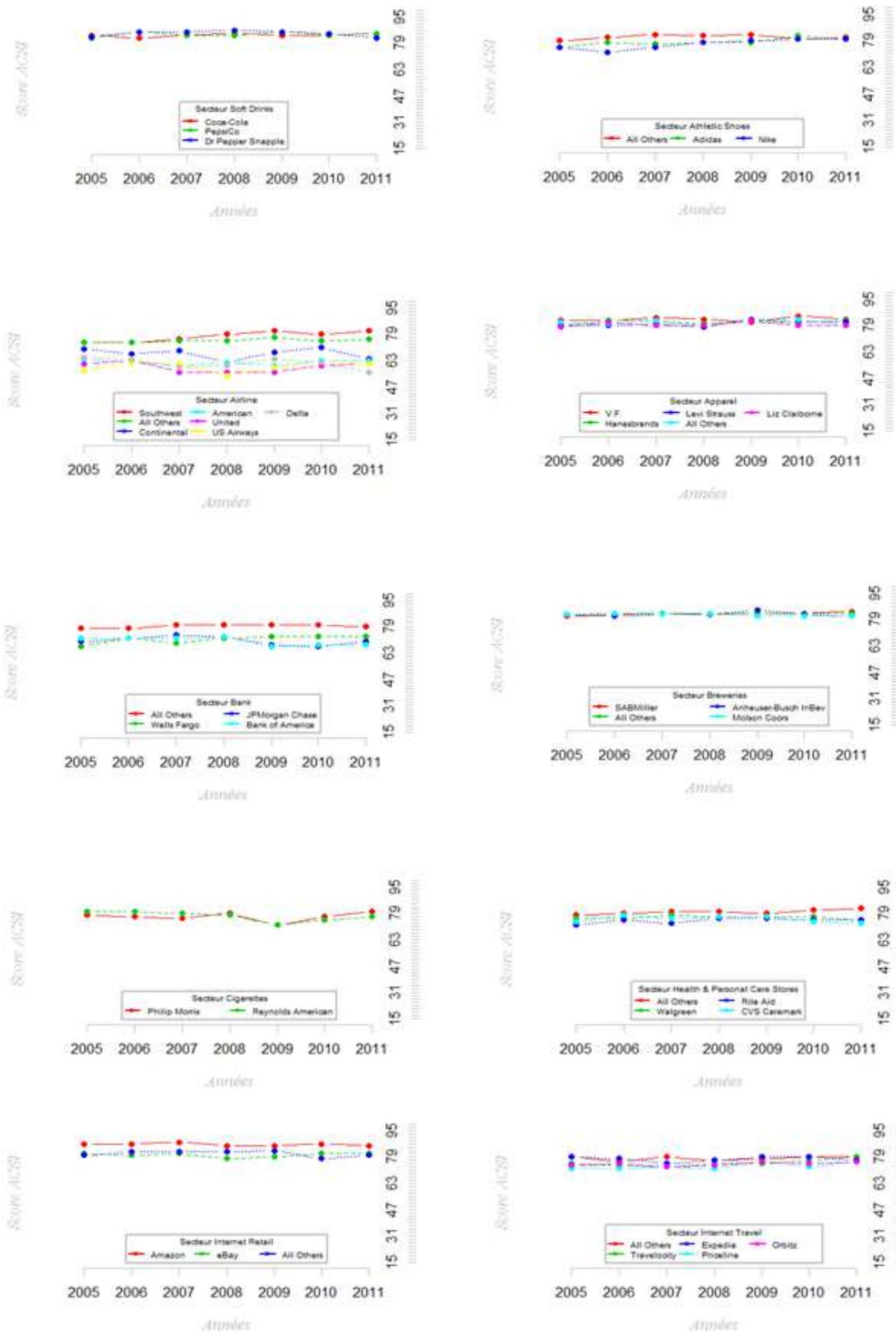


Figure 2 Évolution de l'ACSI par firme suivant 10 secteurs

L'exemple (Figure 1 et 2) ainsi présenté, pointe deux types de problématiques : l'une est longitudinale et l'autre est transversale (ici inter secteur). Elles sont loin d'être spécifiques au baromètre ACSI. Dans des baromètres internes, elles sont aussi présentes et se manifestent par exemple au niveau des magasins.

En général, il n'est pas rare que les managers trouvent que les résultats de satisfaction des baromètres internes présentent une différence systématique dans le niveau de satisfaction entre magasins. On imagine bien la difficulté des managers à fixer des objectifs d'amélioration lorsque les résultats des enquêtes de satisfactions révèlent des différences persistantes entre des unités de mesure (ex : entre magasins) malgré un effort marketing important visant à les corriger. Ce dernier point soulevé est à la base du contrôle. Le besoin de se comparer aux autres est essentiel pour dégager les supériorités et identifier les faiblesses. Sur ce sujet, des difficultés apparaissent pour trouver un indicateur ou un marqueur fiable pour se comparer soit en interne (entre plusieurs magasins), soit en externe (entre concurrents).

Un autre constat fait aujourd'hui est que les indicateurs de satisfaction calculés d'une vague à une autre bougent très peu. Un gain d'un point dans une moyenne ou dans un taux de satisfaits est vraiment très rare. De plus cet événement nous laisse penser qu'il existe un malentendu dans les procédés, car même si les clients semblent garder le même niveau de satisfaction au global, leur comportement individuel¹⁰ quant à lui n'est pas du tout stable. Cela touche des indicateurs globaux (sur l'ensemble des unités¹¹) qui sont produits pour être comparés entre eux dans le temps afin de procéder à des contrôles permanents visant à établir l'évolution des indicateurs et corrigeant si nécessaire les actions qui visent essentiellement une amélioration du produit ou service.

Ces constats mettent à mal les fonctions de l'indicateur de satisfaction, et pointent la faiblesse de la mesure barométrique à rejoindre le cadre académique¹².

¹⁰ Ici nous ferons référence aux comportements du consommateur pris individuellement. Chaque client est caractérisé par une expérience envers l'entreprise, qui par essence évolue dans le temps.

¹¹ L'ensemble des unités correspond ici à un agrégat de tous les avis clients sans distinguer l'unité de gestion. Par exemple produire des indicateurs sur l'ensemble des clients d'une entreprise comme compilation de tous les avis récoltés pour chaque magasin.

¹² Par définition la satisfaction est relative et subjective évolutive (varie dans le temps).

Ainsi l'indicateur de satisfaction est sujet à au moins deux phénomènes qui pointent deux faiblesses lors de l'étape de comparaison, d'où l'intérêt de s'intéresser à son fonctionnement et sa fiabilité comme indicateur de performance. Par cela et compte tenu des remarques précédentes, l'intérêt de s'intéresser aux problématiques des managers et répondre à leurs questionnements se justifie pleinement. Devant ces questions, l'intérêt de la recherche s'est établie naturellement et elle s'est orientée vers un examen des indicateurs de satisfaction dans le contexte barométrique en considérant deux champs : un temporel, plusieurs vagues de sondage, et l'autre transversal, par point de vente.

L'objet donc de cette thèse est d'explorer à travers des études de cas dans la distribution, les conditions de comparaison entre magasins d'une même entreprise, le tout en explorant les indicateurs de satisfaction dans leur fonction. Il s'agit de vérifier la représentativité et la sensibilité de l'indicateur à fournir les résultats escomptés pour gérer l'activité. En d'autres termes, nous nous intéressons à l'usage de l'indicateur de satisfaction comme outil de gestion. En s'appuyant sur les deux problématiques de comparaison longitudinale et transversale, nous cherchons à évaluer si les baromètres sont assez fiables pour produire les indicateurs et pour juger de la performance entre unités, ici entre magasins.

L'unité de prédilection dans cette recherche est l'unité magasin car lorsqu'elle existe, elle est au cœur du contrôle de la performance. C'est en se basant sur cette unité de gestion que l'étude des résultats de comparaison va être effectuée. A partir de cette perspective axée sur des problèmes de comparaison, nous nous positionnons dans un cadre d'inspection de la mesure de satisfaction barométrique à travers l'examen systématique des propriétés d'un bon indicateur de performance.

A ce stade de la présentation de notre problématique, deux points doivent être développés, Le premier est la notion de l'indicateur et le deuxième le cadre de développement de cette recherche qui est celui des facteurs de contingence.

Section 1 – La satisfaction comme indicateur de performance

La mesure barométrique de la satisfaction clients est un appareil de mesure visant à connaître l'opinion des consommateurs sur leurs besoins, leurs attentes et leurs expériences à l'égard du produit ou service. Le tout avec des questionnaires adaptés qui sont répétés trimestriellement ou annuellement.

La mesure est l'acte par lequel on recueille une opinion particulière, ici relative à la satisfaction, auprès d'un ensemble de sujets convenablement choisis. La mesure est un ensemble d'opinions dont on peut analyser ensuite la distribution. La bonne mesure est celle qui mesure sans biais l'objet qu'elle se donne. Le domaine académique s'est donné des outils standardisés¹³ afin d'effectuer des bonnes mesures. Pour l'indicateur de satisfaction barométrique, il pourrait être vu comme la résultante de l'agrégation de toutes les mesures individuelles pour résumer l'information au niveau de la population étudiée. Ainsi nous retiendrons qu'un indicateur est relatif à un groupe d'individus et à une unité, dans notre cas c'est le magasin.

I – Mesure de satisfaction et calcul d'indicateurs de satisfaction

Un point de divergence entre la notion de mesure de satisfaction dans le cadre académique par rapport au champ barométrique est celui de l'usage¹⁴. La finalité, d'une mesure de satisfaction académique, est de capter une information de corrélation sur le concept avec des outils efficaces servant à tirer des conclusions interprétables au niveau de l'individu. Le but de la mesure barométrique de la satisfaction est de fournir des résultats sous forme d'indicateur (score de satisfaction, moyenne ou taux de satisfait) qui sont des représentants de l'état de la satisfaction de la population. Le tout pour un objectif de contrôle de performance. Il s'agit d'une mesure empirique qui pourrait être vue comme une traduction imparfaite du concept de satisfaction sous forme d'indicateurs tangibles. Ainsi la distinction entre l'indicateur de satisfaction et la mesure de satisfaction comme elle est entendue, peut se définir comme suit :

➤ **La mesure de satisfaction**

La mesure consiste à établir des instruments (ici des questionnaires) en construisant des échelles capables de chiffrer la satisfaction d'un individu, le tout en formulant une question habile pour capter l'information recherchée. La mesure de satisfaction a un aspect psychométrique et estime la satisfaction d'un individu.

¹³ Les recherches académiques s'intéressent à inférer au niveau individuel (psychométrique). Ils se sont dotés des outils pour construire des échelles efficaces répondant à l'exigence psychométrique. Voir AURIER P., EVRARD Y. (1998), *Élaboration et validation d'une échelle de mesure de la satisfaction des consommateurs*, Actes du 14^{ème} *International de l'Association Française du Marketing*,

¹⁴ Une recherche de MORGAN A., ANDERSON E. W., MITTAL V., 2005, intitulée : *Understanding firm's Customer satisfaction information usage*, a permis de mettre en lumière les divergences entre les pratiques des entreprises en terme de satisfaction et la différence dans l'usage entre les entreprises suivant l'importance accordée aux enquêtes de satisfaction pour gérer leurs organisations

➤ L'indicateur de satisfaction

Il s'agit du calcul d'une statistique (taux, moyenne, score) obtenu au niveau de l'agrégat de tous les individus. Il complète la mesure en résumant une information agrégée. Il prend le statut d'indicateur car il est considéré comme une donnée utilisée pour juger de la satisfaction de la population de consommateurs. L'indicateur doit apporter une information sur l'ensemble de la population interprétable au niveau de cette dernière¹⁵.

Dans cette recherche, il serait question de travailler sur la pertinence des indicateurs de satisfaction dans leurs fonctions de comparaison. Comme l'indicateur est attaché à l'ensemble des mesures individuelles, il en vient que les problématiques posées concernent l'indicateur de satisfaction et la mesure elle-même.

II – Objectif des mesures barométriques de la satisfaction

Dans la pratique, les baromètres de satisfaction clients sont utilisés pour déceler les faiblesses et pour décider des améliorations. Sur ce principe, différentes questions sont posées aux clients afin de connaître leurs points de vue sur les dimensions du service ou les attributs d'un produit. Suivant leurs fonctions, les baromètres sont des outils essentiels pour une firme qui cherche à s'améliorer, diminuer les coûts de l'insatisfaction et surtout se comparer à d'autres structures. Les baromètres ont donc été créés pour sonder et suivre la satisfaction des clients afin de maîtriser les activités à moyen/long terme. Ce mode de recueil de la satisfaction sert à doter les organisations de deux leviers essentiels qui sont l'amélioration du service/produit et la gestion des attentes clients.

Un baromètre de satisfaction comprend un certain nombre d'items (questions) qui sont déterminés par une première phase qualitative. Cette étape sert avant tout à dégager les dimensions, de services ou de produits, importantes dans l'explication de ce qui est appelé : « la satisfaction générale du consommateur ». Dans la pratique, les questionnaires contiennent plusieurs items structurés sous forme de « thèmes » relatifs à la satisfaction et une question générale utilisée pour capter « la satisfaction générale du consommateur ».

Les résultats de la mesure barométrique doivent répondre à **quatre** objectifs¹⁶ :

¹⁵ Un taux de satisfaction indique la proportion des satisfaits dans la population.

¹⁶ Boss et al : les politiques de satisfaction des clients : une recherche exploratoire. Acte de *XI^{ème}* congrès de L'AFM Reims, p997-1026

- Identifier les leviers de la satisfaction et de l'insatisfaction des clients
- Hiérarchiser les éléments en termes de contribution à la satisfaction et à l'insatisfaction
- Fournir des indicateurs de la satisfaction des consommateurs de l'entreprise
- Dégager les axes d'amélioration.

III – La mesure barométrique de satisfaction comme outils de pilotage

Dans une littérature abondante sur la satisfaction clients, les conséquences économiques et managériales de la satisfaction sont bien établies. Les principaux enseignements de la littérature marketing consacrée à ses effets appuient sur l'importance de la satisfaction des clients sur la performance de l'entreprise et la réussite commerciale (Fornell 1992, Harting 1998, Zairii 2000, Bolton, Kannan et Bramlett 2000, Rampersad 2001, Gounaris et Stathakopoulos 2001, Anderson, Fornell et Mazvancheryl 2004). Une variété d'autres études, souligne que la satisfaction est une condition nécessaire pour la rétention et la fidélité client (Fornell 1992, Bolton 1998, Morgan et Strati 2000), et qu'elle aide à sécuriser les revenus futurs (Rust et Zahorik 1993, Scheuing, 1995; Reichheld, 1996; Helgesen , 2006,). Ces études nous montrent que nous devons donc percevoir des clients satisfaits en tant qu'actifs d'une entreprise qui porte une forte signification économique dans un environnement hautement concurrentiel.

D'autres recherches affirment également que si un haut niveau de satisfaction est garanti, les entreprises peuvent réduire les coûts des transactions futures (Reichheld et Sasser 1990), diminuer la sensibilité au prix (Anderson 1996) et abaisser le comportement des plaintes (Fornell 1992, Bolton 1998). La littérature présente aussi plusieurs effets positifs de la satisfaction sur l'amélioration de l'image de la marque (Glowa 2000), sur la réputation de l'entreprise et des recommandations positives (Anderson et Sullivan, 1993, Zairi 2000, Anderson, Fornell et Mazvancheryl 2004), ainsi que sur la motivation des employés (Muffatto et Panizzolo , 1995, Naumann et Hoisington 2001). Il ressort donc de ces recommandations, que le besoin constant de communication avec les clients et la capacité d'écouter leurs opinions sur les produits ou services, deviennent des enjeux très importants.

Le monde professionnel des études suivant ces références, s'est attelé à appliquer les résultats des scientifiques en donnant à la satisfaction le statut de paramètre de pilotage. Dans cette lancée, les entreprises ont organisé toute une structure managériale

autour de la satisfaction clients via les baromètres. Ce dispositif dans lequel la satisfaction des clients est mesurée et ses résultats sont mis en avant, est devenu une composante essentielle dans le contrôle de gestion moderne.

De plus, il n'est pas rare que dans les baromètres de satisfaction, les résultats soient organisés par centres d'intérêts. Ces mesures permettent au gestionnaire de connaître l'évolution de la satisfaction pour chaque dispositif lié à l'activité (ex : par magasin, par région ou par pays). Le but de ces résultats étant, d'une part d'optimiser la gestion globale liée aux fonctionnements de la firme et d'autre part de définir et d'allouer les ressources nécessaires (humaines et financières) pour atteindre les objectifs voulus. Ainsi la satisfaction est devenue aujourd'hui un élément sur laquelle les managers s'appuient pour prendre leurs décisions.

Mesurer l'avis des consommateurs est aujourd'hui « monnaie courante » dans un contexte où le souci de la performance est prédominant. Grâce au développement des techniques de recueil d'avis, la mesure de satisfaction clients dite microéconomique (au niveau de la population), est devenue un instrument incontournable dans le quotidien des gestionnaires.

IV – Les motivations de la mesure barométrique

La mesure barométrique vient tout d'abord des recherches sur le concept et de la recommandation des chercheurs. En prenant comme une évidence les résultats des recherches, les entreprises ont trouvé en la mesure de satisfaction un élément qui peut pallier les difficultés pratiques des firmes liées à leurs business. La motivation des entreprises pour procéder à une mesure barométrique de la satisfaction peut se résumer par les nécessités suivantes¹⁷ :

- L'avantage concurrentiel et différenciant : devant une concurrence de plus en plus rude, la satisfaction clients est perçue comme un avantage déterminant différenciant les concurrents.
- Le besoin de fidéliser : La fidélisation est devenue un objectif majeur de la politique commerciale et, pour atteindre ce but, il faut connaître la satisfaction de ses clients

¹⁷ Éléments tirés de l'article de J-F BOSS, Y, EVRARD J.LEVY et D, MANCEAU 1995 : Les politiques de satisfaction des clients : une recherche exploratoire *Actes du congrès Volume : 11 Pages : 997-1026.*

- La mobilisation : la politique axée sur la mesure de satisfaction peut transformer une culture du produit vers une culture axée clients. Elle permet ainsi de sensibiliser le personnel à l'importance du client.
- La continuité de la qualité : A la suite des processus de qualité formalisés, la mesure de satisfaction clients s'est installée naturellement pour juger de l'avis clients et mieux organiser les efforts.
- L'apprentissage : connaître les clients en s'informant de leurs attentes.

V –La satisfaction clients comme élément de contrôle de gestion

Afin de mieux cerner la pratique de la satisfaction dans sa mesure barométrique, un passage par le contrôle de gestion est obligatoire. Dans les éléments développés précédemment, nous avons insisté sur l'importance de la satisfaction comme concept. Mais pour mieux connaître son approche dans la pratique (indicateur de satisfaction) où elle est utilisée comme instrument de pilotage, il faut se référer au domaine du contrôle de gestion.

Antony (1965) est parmi les premiers auteurs à définir le contrôle de gestion. Dans un premiers temps, il définit le contrôle de gestion comme un processus par lequel les managers obtiennent l'assurance que les ressources sont obtenues et utilisées de manière efficace pour la réalisation des objectifs. Puis en 1988, l'auteur fait évoluer sa définition pour la faire sortir d'un cadre trop comptable et la présente par : *le contrôle de gestion représente le processus par lequel les managers influencent les autres membres de l'organisation pour mettre en œuvre la stratégie de l'organisation.*

Dans ces définitions, il ressort que le contrôle de gestion est une information sur laquelle les managers se basent pour orienter les comportements des autres membres, le tout vers un objectif d'identification des relations entre les actions et les finalités du pilotage. De ce point de vue le contrôle de gestion est considéré comme un outil du manager pour organiser l'activité afin de mettre en œuvre la stratégie de l'organisation ou de la société.

D'autre part, ce qui est sous-entendu dans les définitions précédentes, est que le contrôle de gestion n'est donc pas un ensemble de pratiques uniformes et que les outils sur lesquels il se base pour parvenir à ces fins peuvent différer (N.Berland 2009)

Le contrôle de gestion oscille entre deux conceptions. La première vise à savoir si le comportement des acteurs de l'organisation est conforme à des normes préétablies de performance, et la seconde suit une logique d'exploitation des voies de croissance et d'organisation des changements.

Dans le contrôle de gestion moderne, deux outils prédominants se complètent souvent. Le premier reste le contrôle budgétaire et le deuxième est le tableau de Bord. Chacun d'eux est utilisé selon le besoin même si beaucoup d'entreprises privilégient le premier moyen.

- Le contrôle budgétaire vise à décliner la stratégie et l'évaluer de manière exclusivement financière et comptable. Cet outil projette sur des sous unités de gestion des impératifs financiers pour un suivi des écarts du réel par rapport aux objectifs à réaliser.
- Le tableau de bord se distingue par une autre démarche, dans laquelle des objectifs financiers et non financiers sont fixés puis évalués à l'aide des indicateurs de suivis. Ce travail est aussi décliné par sous-unité de gestion, le suivi quant à lui s'effectue grâce aux mesures obtenues par les indicateurs qui sont comparés pour assurer un bon déroulement des opérations.

Parmi ces deux outils qui servent au pilotage, le tableau de bord attire tout particulièrement notre attention car il inclue des indicateurs non financiers dont la satisfaction clients est une composante.

VI – Tableau de Bord

Par analogie aux tableaux de bord des voitures ou des avions, les composantes (indicateurs) de cet instrument ne sont jamais fixes et évoluent selon le besoin et selon l'objectif à atteindre. Pour cerner les propriétés d'un tableau de bord en gestion, nous empruntons une définition de Bouquin (1998), qui le définit comme :

« Un tableau de bord est un ensemble d'indicateurs peu nombreux (cinq indicateurs) conçus pour permettre aux gestionnaires (managers) de prendre connaissance de l'état et de l'évolution du système qu'il pilote »

Un tableau de bord doit être construit à la lumière de ce qui doit être réalisé (objectifs), et doit également être établi en fonction des besoins des managers car le contrôle leur est destiné. Nous voyons bien que le tableau de bord d'un avion satisfaisant

les désirs des ingénieurs et non des pilotes, devient vite dangereux. Dans sa mise en forme, un tableau de bord inclut des indicateurs limités en nombre afin de faciliter le pilotage¹⁸. Les indicateurs de type financier et non financier inclus dans le tableau de bord doivent être acceptés par tous, compris par tous, prospectifs et surtout utiles.

En résumé, un tableau de bord doit refléter la complexité du système qu'il présente en montrant les interactions et les logiques d'actions qui sont inhérentes à la performance, le tout en ayant des indicateurs valides. De ce fait, les indicateurs qui le définissent doivent aussi refléter les interactions nécessaires pour le suivi global de la performance. D'ailleurs, c'est à cette tâche que le BSC (Balanced Scorecard) s'est attelé. Le Balanced Scorecard ou tableau de bord prospectif, est un instrument de contrôle initié par Robert Kaplan et David Norton. L'objectif de ce dispositif est de mesurer et améliorer la performance à travers des indicateurs financiers et non financiers directement liés entre eux et à la stratégie de l'entreprise. Il tient son développement des travaux antérieurs de William Ouchi (1979) complétés par Norman Machintosh (1994), qui ont étudié la typologie des systèmes de contrôle¹⁹. Son apport comparé à d'autres instruments est son pouvoir de présenter des indicateurs liés par une chaîne de causalité, qui doit mener à l'atteinte des objectifs. Les indicateurs préconisés par Kaplan et Norton s'organisent en quatre axes²⁰(Figure 3):

- Axe Financier : c'est l'axe central, il sert à définir des objectifs quantifiés : rentabilité, marges, réduction des coûts
- Axe Clients : il traduit la préoccupation d'attirer des nouveaux clients et de fidéliser les anciens
- Axe Processus interne : il concerne l'amélioration de la qualité des produit, du service après-vente et la qualité de service/ et des produits en général.
- Axe Apprentissage organisationnel : il s'intéresse aux performances du travail, la productivité des salariés leurs motivations et la communication de l'entreprise.

¹⁸ Une personne aux capacités cognitives normales ne peut percevoir qu'un nombre limité d'informations à la fois. MILLER G.A. (1956), « The magical number seven, plus or minus two: some limits on our capacity for processing information », *Psychology*, volume 63, pp.81–97.

¹⁹ Nous renvoyons le lecteur aux travaux des deux auteurs qui spécifient la typologie du contrôle

²⁰ Éléments tirés de la communication de Youssef Errami « les apports du Balanced Scorecard à la recherche des performances », 2007.



Figure 3: Le Balanced Scorcard de Kaplan et Norton

Le Balanced Score Card permet ainsi de prendre en compte les différents aspects de la performance au-delà de la performance financière. Il permet aussi de mettre en évidence les relations des causes à effet pour permettre d'expliquer la performance selon la stratégie mise en place. Il existe des critiques vis-à-vis de son fonctionnement car il exige une stratégie définie au départ et qui fonctionne selon une logique «top-down»²¹. En revanche son raisonnement causal qui lie les indicateurs est quant à lui vu comme une innovation²².

Ainsi le pilotage moderne passe aussi par des tableaux de bords incluant des indicateurs non financiers qui, aujourd'hui, deviennent aussi importants par la liaison qui les attache aux indicateurs financiers.

VII – La satisfaction et exigence des instruments de pilotage

Dans les littératures du contrôle de gestion, des auteurs comme Banker, Potter et Sirinivasan (2000) ou encore Itner et Larken (1998), se sont intéressés à corréler les indicateurs non financiers, notamment la satisfaction clients, à la performance financière et ont trouvé que la satisfaction clients est associée à la performance financière à long terme plutôt qu'à court terme. Itner et Larken (1998) montre aussi que les entreprises ayant des clients satisfaits contrôlent mieux les coûts. Ces travaux rejoignent la littérature

²¹ Stratégie élaborée au niveau le plus haut de la hiérarchie pour être exécutée au niveau le plus bas.

²²Choffel. D et Mayssonnier. F « Dix ans de débat autour du Balanced Scorecard » *Comptabilité et Connaissances* (2005)

Marketing sur l'importance de la satisfaction et mettent l'accent sur l'importance des indicateurs non financiers dans le contrôle de gestion.

A la lumière de ces éléments, les indicateurs issus des baromètres de satisfaction ont pris une place prépondérante comme indicateurs clé visant le pilotage de l'activité et la mesure de la performance.

L'intérêt porté aux baromètres de satisfaction s'inscrit avant tout dans une logique de contrôle où la satisfaction devient une composante des tableaux de bords. Les baromètres permettent de fournir des indicateurs non financiers²³, ici la satisfaction clients. De ce fait, un indicateur de satisfaction clients doit répondre à certaines exigences de fiabilité, autant que les autres indicateurs. Les qualités nécessaires pour qu'un indicateur soient de bonne qualité (Berland.N 2009, p131) :

- **Objectif** : c'est à dire qu'un indicateur ne doit pas dépendre de l'interprétation du manager.
- **Fidèle** : l'indicateur doit donner la même réponse lorsque les situations mesurées sont strictement identiques.
- **Simple et compréhensible** : la méthode de calcul doit être simple et compréhensible par tous
- **Sensible** : l'indicateur doit être réactif vis à vis du phénomène qu'il est censé mesurer.

De plus au-delà des qualités d'un bon indicateur, il doit aussi répondre à des exigences pour pouvoir être inclus naturellement dans un tableau de bord. Il se doit ainsi d'être :

- **Corrélé aux leviers** d'action auxquels il est associé et être indépendant des autres influences. Autrement dit, une mesure de qualité de service sera valide dans la mesure où elle est bien corrélée aux variations de la conception du service. S'il s'avérait que la qualité est fortement liée à d'autres variables, il deviendrait difficile de distinguer les éléments dont l'accroissement dépend des actions menées sur cette qualité.
- Sans biais ni bruit, il est **précis**. C'est ce qui est demandé à la mesure, mais l'indicateur étant un agrégat, s'ajoute la question de la composition de la

²³ Nouveaux indicateurs introduits suite à des critiques des indicateurs exclusivement comptables

population dans le cas des sondages de la satisfaction clients, à travers l'éventualité de certains biais et de la présence d'hétérogénéité.

- Il doit être **légitime** et accepté par toutes les parties prenantes. Lorsqu'il est destiné à fonder la légitimité des prétentions des entreprises, le principal moyen de sa légitimité est son indépendance. Au-delà de sa conception, ce qui devient important est l'organisation qui le conçoit.
- Il peut et doit être **incitatif**. Soit par sa simple publication qui encourage l'émulation entre les équipes évaluées par ce critère, soit parce qu'on y attache des sanctions et récompenses. L'indicateur peut ne pas être incitatif quand l'effort accru ne se traduit pas par une variation corrélative.

Implicitement, donner le statut d'indicateur à la satisfaction revient à la faire sortir de sa considération individuelle²⁴ pour la faire évoluer vers une considération globale relative à un ensemble d'individus. Un indicateur doit être complet vis-à-vis de ces propriétés pour que la procédure de comparaison entre unités de gestion, soit la plus juste possible.

Pour l'indicateur de satisfaction : il est supposé suivre les mêmes qualités que les autres indicateurs de performance. Comme l'indicateur n'est pas le résultat d'une mesure physique, mais d'un calcul de statistique issue des procédures de sondage, son objectivité est discutable sur deux points :

- le concept de satisfaction qui est subjectif, chaque individu a une satisfaction propre et définie par lui. L'indicateur dépend donc aussi de sources qui lui sont extérieures, les clients.
- Son usage est une interprétation qui ne possède pas un cadre de référence, sur lequel il peut être comparé. Par exemple une note moyenne de 8 pour un magasin est un excellent résultat, relativement aux autres magasins, mais elle peut aussi être considérée comme note insuffisante lorsque les autres notes enregistrées sont supérieures. L'absence donc de cadre de référence laisse l'indicateur de satisfaction barométrique très dépendant du contexte.

Suivant ces remarques, il faudra s'assurer que les indicateurs de satisfaction représentent bien l'état réel de ce qu'ils doivent mesurer, d'où l'idée d'explorer leur représentativité et

²⁴ Modèle basé sur la comparaison entre attentes et valeurs perçues au niveau de l'individu.

leur sensibilité. L'examen des indicateurs est accompagné par la considération du contexte où l'indicateur est évalué.

Section 2 – Le cadre de recherche : les effets de contingence

En acceptant la satisfaction comme un élément clé pour le pilotage de la performance d'une firme, des obligations concernant l'indicateur apparaissent (ceux définis précédemment). Un indicateur de satisfaction rapproché aux autres indicateurs de performance doit respecter des exigences, dont certaines constituent le socle de cette recherche. Nous distinguons ainsi trois obligations, qui pour nous sont essentielles pour le pilotage par la satisfaction :

- a. Correspondre à un levier d'action.
- b. Être fiable et précis.
- c. Permettre des comparaisons valides.

Le respect de ces exigences fera que l'indicateur fournira une information fiable et robuste du concept, ici la satisfaction microéconomique. Il permettra une analyse fidèle qui le lie avec ce qu'il doit mesurer, comme le montre le schéma suivant :



Figure 4 : Liaison concept et indicateur

De ce fait, l'indicateur doit représenter une information tangible du concept, pour que la généralisation soit sans biais et afin que le gestionnaire puisse conclure au sujet de la satisfaction de leurs clients et non pas au niveau de l'indicateur (un bon taux de satisfaction doit signifier à 100% une bonne satisfaction). D'autre part, les résultats des baromètres structurés sous forme d'indicateurs servent aussi comme instrument de

communication en vue d'envoyer un signal sur la solidité de l'activité. Par leurs fonctions, ils sont utilisés comme des indicateurs de performance qui doivent répondre à des caractéristiques et à des objectifs de pilotage, mais aussi comme des outils de communication. Par cela, les indicateurs de satisfaction doivent permettre d'avoir une information discriminante dans l'aide à la décision et ne doivent en aucun cas faillir à leurs objectifs de contrôle, d'où l'intérêt d'étudier leurs robustesse vis-à-vis des exigences citées auparavant.

L'indicateur de satisfaction barométrique est le fruit d'un calcul basé sur une distribution statistique issue de l'étape du sondage ; dès le départ il paraît présenter un manque de fiabilité compte tenu de son caractère subjectif, issu d'un jugement des clients. Pour illustrer le problème que peut rencontrer un indicateur lors de la mesure du concept de satisfaction, nous proposons de l'examiner sur deux situations²⁵ :

Situation idéale : il y a concordance parfaite entre indicateur et concept: le passage de l'information entre indicateur et concept est optimale (Figure 5).



Figure 5 : Situation de concordance

Situation discordante : l'indicateur recouvre en partie le concept, il existe un manque de sensibilité : le passage de l'information est biaisé et complété par autre chose (Figure 6).



Figure 6: Situation discordante

²⁵ Réseau **Circum**, consultation en gestion et en recherche. Url : <http://www.circum.com>

La question de recherche qui est la nôtre, sera de vérifier laquelle de ces deux situations est la plus réaliste dans la mesure barométrique de la satisfaction clients. Quelques éléments issus de la pratique barométrique nous guident déjà vers des suppositions de discordance entre le concept de satisfaction et l'indicateur. La démarche poursuivie sera entamée par une investigation des indicateurs de satisfaction barométrique en se servant d'une grande masse de données mise à notre disposition. L'examen sera accompagné par une considération des exigences des instruments de pilotage afin de juger si une situation de discordance existe et de connaître l'ampleur de son effet. Ainsi nous vérifierons :

- si l'indicateur de satisfaction est bien corrélé aux actions qui engendrent la satisfaction ;
- s'il est représentatif dans le cas où il détecte une différence entre unités ;
- et surtout s'il est sensible au temps, le tout en mettant en lumière l'impact, l'importance, de la discordance entre indicateurs de satisfaction et concept de satisfaction.

I – Les effets de contingence comme causes des problématiques

Les indicateurs de satisfaction comme leviers d'action permettent d'indiquer si la chaîne **qualité perçue → satisfaction → fidélité → profit** fonctionne comme il faut. Ils sont légitimes car le concept sur lequel ils reposent, la satisfaction, est accepté par tous entreprises/consommateurs. Un bon taux de satisfaction enverra un signal positif aux consommateurs. Un mauvais taux de satisfaction alertera les managers sur l'efficacité de leurs activités. Cependant les indicateurs de satisfaction présentent des caractères subjectifs et surtout ils sont le fruit d'une dépendance externe. Ils sont la conséquence des clients. C'est sur ce point que leur validité peut être mise en défaut.

Contrairement à d'autres indicateurs de performance, la satisfaction clients est mesurée à l'aide de sondages, le tout relatif à un plan de sondage souvent fait par magasin. Les opinions récoltées sont censées représenter la satisfaction des clients. Cependant la satisfaction reste un concept individuel et son agrégation au niveau d'une population ne peut qu'être affectée par les interactions des individus. Chaque individu sondé peut présenter des caractéristiques communes aux autres. Par exemple, comme le fait d'être habitant d'une même région ou encore fréquenter toujours le même magasin. Dans cette situation les opinions données peuvent être perturbées par ce conditionnement qui touchera l'indicateur comme une mesure agrégée. De plus, comme la satisfaction doit

répondre aux exigences d'un instrument de pilotage, des conditionnements par magasin/région peuvent limiter la portée de l'indicateur lorsqu'il est vu globalement²⁶. Le conditionnement peut être aussi interprété par des dépendances individuelles : nous pensons notamment aux variables sociodémographiques. Comme la fonction principale des indicateurs de performance est de permettre des comparaisons et d'attribuer les différences à des actions ou des événements identifiables, le problème principal ici est que les notes de satisfaction enregistrées reflètent les facteurs d'actions (politique marketing), mais aussi des facteurs personnels, sans compter les éléments de climat comme la région, le pays etc...

C'est en suivant ces quelques remarques que notre tâche s'est focalisée sur l'étude des indicateurs afin d'examiner si ces effets de contingences²⁷ ont un impact sur le processus de comparaison transversale, entre unité de gestion, et longitudinale, dans le temps. La comparaison transversale est la plus courante, elle consiste à étudier les différences entre les satisfactions clients enregistrées par unités de gestion (ex : entre magasins). L'analyse longitudinale concerne l'évolution des unités dans le temps. Afin d'arriver à cela nous avons étudié la pertinence des indicateurs de satisfaction face à l'existence de ces facteurs de contingence.

L'indicateur de satisfaction barométrique dans notre recherche est la résultante statistique d'une question qui concerne un attribut ou une satisfaction générale²⁸. D'ailleurs c'est de cette façon que les baromètres sont utilisés pour le pilotage, donc notre intérêt portera sur tous les éléments, attribut ou satisfaction générale, présents dans une étude barométrique de la satisfaction clients.

Pour effectuer ce travail, l'analyse a pris deux directions: la première touche les problématiques de comparaison des indicateurs par unité magasin, et la seconde touche l'évolution de ces indicateurs dans le temps. Notre but est de savoir si l'influence de l'environnement (régions, pays) ou les caractéristiques individuelles affectent ces comparaisons sur les deux perspectives transversales et longitudinales. L'idée est de

²⁶ Il arrive souvent qu'une satisfaction soit cumulée pour tous les consommateurs sans se soucier de la structure hiérarchique, plusieurs magasins, plusieurs régions, plusieurs pays etc.

²⁷ Ces effets sont attachés aux facteurs sociodémographiques et environnementaux. Ils sont qualifiés de contingences car leur effet est subissant à la satisfaction. Des preuves de leurs existences et un cadre théorique de leurs actions, seront présentés dans la suite.

²⁸ En règle générale dans des mesures barométriques, plusieurs questions relatives à la satisfaction sont présentes. Une question sur la satisfaction globale (ou générale) est posée aux interviewés accompagnée par d'autres relatives aux attributs de cette dernière : satisfaction concernant l'accueil, satisfaction vis-à-vis du personnels etc.

savoir si l'indicateur de satisfaction calculé sous forme d'une moyenne, taux ou score, est robuste face à l'existence de certains facteurs. Cette robustesse est nécessaire pour que les comparaisons (transversale et longitudinale) effectuées à travers des indicateurs soient valides. Nous étudierons ces comparaisons pour tester si l'indicateur de satisfaction est fiable, d'une part pour accomplir son rôle (comparaison) et d'autre part pour savoir s'il représente bien le concept de satisfaction microéconomique (situation de discordance ou non).

L'essentiel du projet de cette thèse, s'appuie sur l'idée d'exploiter l'ensemble des enquêtes réalisées par **INIT-conseil** depuis plusieurs années, ce qui représente un corpus d'enquêtes de l'ordre de milliers de personnes. Le départ de cette thèse vient d'une considération purement professionnelle émanant des besoins des praticiens se trouvant confrontés à des situations rarement décrites de façon complète dans la littérature marketing. Cette dernière concerne les variations dans les niveaux de satisfaction suivant certaines caractéristiques de mesure ou encore la stabilité dans le temps des indicateurs de satisfaction. Le travail de cette thèse concerne une investigation visant à contrôler le processus de comparaison temporelle et spatiale, le tout en donnant les pistes de compréhension des problèmes rencontrés en se basant sur l'exploitation des bases récoltées par la société partenaire (INIT). Comme notre problématique de recherche vise à faire un diagnostic des mesures de baromètres de satisfaction clients, le fait d'utiliser des baromètres comme instrument de recherche est naturel. Nous avons sélectionné deux baromètres relatifs à deux entreprises : Un grande enseigne de bricolage et une grande enseigne de prêt à porter. Le cumul en données disponibles pour le premier baromètre est de plus 400000 individus, et le deuxième est de plus de 175000. Comme on le verra dans la suite, le choix de ces deux baromètres était conditionné par la question traitée²⁹ et par la présence de l'unité magasin. Ces deux baromètres constituent les instruments de recherche afin de comprendre le fonctionnement de l'indicateur dans la pratique des managers face à l'existence des problématiques de comparaison.

L'objectif principal de notre recherche est d'apporter des éclaircissements sur les fonctions de comparaisons via les indicateurs de satisfaction dans le cadre barométrique, avec comme intention de tester la robustesse de ces derniers. Nous avons voulu savoir si l'indicateur de satisfaction comme résultante sur laquelle la comparaison est effectuée, est

²⁹ Un pour les problématiques longitudinales et l'autre pour les problématiques transversales.

assez résistance³⁰ face à l'existence des facteurs de contingence. Dans notre démarche, nous nous sommes focalisés sur deux types de facteurs : ceux attachés à l'environnement des individus (Région et pays) et ceux liés à ses caractéristiques sociodémographiques (âge et sexe). Toutefois la procédure générale, théorie et méthodologie, sera valable pour tous les facteurs présentant les mêmes propriétés³¹.

➤ **Les baromètres de satisfaction comme instruments de recherche**

Les facteurs de contingence sont des sources d'erreurs présentes dans une mesure de type barométrique attachée à une population de consommateurs. Essayer de prouver leurs effets ne peut passer que par l'exploitation des données barométriques. Cependant cette exploitation pose une première contrainte relative à la nature de la mesure. Dans le cadre barométrique, elle est souvent unidimensionnelle³², alors que le concept est très complexe pour être capté par un seul item. Bien sûr, il paraît très réducteur de considérer une seule question comme représentant du phénomène, néanmoins comme les facteurs contingents touchent la mesure elle-même, ils seront présents dans cette considération unidimensionnelle (une question globale de satisfaction) aussi bien que dans le principe multidimensionnel, plusieurs items considérés comme des manifestes qui résument le concept qui lui est considéré comme latent. La deuxième contrainte est que les questions de satisfaction générale ou par attributs dans les baromètres peuvent englober plusieurs concepts voisins comme la qualité, l'image, la performance perçue. Ainsi la mesure obtenue est un résumé englobant plus que la satisfaction. Cependant comme notre démarche est axée sur les facteurs de contingence, le fait d'utiliser les résultats des enquêtes de satisfaction barométrique pour prouver leur existence se justifie par le fait que leurs actions sont postérieures à la formation des concepts, donc touchant la mesure barométrique (la satisfaction microéconomique).

D'un autre côté, pour explorer les fonctions des indicateurs, les baromètres deviennent les instruments de prédilection parce que leurs résultats concernent les clients et le pilotage leur est destiné. De plus, pour voir si un indicateur est sensible dans le temps, les baromètres présentent l'avantage de capter la même information (même

³⁰ Nous pensons ici à la capacité de l'indicateur à fournir des informations discriminantes, pour que la comparaison soit fiable. Ceci renvoie aussi à la capacité de l'indicateur à être fidèle à ses qualités d'indice de performance.

³¹ Notre volonté est également de donner une procédure d'identification et de test, applicable à toute source qui perturbe l'indicateur de satisfaction lors de la mesure barométrique.

³² Une seule question souvent captée par une échelle de note de 1 à 10. L'unidimensionnalité est à opposer à la multi-dimensionnalité qui elle sous-entend l'utilisation de plusieurs questions pour capter un seul concept.

questionnaire) dans le temps en s'appuyant sur la représentativité des échantillons sélectionnés.

En résumé, notre problématique concerne les baromètres de satisfaction et les indicateurs produits, donc l'instrument de recherche ne peut être que les baromètres eux-mêmes. L'avantage principal des baromètres est le nombre de réponses (plus de 100 000 individus dans notre cas d'étude) ce qui confortera les étapes de l'analyse statistique.

Pour ce qui est de l'échantillonnage lors du sondage effectué, il est représentatif par construction, et la représentativité est respectée dans le temps (vague par vague). Pour notre recherche, nous avons récupéré des baromètres a posteriori et ces derniers ont prouvé leur efficacité dans la pratique. Comme notre sujet concerne un problème de comparaisons, le fait que les baromètres restent identiques dans le temps, et que les données recueillies soient conséquentes, atténue le problème des biais de mesure sur le sujet de notre problématique.

Section 3 – Structure de la thèse

L'architecture générale de cette thèse comprend 5 chapitres. Le premier consiste en état de l'art de la satisfaction et sa mesure barométrique. Ce chapitre permet de présenter entre autres l'intégration de la satisfaction comme un élément de contrôle de gestion sous forme d'indicateur de performance.

Le chapitre 2 définit la problématique de notre recherche, à savoir les problématiques rencontrées lors de l'utilisation des indicateurs de satisfaction, comme instruments de comparaison transversale et longitudinale. Une définition « mathématique » des facteurs de contingence est donnée, et un cadre de test (diagnostic des indicateurs de satisfaction) est spécifié à l'aide de trois propositions qui vont guider nos études empiriques.

Le chapitre 3 définit les études de cas et présente la méthodologie de traitement de ces données. À l'aide de méthodologies statistiques qui sont la méta-analyse et l'analyse de comparaison fiducio-bayésienne, une adaptation est effectuée pour tester nos propositions

Dans le chapitre 4, les résultats des tests des trois propositions sont présentés, et les conclusions mettant en défaut la représentativité et la sensibilité des indicateurs de satisfaction sont avancées.

Dans le chapitre 5, nous discutons un autre indicateur de performance qui est « le Net Promotor Score » (NPS) et de sa relation avec l'indicateur de satisfaction clients. Ce chapitre a pour but d'éclaircir les fonctions du NPS et nous montre l'insuffisance de ce dernier à juger de la performance, issue de l'axe clients.

Le document sera clôturé par une discussion générale ainsi qu'une conclusion générale. L'architecture de ce travail doctoral est présentée dans la figure suivante :

<i>Introduction Générale</i>	
<p>Chapitre 1. La satisfaction clients et sa mesure barométrique</p>	<ul style="list-style-type: none"> - Définition du concept de satisfaction déterminants et conséquences -La mesure barométrique de la satisfaction -La satisfaction dans le contrôle de gestion : l'indicateur de performance
<p>Chapitre 2. Facteurs de contingence face à l'évaluation des indicateurs</p>	<ul style="list-style-type: none"> -Présence des facteurs de contingence dans la mesure de satisfaction -Cadre explicatifs des facteurs de contingence -L'indicateur de satisfaction et facteurs de contingences -L'usage des indicateurs de satisfaction comme instruments de comparaison, formulation des propositions de test.
<p>Chapitre 3. Méthodologie de test des propositions</p>	<ul style="list-style-type: none"> -Présentation des données de l'analyse -les techniques des méta-analyses -Méthodologie fiducio-bayésienne d'analyse -Proposition de méthodologie de test des propositions par double procédé : méta-analyse et procédures fiducio-bayésiennes
<p>Chapitre 4. Résultats de test des propositions</p>	<ul style="list-style-type: none"> -Définitions des indices traduisant les propositions -Analyse des variations des indices et facteurs de contingence : test des propositions
<p>Chapitre 5. L'indicateur de satisfaction face à son principal concurrent le cas du Net Promotor Scor</p>	<ul style="list-style-type: none"> - L'émergence du Net Promotor Score face au taux de satisfaction - L'analyse des liaisons entre les deux indicateurs - Discussion sur leurs exploitations communes
<i>Discussion Générale</i>	
<i>Conclusion Générale</i>	

Chapitre 1 : La satisfaction clients et sa mesure barométrique

Depuis plus de 30 années de recherche en Marketing, le concept de satisfaction a bien évolué et semble atteindre son achèvement en termes de définition et de caractérisation du concept (déterminants et conséquences). La littérature scientifique a parfaitement pensé à la notion de satisfaction et a surtout dégagé une structure explicative de sa réalisation. Depuis les travaux fondateurs (Anderson 1973 ; Hunt 1977 ; Oliver, 1980 et 1981) la conceptualisation a permis de prendre conscience de l'importance de la satisfaction des consommateurs comme un paramètre clé menant au profit. Cette correspondance est le fruit d'un schéma mental qui se décline en trois équivalences (Boss, 1993) :

1. La qualité du service/produit, fourni par l'entreprise, se traduit par la satisfaction
2. La satisfaction se traduit par une meilleure fidélisation de la clientèle
3. Une clientèle fidèle est plus profitable pour l'entreprise à court terme et à long terme.

La satisfaction est devenue un des thèmes majeurs dans lequel les recherches se sont multipliées pour mieux éclaircir la théorie et encadrer la pratique ; c'est-à-dire mesurer la satisfaction. Pour mesurer cette dernière, les chercheurs se sont directement intéressés à en modéliser le processus de formation. Ils ont établi une structure sous la forme d'un modèle d'écart entre les attentes et l'évaluation de la qualité du produit ou du service (Oliver, 1980). Cette schématisation a donné lieu au paradigme dit de « disconfirmation » qui s'est imposé comme le modèle dominant expliquant le mécanisme de la satisfaction. Même si ce modèle a subi quelques changements au cours des années, il demeure néanmoins le support expliquant la formation de la satisfaction clients.

Sur cette lancée, la pratique de la satisfaction clients s'est dotée d'un instrument de mesure connu sous le nom de baromètre qui vise à suivre la satisfaction dans le temps et à expliquer sa formation. Il est devenu un outil essentiel pour toute entreprise soucieuse de l'avis de ces clients. Il permet aux acteurs décisionnels dans chaque entreprise, d'orienter leurs actions pour satisfaire et fidéliser leurs clients.

Section 1 – Le concept de satisfaction

De manière spontanée, la satisfaction peut se définir comme l'incarnation d'un sentiment positif exprimé par les consommateurs, suite à l'achat ou utilisation d'un produit /service. Elle correspond à un état affectif ressenti par l'individu, quand il obtient ce qu'il souhaite. De manière plus formelle dans une vision Marketing, cette définition reste insuffisante, vu l'aspect latent du concept. Le caractère latent fait référence aux divers processus psychologiques relatifs à ce concept. A relier toutes les définitions proposées, la satisfaction a surtout été attachée soit à des concepts purement émotionnels (Westbrook, 1980, Woodruff, Cadotte et Jekins, 1983, Fournier et Glen Mick, 1999), soit à une pure comparaison de nature cognitive (Bloemer et Kasper, 1995, Churchill et Suprenant, 1982). Cette distinction a amené des chercheurs à considérer plusieurs définitions de la satisfaction comme concept. Au lieu d'une définition générale de nature émotive ou cognitive, les définitions créent une typologie qui situe le concept à plusieurs niveaux.

I – Les types de satisfaction

En 2002, Joëlle Vanhamme a formulé la satisfaction selon deux plans : vertical et horizontal, qui se complètent pour distinguer les différentes formes. Le plan horizontal porterait sur le stade de l'expérience de consommation/achat sur lequel le jugement de satisfaction est donné. Cette dimension représente tous les éléments qui interviennent au moment où pendant la consommation/ou l'achat. L'auteur donne trois caractéristiques de ce plan horizontal :

- La satisfaction par rapport aux éléments engendrés lors de l'expérience.
- La satisfaction spécifique à la transaction (l'achat ou l'expérience avec le produit ou service).
- La satisfaction par rapport au niveau de satisfaction obtenu.

D'autre part, l'auteur a défini un plan vertical qui renvoie quant à lui, aux formes qui sont agrégées à partir de la satisfaction spécifique à la transaction elle-même. Cette hiérarchisation est décrite par Anderson et Fornell (1994) et par Oliver (1997) et Elle inclut quatre situations qui sont :

- La satisfaction spécifique à une transaction
- La satisfaction par rapport à la marque, ou satisfaction relationnelle
- La satisfaction microéconomique
- La satisfaction macroéconomique

A – La satisfaction relative à une transaction

La satisfaction client spécifique à une transaction, envers un produit et/ou service repose sur trois éléments qui sont la subjectivité, la relativité et l'évolutivité.

La satisfaction est subjective : La satisfaction des clients dépend de leurs perceptions des produits/services, et de leurs vécus en tant que consommateur.

La satisfaction est relative : Cet aspect est le résultat du jugement de la satisfaction comme une comparaison entre l'expérience du consommateur (subjective) et un standard de comparaison (Ngobo 1998).

La satisfaction est évolutive : La satisfaction varie avec le temps en fonction des attentes et des standards de comparaison, et du cycle d'utilisation des produits. En d'autres termes pour les clients, les standards de référence évoluent avec le temps, ce qui entraîne une modification de la satisfaction vu son caractère relatif. Ces trois éléments donnent la définition de la satisfaction relative à une transaction spécifique, définie par Joëlle VANHAMME, comme : « *la satisfaction peut être vue comme un état psychologique résultant d'une expérience d'achat et/ou de consommation* ».

B – La satisfaction relationnelle/cumulée ou par rapport à une marque

Elle est relative à plusieurs occurrences du même type d'expérience. Elle est qualifiée de satisfaction de long terme et englobe une évaluation groupée de toutes les expériences de consommation et /ou d'achat (Oliver 1997).

C – La satisfaction cumulée, microéconomique et macroéconomique

La satisfaction cumulée traduit une évaluation générale de l'expérience de consommation pour un ensemble de consommateurs d'un produit donné (Fornell, 1992). Elle renvoie à la satisfaction cumulée de tous les consommateurs d'une entreprise « baromètre de satisfaction » (microéconomique), ou à une satisfaction cumulée des consommateurs d'une même culture/société (macroéconomique).

La distinction sur les deux axes, vertical et horizontal de la satisfaction, permet de bien spécifier ce qui est mesurable de ce qui ne l'est pas. Ainsi comme la distinction entre les différents niveaux est très difficile dans la pratique (étude quantitative), la majorité des recherches étudiant les causes (antécédents de satisfaction) se sont intéressée à la satisfaction spécifique à une transaction. Elle reste non agrégée et vise le niveau

individuel³³. Cependant en pratique³⁴ pour mesurer la satisfaction, on la considère sous sa forme agrégée (suivant un segment par exemple) car elle présente l'avantage de donner une mesure directe et générale sur l'ensemble des consommateurs à l'égard du produit ou du service.

A travers ces éléments, nous pouvons remarquer que la satisfaction relative à une transaction, est définie comme un concept non directement observable, d'où la confusion qui peut se faire avec d'autres concepts voisins comme l'attitude, ou encore la qualité perçue.

En somme la satisfaction dans la littérature a été considérée dans sa version relative à une transaction. Sa définition s'est établie autour de l'idée qu'elle puisse résulter de processus affectifs et cognitifs qui peuvent interagir entre eux. Cette considération est aujourd'hui établie dans le domaine de la recherche sur le concept, sans autant qu'une distinction claire soit attribuée à telle ou à telle dimension dans le processus dual : affectif et cognitif.

II – La satisfaction et les concepts voisins

La satisfaction n'est pas le seul concept qui a intéressé les recherches en Marketing pour expliquer le comportement des consommateurs. D'autres concepts ont dû coexister avec la satisfaction et leur développement n'a pas pu être fait indépendamment de la satisfaction. Ce fait a créé certaines confusions entre la satisfaction, l'attitude, la qualité perçue ou encore la performance perçue. Cependant ces confusions ne sont qu'accessoiries et les recherches sur le sujet ont bien établi des différences entre le concept de satisfaction et ses voisins.

A – La satisfaction et l'attitude

L'attitude, pourra être vue comme « une tendance psychologique s'exprimant par l'évaluation d'un objet particulier avec un certain degré de faveur ou défaveur » (Eagly et Chaiken, 1993). Elle traduit la disposition que l'individu peut avoir avant l'accomplissement de certaines tâches (ici évaluer le produit /ou le service). L'attitude vise donc une évaluation relativement stable, ne nécessitant pas forcément d'expériences spécifiques avec le produit/service. La satisfaction du consommateur quant à elle, même

³³ Elle concerne la formation de la satisfaction au niveau de la psychologie de l'individu. Le résultat concerne donc l'individu en tant qu'unité.

³⁴ La pratique dans ce cas renvoie au pratique barométrique. (Enquêtes de satisfaction clients d'une entreprise)

si elle rentre dans le cadre d'une évaluation d'un objet, reste transitoire et spécifique à l'acte d'achat ou la consommation et elle ne peut exister avant. Cependant les deux concepts sont liés. Selon Oliver (1981), la satisfaction s'éteint avec le temps, avec l'extinction de l'activation liée à l'expérience. Elle peut se confondre avec l'attitude générale vis-à-vis de l'achat. Nous voyons ici que la satisfaction se distingue de l'attitude en ce qu'elle est situationnelle. Néanmoins les attitudes jouent un rôle de modérateur agissant sur les attentes préalables à l'expérience d'achat/consommation, qui influencent à leur tour la satisfaction. L'attitude et la satisfaction, relative à une transaction, restent distinctes contrairement à la satisfaction cumulée, sous ses différentes formes qui est souvent assimilée à une attitude.

B – La satisfaction et la qualité perçue

Pour la majorité des spécialistes, la qualité perçue du service résulte de l'adéquation entre les « Attentes » du client et la « performance perçue » vis à vis du service (Prasuraman et Zeithaml et Berry 1985). Cette définition de la qualité perçue a créé la confusion avec le concept de satisfaction, qui lui aussi est considéré comme « l'état d'un client résultant d'un jugement comparant les performances d'un produit au niveau de ses attentes ».

Cependant, il existe bien une distinction entre les deux concepts qui consiste dans le fait que la qualité est un jugement évolutif global vis-à-vis d'un produit ou d'un service concernant sa supériorité par rapport à un standard d'excellence. La satisfaction résulte d'un jugement évaluatif sur le produit ou le service à condition d'avoir une expérience avec le produit. De plus, d'un point de vue psychologique, la satisfaction est un état propre à l'individu (dual affectif et cognitif) et la qualité perçue est un état décrivant le produit ou service en jugeant son excellence (résultat purement cognitif) (Vanhamme, 2002)

C – La satisfaction et la performance perçue

La performance perçue a été définie comme étant « le niveau perçu de résultat du produit ou service, habituellement mentionné sur une échelle objective bornée par des niveaux de bonnes et mauvaises performances. » (Oliver, 1997). Dans les premiers travaux sur la satisfaction (Cardozo, 1965 ; Olshavski et Miller, 1972 ; Anderson, 1973), la satisfaction a été confondue au concept de performance perçue par le consommateur. Cependant bien après elle s'est différenciée de la performance perçue sur le fait qu'elle est consécutive de

cette dernière qui est incluse dans le processus de comparaison avec les attentes. La performance perçue a été introduite dans le paradigme de la « disconfirmation » des attentes comme étant un antécédent de la satisfaction. D'autres recherches l'ont même présentée comme l'unique déterminant de satisfaction, se substituant ainsi au processus de « disconfirmation » des attentes »(Churchill et Surprenant, 1988). D'autres éléments de distinction se situent sur le fait que la satisfaction est une résultante affective et cognitive et que la performance perçue est de nature purement cognitive.

La satisfaction dans son aspect spécifique à une transaction, pourrait être vue comme un composite cognitif-affectif. Ceci est l'élément clé qui permet d'isoler le concept de satisfaction de ses voisins.

Dans la pratique, la satisfaction se trouve englobée avec d'autres concepts et reste difficilement identifiable. Néanmoins, d'un point de vue académique, les chercheurs se sont attelés à dégager ses déterminants et ses conséquences sous forme d'objets psychologiques (attente, valeur perçue, émotion, attribution, équité etc.). Cela a permis de bien déterminer les facteurs de sa formation et de comprendre ses effets sur le rachat, le bouche-à-oreille favorable et la fidélité.

Section 2 – Modèle de la formation de la satisfaction : ses déterminants et conséquences

Les premiers travaux visant à cerner les antécédents de la satisfaction des consommateurs, se sont intéressés à analyser les liens entre les attentes et la perception de la performance des produits ou services. Cette modélisation a donné naissance à une comparaison de type cognitif et s'est approfondie au cours des années en incluant d'autres antécédents que les attentes et la performance perçue, tels que les réactions émotionnelles, l'équité, l'attribution, la qualité et la performance perçue. Le modèle cognitif de base connu sous le nom de « disconfirmation », se voit argumenter par des attributs affectifs ce qui a amené à considérer le concept de la satisfaction comme un dual entre un aspect cognitif et affectif. Le côté affectif se traduit comme un antécédent au même titre que les attentes préalables et la disconfirmation des croyances. Lorsque l'émotion ressentie laisse des traces affectives, il s'intègre lui aussi au jugement évaluatif (Westbrook et Oliver 1991). Ce ressenti affectif, devient donc une composante de la satisfaction et reste

insuffisant pour l'expliquer complètement, contrairement à certaines définitions préconisées par Bagozzi et al (1999).

I – Le modèle cognitif de base

Selon le paradigme de « disconfirmation » ou non–confirmation présenté par (Oliver 1980), la satisfaction d'un consommateur est la résultante de la comparaison entre les attentes avant la consommation/ ou l'achat d'un produit/ou service, et leurs observations concernant les performances de ces derniers. Cet écart entre les attentes et performance perçue forme une perception de la non-confirmation. Cette dernière peut être neutre quand les attentes sont au niveau de la performance perçue, positive si cette dernière est supérieure, négative si les attentes dépassent la performance perçue. Ainsi ce modèle de « disconfirmation » comporte trois éléments qui sont :

Les attentes préalables à la consommation.

Le jugement porté sur la performance

La « disconfirmation » ou la non-confirmation qui résulte de la comparaison entre performances et attentes.

Ainsi le processus de non-confirmation quand il est positif entraîne une satisfaction et quand il est négatif une insatisfaction. Ici la non-confirmation joue le rôle d'éléments modérateurs amenant soit à une satisfaction soit à une insatisfaction. Cependant ce modèle a été critiqué sur ses insuffisances qui concernent un seul standard de comparaison (attentes/valeur perçue), et sur le fait que la disconfirmation est vue comme la variable clé médiatrice amenant la satisfaction ou l'insatisfaction.

II – Le modèle cognitif amélioré

Il s'agit d'une modification, pour conjurer l'hypothèse que seule la non-confirmation est le médiateur de l'influence des attentes et la valeur perçue sur la satisfaction. Ainsi différents chercheurs ont montré que les attentes ont des effets directs sur la satisfaction (Anderson 1973, Oliver et Linda 1981, Churchill et Surprenant 1982). D'autre part il a été aussi montré qu'il existe un lien direct entre la performance et la satisfaction (Oliver et Sarbo 1988 ; Oliver 1993). Yi (1993) expose que les attentes jouent un rôle important en l'absence de critères objectifs pour évaluer le produit, et que

la performance à un impact important lorsque le jugement du produit par le consommateur est aisé.

On a pu voir dans d'autres recherches (Churchill et Surprenant 1982 ,Tse et Wilton, 1988 , Bolton et Drew 1991) qu'il existe des liens directs entre la performance sur la satisfaction sans passer par la non confirmation. Cependant d'autres études préconisent l'existence de liens directs et indirects (via la disconfirmation) entre les attentes et la performance sur la satisfaction (Oliver et Desarbo, 1988).

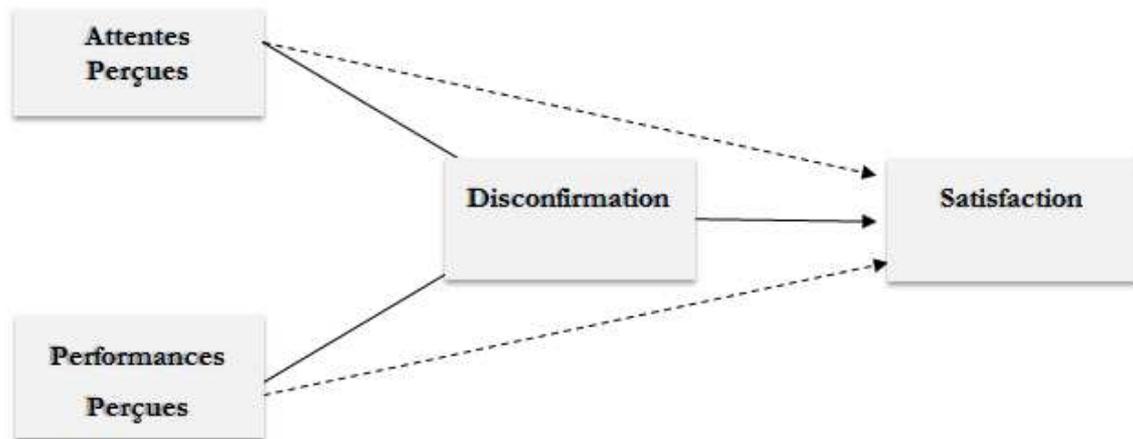


Figure 7: Modèle de disconfirmation

Source: Churchill et Surprenant (1982): "An Investigation Into the Determinants of Customer Satisfaction" Journal of Marketing Research,Nov, p. 492.

Remarque: lors de la mesure, la « disconfirmation » se distingue en deux types : la non confirmation objective et la non-confirmation subjective. La première est la résultante d'une comparaison algébrique entre attentes et performance. Quant à la deuxième, elle est produite en posant la question directement au consommateur, en demandant au interviewé un score sur une échelle de type « mieux/moins bien attendu ». Ces deux formes se présentent comme un problème de mesure de la non confirmation.

III – Autres antécédents cognitifs de la satisfaction

Il s'agit ici des déterminants ajoutés aux modèles de « disconfirmation ». Parmi ces déterminants on retrouve l'équité et l'attribution, la qualité et la valeur perçue.

A – L'équité comme antécédent à la satisfaction

L'équité ou encore de la justice distributive est une théorie issue de la psychologie sociale dans le domaine de l'organisation et l'échange social. Les premiers travaux de recherche sont attachés à Homans (1961). L'auteur s'est intéressé à l'analyse des mécanismes psychologiques et économiques liés au comportement de l'individu quand il est récompensé ou puni par le comportement d'un tiers lors d'un échange.

Comme l'acte de consommation/achat met en scène deux parties : consommateur et entreprise, l'équité est devenue une variable à considérer pour expliquer la satisfaction, à tel point que des chercheurs comme Oliver et Desarbo (1988) ont estimé qu'elle est le quatrième déterminant le plus significatif après la disconfirmation, la performance et les attentes.

L'équité se fonde sur la comparaison entre les coûts engagés dans la transaction (input), et les bénéfices attendus pour chaque partie de l'échange (output). Elle possède donc un caractère visant à évaluer le rapport (input/output). Selon cette théorie l'acheteur est censé comparer son propre rapport (input/output), à celui d'un autre acheteur qui a acheté le même produit ou a reçu le même service (Oliver Swan 1998). En d'autres termes, le consommateur juge si l'échange en cours lui a permis d'être équitablement traité ou s'il a été traité de façon non équitable.

L'équité a été considérée comme un antécédent à la satisfaction au même titre que la disconfirmation (Swan Oliver 1985 ; Smith, Bolton et Wagner 1999). De manière générale, si les consommateurs se sentent équitablement traités, ils seront satisfaits. Il faut rappeler que le sentiment d'équité est basé sur un rapport perçu. Cette perception, peut être aussi influencée par le prix à payer, les efforts consentis, les expériences préalables (Tse, Wilton 1988).

B – L'attribution comme antécédent

L'attribution peut se définir comme l'affectation d'un individu donné, le fruit de son comportement ou celui des autres, à des causes internes ou externes. Cette théorie qui était à l'origine introduite pour expliquer le comportement des réclamations, a été depuis étendue pour devenir un antécédent direct de la satisfaction (Folkes, 1984). Wiener (1986) rapporte que le « succès », l'achat d'un bon produit, et « l'échec », l'achat d'un produit défectueux, génèrent des inférences causales qui peuvent être classifiées en fonction de trois dimensions : la stabilité, la contrôlabilité et le lieu de causalité. La stabilité indique si la cause de l'évènement est permanente ou temporaire, et informe le

consommateur quant à la possibilité que le même évènement se reproduise. La dimension de lieu de causalité, spécifie si les causes sont externes (entreprise) ou internes (le consommateur lui-même). Enfin, la dimension de contrôlabilité indique si ces causes sont volontaires ou involontaires.

Dans le domaine de la satisfaction, l'attribution consiste à identifier des causes pour justifier l'échec/ou pas du produit à satisfaire. Selon la dimension de l'attribution (stabilité, lieu de causalité, contrôlabilité), les effets sur la satisfaction varient. Richins (1983), indique que les « échecs » conduisent à une insatisfaction plus marquante si l'attribution est externe que si elle est interne (dimension de lieu de causalité). Bitner (1990) a montré empiriquement que la stabilité et le contrôle influent négativement sur la satisfaction. Les consommateurs sont insatisfaits lorsqu'ils pensent que les fournisseurs ont un contrôle sur la raison du mauvais service et que cela n'est pas dû à quelque chose d'involontaire. Aussi l'insatisfaction s'aggrave quand le consommateur estime que les raisons du mauvais service sont stables.

C – La qualité perçue comme antécédent

Étant donné le lien étroit entre la satisfaction (S) et la qualité perçue (Q), les chercheurs se sont intéressés à définir le sens de causalité entre ces deux concepts. Sureshchander et al (2002) rapportent que les deux concepts sont différents mais fortement corrélés. Cependant le sens de leur liaison, n'est pas vraiment établi. Bitner (1990), Drew et Bolton (1991) sont pour considérer la satisfaction comme antécédent de la qualité $S \Rightarrow Q$, alors que Cronin et Taylor (1992), ou encore Parasuraman, Zeithaml et Berry (1988) tranchèrent pour le sens inverse ($Q \Rightarrow S$). Selon Teas (1993), ces résultats ne sont pas nécessairement contradictoires, et le sens dépend du cadre d'analyse. Ainsi le sens $Q \Rightarrow S$ semble plus adapté à une approche transactionnelle (une transaction spécifique) car le jugement de la qualité peut préexister à la satisfaction, il ne requiert pas forcément une expérience d'achat ou de consommation. D'autre part, la liaison $S \Rightarrow Q$ est plus valable dans une approche relationnelle (plusieurs transactions), la qualité devenant fonction de la satisfaction.

D – La valeur perçue comme antécédent

La valeur perçue est le résultat d'une comparaison entre les bénéfices et les coûts perçus (l'effort financier et psychologique). Evrard et Ngoala (1998) définissent la valeur

perçue comme un concept purement cognitif. Comme ce concept reste proche de la satisfaction spécifique à une transaction dans sa nature cognitive, mais différent dans sa nature psychologique, il a été considéré comme un antécédent de la satisfaction (Rust, Oliver 1994). Une relation positive existe entre ces deux concepts (Huber Herman 2000). Cependant, le lien de causalité pourrait être inversé à l'instar de la qualité perçue dans une approche relationnelle, comme le précise Oliver (1999). La valeur perçue de type relationnelle est une valeur cumulée, donc la satisfaction qu'on peut tirer de chaque expérience prise séparément peut agir positivement sur la perception de la valeur cumulée. Ainsi la satisfaction peut être considérée comme antécédent de la valeur perçue relationnelle.

E – Les dimensions affectives comme antécédent

Les éléments affectifs englobent l'ensemble des réactions émotionnelles telles que : l'humeur, l'optimisme/pessimisme, la satisfaction envers la vie, le mécontentement généralisé etc. Leurs effets sur le comportement des consommateurs touchent essentiellement le traitement de l'information (Batra et Ray, 1986, Holbrook et Barta 1987). Pour ce qui concerne la satisfaction des consommateurs les éléments affectifs ont été introduits pour combler le manque d'explication du paradigme de non-confirmation sous son aspect purement cognitif. Westbrook, (1980,1987) a montré que les réactions émotionnelles sont des déterminants et améliorent le pouvoir explicatif du modèle cognitif et leurs intégrations comme variables accroissent significativement l'explication de la satisfaction.

Plusieurs recherches ont étudié les effets des dimensions affectives et ont énuméré la nature des effets. Par exemple, l'étude de Spies, Hesse et Loesch (1997) a confirmé l'existence d'un effet de l'humeur sur la satisfaction. Les auteurs affirment que les clients ont des scores de satisfaction élevés lorsqu'ils sont de bonne humeur et inversement. Cependant, il a été montré que l'humeur n'a pas la même influence pour tous les types de produit /service et dépend de l'échelle de mesure de satisfaction utilisée (Wesbrook, 1980 ; Brockman, 1998). L'émotion a elle aussi été étudiée et son influence sur la satisfaction a été mise en avant (Babin, Griffin ,1994). WestBrook, (1987) a montré que les émotions positives ou négatives accroissaient le pouvoir explicatif du modèle de « disconfirmation ». Depuis les chercheurs ont mis en évidence deux issues pour les liens entre les réactions émotionnelles et la satisfaction. La première postule que les réactions

émotionnelles ont des effets parallèles et indépendants du modèle de « disconfirmation » des attentes (Wesbrook, 1987 ; DubéRioux, 1990). La seconde conception, est résumée par les chercheurs comme une séquence d'effets : « disconfirmation » des attentes => réaction émotionnelle => satisfaction (Muller, Tse et Venkatasubramaniam, 1991 ; Oliver 1993 ; Oliver Rust et Varki 1997).

Il est aussi nécessaire de rappeler que l'émotion a souvent été confondue avec satisfaction, jusqu'à définir la satisfaction comme une émotion parmi d'autres. Cependant des chercheurs comme Cohen et Arseni, Wesbrook et Oliver (1991) ont rapporté que les réactions émotionnelles lors de l'expérience laissent des traces affectives qui sont à leur tour intégrées dans le jugement évaluatif d'un produit ou service, au même titre que les éléments cognitifs : les attentes préalables, et la « disconfirmation » des attentes. Donc la satisfaction, ne peut pas être ni purement cognitive ni exclusivement de nature émotionnelle, mais une sorte de dual **affectif-cognitif** qui se complètent lors de sa réalisation.

IV – La satisfaction et l'insatisfaction

Traditionnellement la satisfaction est considérée comme un jugement fondé sur la comparaison entre les attentes et les performances perçues du produit ou service. Sous cette forme la satisfaction d'un consommateur devient la résultante d'une comparaison positive (attentes > performance perçue), et l'insatisfaction est la conséquence d'une comparaison négative (performance perçue > attentes). Ce schéma considère que la satisfaction et l'insatisfaction sont opposées tout en appartenant à un même plan (continuum unidimensionnel). Cependant d'autres recherches considèrent une vision différente basée sur un plan bi factoriel qui est motivé par le fait que ce n'est pas parce qu'un client n'est pas insatisfait sur un critère X, qu'il est nécessairement satisfait. Cette vision est justifiée par le fait que les dimensions des déterminants de la satisfaction et l'insatisfaction ne sont pas les mêmes. Cependant les recherches basées sur ce concept (Swan et Combs 1976 ; Maddox, 1981 ; Mackoy et Spreng, 1995) ne permettent pas d'établir une différence notable entre les deux visions. La distinction entre ces deux formes, est surtout basée sur la distinction entre les critères qui engendrent la satisfaction et l'insatisfaction. Par exemple, si un produit et/ou un service remplit un critère de base, il peut ne pas engendrer de la satisfaction, mais si ce critère disparaît c'est l'insatisfaction assurée. Ainsi on pourra considérer ces deux visions comme complémentaires pour

comprendre la formation du concept de satisfaction et de l'insatisfaction, en admettant qu'ils ne sont pas nécessairement opposés, mais proviennent d'éléments différents, par le fait que les éléments destinés à réduire l'insatisfaction ne sont pas généralement les mêmes que ceux cherchant à optimiser la satisfaction.

V – Conséquences de la satisfaction

L'intérêt de la satisfaction se rapporte surtout à ces conséquences. Depuis les travaux fondateurs, la satisfaction a été attachée à la fidélité clients et le bouche-à-oreille positif (recommandation de l'entreprise), et l'insatisfaction a été attachée au départ clients, à la réclamation et au bouche-à-oreille négatif (prévenir des amis et des membres de la famille de son insatisfaction). Ce type de conséquences a fait l'objet de différents travaux et les chercheurs ont permis de statuer sur le lien entre la satisfaction/insatisfaction et l'ampleur de la réaction qui suit la réalisation du phénomène.

A – La Fidélité

Cette relation entre satisfaction et fidélité est le sujet le plus étudié depuis les premiers travaux sur la satisfaction. Plusieurs chercheurs rapportent une relation linéaire liant les deux concepts à travers des tests empiriques liant positivement la satisfaction à la fidélité (Bitner, 1990, Taylor et Baker, 1994; Rust, Zahorik, and Keiningham 1995 ; Fornell *et al.*, 1996; Host et Knie-Andersen, 2004). D'autres plus récemment, rapportent une relation non linéaire (Ngobo, 1999; Mittal & Kamakura, 2001; Homburg, Koschate, & Hoyer, 2005) ou encore une relation modulée par les caractéristiques des individus, du produit et celle du marché (Homburg et Giering 2001, Songting Dong & all 2011) Cependant dans tous les travaux, et depuis les premières recherches sur la question (Howard et Sheth 1969, Cardozo 1973, Howard 1977, Oliver 1980), le lien entre la satisfaction s'est imposé naturellement avec une approche partagée qui est celle de la relation positive entre la satisfaction et la fidélité.

B – Bouche-à-oreille positif/ négatif

La deuxième conséquence naturelle de la satisfaction est de parler de son expérience à autrui. Anderson (1998) avance que l'insatisfaction conduit à favoriser le bouche-à-oreille négatif. D'autres chercheurs à l'exemple de Richins et Bloch (1986) pensent que cette relation peut être modulée par l'expertise du consommateur et son

implication envers le produit. La liaison est plus manifeste quand l'expertise et l'implication sont fortes. D'autres études comme celle de Singh (1990), avancent que l'attribution de cause de l'échec est facteur modulateur de la liaison entre l'insatisfaction et le bouche-à-oreille négatif. Ainsi, la tendance des clients à parler à d'autres de leur insatisfaction est plus importante dans le cas où la responsabilité du problème est imputée au fournisseur (une cause externe), contrairement au cas où la responsabilité du problème est attribuée à soi-même (une cause interne comme par exemple, une mauvaise utilisation du produit).

C – La relation entre la satisfaction et le bouche à oreille positif

Pour ce qui est de la relation entre la satisfaction et le bouche-à-oreille positif, l'exemple de Kraft et Martin (2001) propose que le bouche à oreille positif soit fonction de l'enchantement (une satisfaction très élevée), des bénéfices attendus, de l'implication dans le produit/service, des normes sociales, et de facteurs personnels et situationnels. Ainsi la satisfaction/insatisfaction est rattachée à la recommandation positive/négative, le tout avec des facteurs modulateurs de l'intensité de la relation.

D – La réclamation

Le phénomène de réclamation a été rapproché à l'insatisfaction des consommateurs. Le comportement de réclamation peut être défini comme : « une action entreprise par un individu qui entraîne la communication d'éléments négatifs concernant un produit ou un service vers un producteur, un distributeur ou une troisième entité » (Ladhari 2005). Cette réaction est adjacente au phénomène de départ clients. Un phénomène de départ serait plutôt corrélé à l'insatisfaction et se situerait comme le contraire de la fidélité.

Comme le rapporte Ladhari (2005) la littérature sur le comportement de réclamation montre qu'il y a prédisposition des consommateurs insatisfaits à s'engager dans un comportement de réclamation. Ce comportement est influencé par plusieurs facteurs incluant l'attribution de la cause de l'échec du produit à satisfaire, les attentes ainsi que les caractéristiques psychologiques, culturelles et démographiques.

La mesure de satisfaction sous-entend deux processus : un processus de construction d'échelles par des questions capables de capter la satisfaction et un processus de calcul statistique qui résume les informations sous forme d'indicateurs de satisfaction.

Dans un cadre barométrique, la mesure prédominante est dite subjective qui s'intéresse à sonder l'expérience des clients. Ces mesures élaborées à partir des enquêtes répétées dans le temps, visent d'une part à capter le niveau de satisfaction et d'autre part à expliquer la variation de ce niveau en fonction d'un ensemble d'attributs de la satisfaction. Il s'agit en effet d'interroger les clients sur leur satisfaction générale envers le produit/et ou service, en leur demandant un arbitrage, exprimant un avis de satisfaction sur les différentes dimensions qui caractérisent le service ou le produit (ex : rapport qualité/prix, l'accueil, service après-vente). En procédant ainsi, les entreprises obtiennent des indicateurs statistiques sur lesquels le suivi de l'évolution du niveau de satisfaction est basé. Ils obtiennent aussi un résumé des informations par attribut, qui modulent la satisfaction générale. Ces résultats servent aux entreprises pour orienter les améliorations qu'ils considèrent comme prioritaires.

En règle générale, la mesure de satisfaction clients dans les baromètres consiste à établir des questionnaires mono-items (une idée représentée par une question suivant une échelle) exprimant à chaque fois soit une satisfaction générale soit une satisfaction envers un attribut. Pour ce faire, les questions sont souvent formulées comme : « Comment évaluez-vous le degré de satisfaction ? », soit au global, soit par attribut. La mesure barométrique est une mesure de diagnostic très éloignée du modèle académique, où l'intérêt est porté sur la définition d'un modèle expliquant la formation de la satisfaction en identifiant ses antécédents et ses conséquences. Il faut dire aussi que la mesure barométrique s'intéresse plus à un niveau de satisfaction par segment (une population de clients), là où le monde académique établit des mesures de satisfaction relative à une transaction. La mesure dans des baromètres est une composante qui peut être attachée à une population alors que la mesure académique est une mesure psychométrique qui a pour but d'expliquer le phénomène au niveau de l'individu.

Dans le monde académique, l'élaboration de la mesure est multi-items (Evrard, Audrain 1998). Cette dernière répond mieux aux exigences psychométriques de la fiabilité³⁵,

³⁵ La fiabilité dans ce cas signifie la construction d'échelles qui représentent le concept psychologique le plus fidèlement possible et sans biais.

contrairement aux mesures barométriques où les échelles sont rarement contrôlées pour rejoindre ces exigences, à l'exception de certains baromètres nationaux comme L'ACSI³⁶. Ceci nous laisse penser que les données barométriques ont été rarement utilisées dans le domaine de la recherche à cause du manque d'adéquation des échelles au domaine académique. Ce fait, qui est bien-sûr préoccupant, nous paraît insuffisant pour exclure les baromètres d'une utilisation orientée recherche. De plus, dans les baromètres, les questionnaires restent identiques par vague, ce qui leur donne l'avantage d'être longitudinaux (représentant le temps) et de posséder une importante quantité de données. La quantité d'informations recueillie peut avoir un effet de diminution des erreurs aléatoires de mesure, mais sans diminuer les erreurs systématiques de toutes sortes. Ainsi en considérant que la fiabilité des échelles est une erreur systématique parmi d'autres, du fait que les instruments recueillis par vague sont identiques (même questionnaire), il y a possibilité de détecter les erreurs systématiques et de pouvoir raisonner en fonction. Cela peut nous amener à utiliser les baromètres de satisfaction comme des instruments de recherche, en commençant bien sûr par identifier les sources d'erreurs pour augmenter leur fiabilité au sens large. Un autre point aussi qui nous semble en faveur de l'utilisation des baromètres comme instrument de recherche reste le caractère de ces résultats, c'est à dire la production des indicateurs de satisfaction.

I – L'usage des baromètres de satisfaction

Les plans d'action marketing visant à améliorer la relation clients et à garantir un **bon échange entreprises/clients** sont axés sur les résultats des enquêtes de satisfaction. L'accroissement de ces enquêtes de satisfaction (dans le temps et entre entreprises) a donné lieu à des baromètres de satisfaction dont l'utilité première est de produire un ensemble d'indicateurs de résultats et de suivis³⁷, destinés au manager. Ces indicateurs de satisfaction clients sont aujourd'hui une partie intégrante des tableaux de bord prospectifs³⁸ nécessaires au pilotage de l'activité. .

³⁶ ACSI pour American Customer Satisfaction index

³⁷ Les indicateurs de résultats **servent** à juger l'accomplissement final d'une action : réalisation d'un objectif. Les indicateurs de suivis permettent d'observer dans le temps ou par rapport à la concurrence, la satisfaction clients.

³⁸ Nous pensons notamment au Blanced Scorecard de Kaplan et Norton qui est constitué d'un ensemble d'indicateurs financiers et non financiers (satisfaction clients) mis en relation, et qui ont pour objectif de mener à bien le pilotage de la performance.

Deux types de baromètres coexistent :

- les baromètres internes : chaque entreprise en possède un spécifique à son activité et à ses besoins. Souvent ces baromètres contiennent des informations sur la satisfaction générale des clients et les satisfactions par attributs du produit et/ou service de la firme (ex : satisfaction envers l'accueil, le prix du produit etc.).
- Les baromètres externes : ils sont souvent qualifiés de baromètres nationaux, car ils sont relatifs à un ensemble d'entreprises d'un même pays. Le but de ce baromètre est d'avoir un indice synthétique de la satisfaction clients, capable d'évaluer des performances économiques à un niveau national..

L'attention portée aux baromètres de satisfaction s'inscrit avant tout dans une logique de contrôle. Les baromètres permettent de fournir des indicateurs non financiers³⁹, ici la satisfaction clients.

II – Les recherches basées sur les baromètres de satisfaction

Les baromètres, comme enquêtes répétées dans le temps auprès des clients, ont pris deux configurations : les baromètres internes spécifiques à chaque entreprise, et les baromètres externes (entre entreprises). Ces derniers sont souvent qualifiés de baromètres nationaux car ils sont relatifs à un ensemble d'entreprises d'un même pays. Les baromètres nationaux sont différents des baromètres internes dans leur fonction, et dans leur construction. Les baromètres externes servant à obtenir des indices nationaux ont été dévoilés pour la première fois par Fornell (1992) dans son article de présentation du baromètre suédois (SCDB). Les baromètres sous cette forme ont permis de donner des indicateurs supplémentaires aux mesures traditionnelles de la performance. Ces travaux, qui concernaient au départ la Suède et les États-Unis, ont vite été étendus à d'autres pays comme la Norvège ou encore au niveau de l'Europe.

Créé en 1989, le Baromètre de satisfaction des clients suédois SCSB (*The Swedish Customer Satisfaction Barometer*) a été le premier outil à donner un véritable indice national de satisfaction clients pour les produits achetés et consommés localement ainsi que les services (Fornell, 1992). L'indice est estimé à partir du baromètre où la

³⁹ Appellation qui pointe les nouveaux indicateurs introduits suite à des critiques des indicateurs exclusivement comptables

satisfaction est modélisée sous forme de variable latente⁴⁰ en interaction avec ses antécédents telles que la qualité perçue, les attentes et la valeur perçue, et des variables vues comme des conséquences telles que la fidélité (loyauté) et le comportement plaignant. Dans le modèle, toutes ces variables sont considérées comme des concepts latents. Le baromètre suédois est construit de façon uniforme entre plusieurs secteurs. Il a toujours compté environ 130 entreprises de 32 des plus grandes industries de la Suède. D'autres indices et baromètres de satisfaction se sont développés sur le même principe comme : le baromètre de clientèle américaine (ACSI) qui a été introduit à l'automne 1994 et qui concerne environ 200 entreprises de 34 secteurs d'activité (Fornell et al., 1996)⁴¹, ou encore le baromètre de satisfaction clientèle norvégienne (Andreasen et Lervik, 1999; Andreassen et Lindestad, 1998) qui a été introduit en 1996 et publie ses résultats pour 42 entreprises dans 12 industries différentes. Le développement a même été étendu à la clientèle européenne (ECSI) dans quatre secteurs et 11 pays de l'Union européenne.

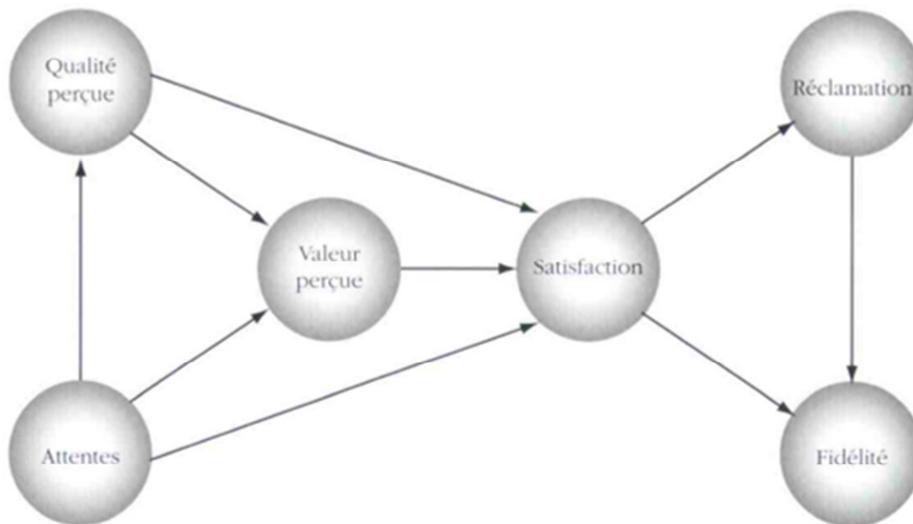


Figure 8: Le modèle A.C.S.I, source article de D. Ray (2004)

⁴⁰ La notion de variable latente suppose que la quantité à mesurer est insondable directement et que pour l'évaluer il faut lui faire correspondre un groupe de variables dite manifestes (observées) qui sont utilisées pour estimer la variable latente.

⁴¹ FORNELL (C.), JOHNSON (M. D.), ANDERSON (E. W.), CHA (J.), BRUYAN (B. E.), The American customer satisfaction index : nature, purpose and findings, *Journal of marketing*, 1996, (60), p. 7-18.

Ces différents baromètres utilisant la modélisation théorique de la satisfaction introduite dans le baromètre norvégien ont été adaptés à chaque pays avec de petites variantes soit méthodologiques soit structurelles (le modèle) pour le besoin du pays concerné. D'ailleurs, en excluant toutes les recherches qu'utilisent ces baromètres comme des outils empiriques, la recherche académique s'est axée sur les baromètres nationaux et rarement sur les baromètres internes. En effet, la recherche s'est intéressée à la fiabilité de ces baromètres dans leurs applications à un pays précis, et surtout elle a ajouté d'autres éléments dans le modèle susceptible d'influencer la satisfaction, pour améliorer l'estimation.

De manière plus quantitative, dans un moteur de recherche tel qu'Ebsco, sur la base d'une recherche où le terme « baromètre » apparaît dans le titre et pour les recherches qui contiennent le terme « satisfaction » tout en se limitant aux revues à comité de lecture, on trouve peu d'articles présentant des recherches sur des baromètres. Ceux qui sont disponibles s'intéressent à une amélioration des baromètres nationaux dans un cadre conceptuel (sur le modèle théorique) (Johnson M. D. et al ; 2001)⁴². Les autres s'intéressent à la méthodologie d'analyse dans un point de vue statistique (Grigoroudis. E, et al ; 2008)⁴³, ou encore visent à montrer l'intérêt des baromètres de satisfaction nationaux notamment sur la performance économique et la nécessité de leur généralisation⁴⁴.

Ces recherches s'inscrivent donc dans une vision d'amélioration et de généralisation des baromètres nationaux pour donner plus de légitimité à leur extension à

⁴² JOHNSON, M. D. et al. The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, v. 22, n. 2, p. 217-245, Apr. 2001.

⁴³ E. GRIGOROUDIS, G. NIKOLOPOULOU AND C. ZOPOUNIDIS, Customer satisfaction barometers and economic development: an explorative ordinal regression analysis, *Total Quality Management* Vol. 19, No. 5, May 2008, 441-460

⁴⁴ ANDERSON, E.W, FORNELL C ; Foundations of the American Customer Satisfaction Index ; *Total Quality Management* ; Sep2000, Vol. 11 Issue 7, p 869, 14

Anderson, Eugene W., Claes Fornell, and Donald R. Lehmann (1994), "Customer Satisfaction, Market Share and Profitability: Findings from Sweden," *Journal of Marketing*, 58 (July), 53-66.

FORNELL (C.), ITTNER (C.D.), LARCKER (D. F.), Understanding and Using the American customer satisfaction index (A.C.S.I), assessing the financial impact of quality initiatives, IMPRO 95 Juran institute's Conference on managing for total quality, 1995, 14 p.

ANDERSON, EUGENE W., CLAESFORNELL, AND ROLAND RUST (1997), "Customer Satisfaction, Productivity, and Profitability: Differences between Goods and Services," *Marketing Science*, 16 (2), 129-145.

ANDERSON (E.), FORNELL (C.), MAZVANCHERY (S. K.), Customer satisfaction and shareholder value, *Journal of marketing*, Octobre 2004, (68), p. 172-185.

d'autres pays. Cependant en prenant l'exemple de la France, cette tentative a échoué à ce jour, au profit d'une multitude de baromètres internes souvent utilisés à tort pour se comparer entre entreprises⁴⁵. Suivant notre point de vue, les baromètres nationaux et les recherches les concernant sont primordiaux, mais à l'instar de ces derniers les baromètres internes doivent aussi avoir de l'intérêt. Leurs résultats sont aussi importants que ceux des baromètres nationaux car ils sont utilisés pour le contrôle de la performance. L'intérêt de la recherche académique sur l'utilisation et l'exploration des baromètres nationaux vient du fait que les modèles sous-jacents sont plus proches des recommandations académiques, notamment dans la modélisation de la satisfaction suivant ses antécédents et ses conséquences, le tout avec une méthodologie statistique basée sur l'estimation via les variables latentes qui respectent les exigences psychométriques. De plus, ces baromètres nationaux sont faits de façon uniforme avec un modèle solide théoriquement. D'un autre côté, les baromètres internes qui sont propres à chaque besoin des entreprises, n'utilisent pas forcément les échelles multi-items et ne font pas forcément référence aux antécédents et aux conséquences de la satisfaction. En d'autres termes, ils se sont éloignés de ce qui est recommandé dans la recherche académique.

Les baromètres dits nationaux ont été vite adoptés dans la recherche académique mais les travaux concernant leur fondement n'a pas été le centre des préoccupations. En effet, comme ces baromètres sont les résultats de sondage, rares sont les recherches qui se sont intéressées à identifier les sources qui perturbent leurs résultats. Mis à part les fondateurs eux-mêmes (Anderson, Fornell, et Rust, 1997) qui se sont questionnés sur les liens entre l'évolution de la satisfaction du client et de l'évolution de la productivité. Leurs résultats, basés sur l'exploitation du baromètre de satisfaction Suédois (SCSB), indiquent que la liaison est positive pour les marchandises, mais nettement moins pour les services. Par ces conclusions, les auteurs pointent des différences qui sont contingentes au secteur. Dans la même démarche, nous trouvons aussi l'article de Johnson M.D, A. Hermann et A. Gustafsson 2001), avec leurs études où ils exploitent les résultats des baromètres nationaux sur trois Pays : Suède, Allemagne et États-Unis. Les auteurs s'interrogent sur les indices de satisfaction comme résultats de comparaison par pays et par secteur. Dans cette recherche, basée sur des arguments de l'économie, de la psychologie, de la sociologie, ils étudient les différences systématiques de la satisfaction clients dans deux secteurs à travers les trois pays. Leurs résultats révèlent que, selon les pays, la satisfaction

⁴⁵ A partir des taux de satisfaction communiqués les entreprises jugent leur position par rapport aux concurrents.

est plus élevée pour des produits compétitifs, inférieure pour des services concurrentiels et encore plus faible pour les organismes gouvernementaux et publics. Toutefois des interactions suivant les pays sont aussi révélées. Dans leurs conclusions, ils remarquent que la satisfaction est inférieure en Suède et en Allemagne par rapport aux États-Unis. Ces recherches et celles de Fornell et al (1996), basées sur l'utilisation d'indices nationaux, sont parmi les seules qui s'intéressent à la question de comparer la satisfaction issue des baromètres nationaux suivant les catégories de produits, secteurs ou encore les pays. Cela nous semble très peu, par rapport à l'utilisation des résultats de baromètres dans la recherche académique, notamment pour établir les liaisons entre la satisfaction clients et la performance. Ainsi, nous constatons que les recherches basées sur les baromètres nationaux s'intéressent marginalement aux résultats, c'est à dire aux indices eux même qui sont considérés comme des indicateurs fiables. Ce postulat, de fiabilité des indices, admis par tous doit être aujourd'hui étudié car il se peut que la légitimité de ces indices comme indicateurs de performance soit remise en cause, surtout lorsqu'on pointe les différences entre pays. D'où l'intérêt de baser les recherches futures sur l'explication de ces phénomènes (différence systématique entre niveaux de satisfactions) d'un point de vue théorique et méthodologique.

Il faut rappeler que la force des indices comme l'ACSI vient de ce qu'ils sont effectués aux niveaux nationaux, avec la même méthodologie pour toutes les industries, ce qui conforte leur utilisation comme outils de comparaison (on compare des choses comparables). Ces indices constituent donc des standards très utiles lorsqu'on veut se comparer. Toutefois, il ne faut pas perdre de vue l'objectif même de ces indicateurs, qui est leur capacité de juger de la satisfaction clients. Donc toute source qui est susceptible de les déstabiliser doit être identifiée et surtout comprise, pour qu'ils évaluent bien ce qu'ils doivent mesurer, c'est-à-dire la satisfaction clients

D'un autre côté, l'inattention touchant les baromètres internes (spécifique à chaque entreprise) fait que des problématiques importantes se trouvent négligées, notamment celles touchant leur fiabilité dans le contrôle de gestion via les indicateurs de satisfaction. Bien sûr l'idéal serait que la méthodologie de ces baromètres internes rejoigne méthodologiquement les baromètres nationaux⁴⁶. Toutefois comme leur utilité première est de contrôler l'activité de l'entreprise en interne, l'intérêt de cette modification

⁴⁶ La méthodologie ici, signifie que la mesure de satisfaction soit multi-items et soit effectuée à la lumière des antécédents et conséquences de la satisfaction.

n'est pas considéré comme pertinent par les entreprises. Ainsi à notre sens, malgré ces défauts, une meilleure attention doit être portée aux baromètres internes autant qu'aux baromètres nationaux en s'intéressant à leurs problématiques, car c'est dans ce cadre-là que les entreprises prennent leurs décisions. De plus ces baromètres internes possèdent une valeur ajoutée qui est celle du volume de données récoltées auprès des clients et la dimension temporelle. Ce changement de point de vue peut servir autant pour les entreprises que pour la recherche future en lui apportant des nouveaux sujets jusque-là inexplorés.

De ce fait, les baromètres de satisfaction, qu'ils soient internes ou externes, sont utilisés pour produire des indicateurs de performance, et la question de leur capacité à remplir les objectifs du contrôle ou le pilotage est toujours en suspens.

Section 4 – **La satisfaction clients dans le contrôle de gestion**

Comme nous l'avons évoqué dans l'introduction de ce texte, la satisfaction clients est aujourd'hui considérée comme un indicateur de performance et rejoint ainsi le domaine du contrôle de gestion, où son exploitation a attiré l'attention des chercheurs qui lui ont donné ce statut d'indicateur.

Bien sûr, le marqueur de la satisfaction clients n'est pas le seul indicateur sur lequel le contrôle de gestion est basé. Avant tout, les indicateurs les plus populaires sont les indicateurs financiers et comptables. Comme le contrôle de gestion est avant tout un système établi pour veiller à la qualité de gestion d'une entreprise, ces indicateurs sont en grande partie privilégiés par leur caractère compréhensible par tous, et leur facilité à être mesurés. Cependant le fait de se réduire à ces indicateurs purement financiers est très réducteur, et un manager ne peut s'en contenter. Dans plusieurs cas, la réussite d'une activité peut dépendre de certains facteurs qui lui sont externes (concurrence, état du marché etc.) et connaître l'évolution de ces derniers peut s'avérer essentiel.

Personne ne conteste l'utilité des indicateurs financiers ou de contrôle budgétaire, car ils présentent l'avantage majeur d'être :

1. Opérationnels en engageant des objectifs et en permettant aux gestionnaires de prendre conscience de l'impact de leurs actes.
2. Produisant un chiffrage en langage commun aux responsables financiers.
3. Présentant des résultats qui synthétisent l'activité pour les dirigeants.

Cependant ils restent relativement limités pour produire plus d'information sur les causes des écarts (prévisions/réalisations) qui peuvent survenir lors du contrôle de l'activité. Un outil de ce type est indispensable pour le contrôle de la planification à court terme, mais ne peut suffire car la prise de décision exige (Bouquin, 1998, p404) :

- Une information continue
- Des données anticipatrices
- Des indicateurs peu nombreux
- Un langage diversifié mais souvent technique plutôt que financier.

De ce fait, il est clair que l'outil budgétaire, s'il est indispensable, paraît comme une condition nécessaire, mais non suffisante pour faire fonctionner le pilotage de la performance. Cette logique a conduit le gestionnaire à se doter de plusieurs d'indicateurs supposés pertinents lui permettant d'anticiper tout changement qui met en cause le fonctionnement du système. D'autre part, l'émergence des indicateurs non financiers (comme la satisfaction) vient aussi de certaines critiques concernant parfois le manque de clarté des indicateurs financiers auprès des managers. Ils sont parfois trop agrégés et n'informent pas sur les changements clés. De plus, devant une concurrence accrue qui rend l'environnement très incertain, leur pouvoir devient très limité. En ajoutant à cela le développement des politiques de management par la qualité (TQM) et la gestion des relations clients (CRM), cela a conduit à s'intéresser à des informations extracomptables pour mesurer la performance.

C'est sur cette lignée, que Kaplan et Norton ont proposé, quatre types d'indicateurs à privilégier dont la satisfaction clients est une composante. Dans leurs BSC, ils ont préconisé que les indicateurs financiers sont des indicateurs de résultats et que les indicateurs non financiers sont des indicateurs avancés⁴⁷.

Les auteurs déclinent donc quatre indicateurs qui sont :

1. Des indicateurs financiers
2. Des indicateurs internes de l'efficacité, et l'efficience des processus clés.
3. Des indicateurs tournés vers le marché rendant compte de la satisfaction des clients.
4. Des indicateurs de progrès continu indiquant la manière dont l'innovation se poursuit.

⁴⁷ Livre de Berland N, 2009 : Mesurer la performance.

Ces indicateurs doivent avant tout fournir des mesures essentielles pour l'étape de comparaison (Benchmarking), dans le temps, en interne et par rapport au meilleur. Cette tournure qui se base sur des informations non financières a fait basculer le contrôle de gestion vers des considérations stratégiques qui tiennent compte de certains facteurs clés, dont bien peu se prêtent à des mesures financières (Bouquin, 1998, p408)

Dans la section précédente, nous avons abordé la nécessité des baromètres de satisfaction dans les entreprises modernes comme outils de production d'indicateur de satisfaction clients. Ces baromètres tirent donc leurs origines du besoin de mesure et d'évaluateur l'environnement externe relative aux clients afin d'assurer une qualité répondant aux exigences des clients (internes ou externes)⁴⁸. Pour appréhender des marchés qui sont de plus en plus concurrentiels, le besoin de se doter d'un système capable de sonder la performance externe est devenu inévitable.

C'est dans ce contexte, que les indicateurs non financiers se sont vus exploités comme information nécessaire pour évaluer la performance des entreprises d'un point de vue externe en complément de la vision interne visant à minimiser les coûts et optimiser le rendement, la productivité et respecter la qualité et le délai. Ces deux dimensions fonctionnent ensemble pour la création de la valeur et pose la notion de leurs dépendances. Comme le précise Berrah (200, p22) « une bonne performance interne est indispensable à une performance externe, puisque elle n'est que la projection ».

Les indicateurs non financiers (comme la satisfaction clients) sont donc des instruments qui reflètent la performance externe ; ils se structurent d'une part d'un jugement du client au regard du produit/service qu'il perçoit et d'autre part du jugement du fournisseur (suivant par exemple les réclamations). Par cet aspect, la satisfaction clients comme concept s'est imposée naturellement dans le sens où par définition, la satisfaction clients est la résultante d'une comparaison entre les attentes des clients et la performance perçue du produit ou du service. Par sa nature même, la satisfaction clients s'est vue projetée comme outil de pilotage. La satisfaction clients mesurée permettra ainsi de capter l'avis clients ce qui permettra de juger la qualité perçue et de la capacité de

⁴⁸ Les clients externes sont les clients finaux, à ceux que le produit ou le service est vendu. Les clients internes représentent les différents services d'une même entité. Chaque composante de l'organisation est soit un fournisseur soit un client, qui fonctionne sous forme de chaînon qui conduit à la performance. Ce modèle « contractuel » suivant Burlaud (2013, p20) consiste à considérer l'entreprise en somme de contrats mettant en relation ces différentes cellules sous une forme : clients-fournisseur. Comme le contrôle vise à influencer les comportements, tenir des exigences de chaque niveau permet de produire de la valeur tout en atteignant les objectifs par un contrôle multidimensionnel : qualité, délai de mise à disposition, le respect des normes etc.

l'entreprise à répondre aux demandes de ses clients. Rappelons aussi que l'élaboration de la mesure dans ce cas est le résultat d'une collecte de données, par exemple dans la satisfaction c'est les questionnaires clients organisés sous forme de baromètres, qui sont synthétisées par à des outils statistiques permettant de produire des résultantes agrégées (au niveau d'un ensemble de clients) qui représente l'indicateur.

Ces indicateurs extracomptables, ont émergé après la prise de conscience que l'appréhension de la performance ne peut plus se restreindre à une minimisation des coûts et une augmentation du volume de production. Les indicateurs financiers, ne permettent plus à eux seuls d'anticiper l'impact global d'une décision. Par conséquent, les indicateurs non financiers ont vu le jour comme outils de gestion capable de combler les faiblesses des indicateurs purement financiers. Nous attirons l'attention sur le fait que ces indicateurs non financier ne doivent pas être perçus comme des substituts, aux indicateurs financier, mais plutôt comme des compléments qui ont des liens de cause à effet (Cauvin et Bescos 2005).

C'est dans cette dynamique de contrôle où la performance est distinguée en approche multicritères⁴⁹, que le Blanced Scorecard de Kaplan et Norton, s'est développé au tour de la liaison entre indicateur financiers et non financiers pour servir un pilotage réactif qui permet à son titulaire d'atteindre ces objectifs. C'est dans cette lancé que la satisfaction clients s'est trouvé intégré comme un indicateur de performance relative à l'axe clients.

Ainsi dans une perspective moderne du contrôle de gestion alliant une vision financière (contrôle budgétaire) et un spectre non financier (tournés vers le marché), la littérature comptable s'est intéressée à l'étude des indicateurs basés sur la satisfaction. A l'instar des recherches en marketing qui ont étudié la relation de la satisfaction clients et performance financière (Anderson, Fornell, et Lehmann (1994), Anderson, Fornell et Mazvancheryl (2004) and Reichheld et Sasser (1990); Ngobo, Ramarasson 2004), des recherches comme celles Ittner et Larker (1998a), Banker, Potter et Srinivasan (2000), se sont intéressées à étudier l'impact de l'utilisation des indicateurs, notamment la satisfaction, sur la performance de l'entreprise. Les auteurs cherchant à savoir si d'une part la satisfaction clients est un indicateur avancé de la performance (financière) et

⁴⁹ La notion de performance étant devenu complexe dans un environnement hautement compétitif, l'atteinte des objectifs ne peut être contrôlées par un seul élément. Les leviers sur lesquels le contrôle de la performance est basé, doivent être actionnables permettant ainsi de fournir des informations sur les causes des phénomènes.

d'autre part si ces indicateurs apportent une information pertinente aux marchés financiers. Les conclusions convergentes de ces recherches indiquent d'une part que la satisfaction clients a des effets sur le comportement future des clients et qu'elle peut être associée à la performance financière à long terme. D'autre part, les conclusions de ces travaux vont dans le sens que les entreprises ayant des clients satisfaits contrôlent mieux leurs coûts. Dans l'étude d'Ittner et Larker(1998b), les auteurs basent leur raisonnement sur trois études distinctes. Les approches utilisées étudient la satisfaction client sur différents niveaux d'agrégation (l'individu ou client, unité d'affaire et l'entreprise), pour examiner les différentes hypothèses. L'article de Lambert (1998), qui discute les résultats des recherches d'Ittner et Larker, appuie sur le fait que cet article apporte une contribution dans le sens où chaque étude vise à analyser une composante de la relation satisfaction clients et performance financière. La première étude examine la relation entre la satisfaction du client et les revenus futurs au niveau du consommateur, le cas concernait le secteur des télécommunications. La seconde étude porte sur les niveaux de satisfactions des clients agrégées au niveau de la branche d'une banque. Cela a permis aux auteurs d'examiner la relation possible ; satisfaction clients, rentabilité future, coûts. La troisième étude teste l'association des mesures de satisfaction et les cours des actions. L'ensemble de ces études donnant des résultats quelques peu contrastés, mais soutenant l'impact positive de la satisfaction et la performance financière, montre bien la difficulté à trancher sur l'effet de la satisfaction (indicateur non financier), surtout quand on travaille sur différents niveaux. Lambert s'interroge alors sur les erreurs de mesure qui peuvent expliquer les différences dans l'importance des effets et continue son examen en discutant la forme fonctionnelle de la relation satisfaction du client et la performance. Dans littérature de gestion, la satisfaction clients a aussi été étudiée en relation avec les coûts engagés. L'article de Cugini et al (2007), étudie cette la relation satisfaction et la maîtrise des coûts, dans le secteur du service. Cette recherche rejoint les conclusions des travaux dans le marketing comme ceux de Reichheld et Sasser (1990), et contribue ainsi à favoriser la satisfaction comme élément du contrôle de gestion.

Des recherches empiriques (Banker Chang Majumdar 1993 ; Amir et Lev 1996 ; Cauvin et Bescos 2005), ont aussi montré des relations positives entre les mesures non financières et financières. Compte tenu de tout cela, en plus de l'introduction du BSC, les entreprises ont adopté les indicateurs non financiers (comme la satisfaction clients) pour appuyer leurs décisions et évaluer la performance. Ces conclusions rejoignent naturellement les préconisations établies dans une littérature Marketing travaillant sur la

relation : satisfaction clients et performance financière. Ce qui a conforté l'utilisation la satisfaction comme indicateur de performance par les entreprises, dont les baromètres constituent des instruments de mesure. Dans cette approche axée sur le contrôle de la performance alliant des indicateurs financiers et non financiers (tournés vers le marché), la mesure et le suivi de la satisfaction clients s'est imposée comme un élément discriminant. L'utilisation quasi systématique des tableaux de bords dans les entreprises modernes, a permis son intégration comme indicateur capable de veiller sur l'axe clients.

➤ **La nécessité des indicateurs non financiers pour le pilotage de la performance**

La notion de pilotage consiste à identifier l'ensemble des actions correctives nécessaires à l'atteinte des objectifs. En vue la complexité des processus physiques à contrôler, le pilotage doit actionner différents niveaux, qui doivent interagir ensemble dans le seul but de contrôler et avoir la main mise sur le système. Pour arriver à ces fins, la performance est pilotée par un ensemble de composantes (d'outils) qui sert la gestion de l'entreprise comme : la comptabilité générale, la comptabilité analytique, contrôle des investissements par la rentabilité, contrôle de opérations par le coût, les tableaux de bord. Dans cette démarche, le pilotage permet de chiffrer (à travers la mesure) les résultats de toutes actions pouvant être prises de façon à permettre une gestion efficace. Berrah (200, p30). Pour atteindre les objectifs, les entreprises ont considérées des axes de contrôle : l'un pour surveiller l'efficience (la productivité) et l'autre pour gérer la relation avec le marché. Le deuxième axe, s'est beaucoup développé car le contexte où les entreprises évoluent est devenu incertain et les évolutions sont continues et peu distantes dans le temps. Dans ce cadre, la valeur dépend beaucoup de ce qui est externe à l'entreprise⁵⁰, négliger cet aspect serait une erreur. Ainsi cette dépendance du marché à fait accélérer la considération des informations extracomptables (inefficace face aux éléments externes à l'entreprise), dont la satisfaction est une des composantes.

La satisfaction clients et les indicateurs non financiers, se sont aussi développé car les indicateurs financiers peuvent aller à l'encontre d'un pilotage réactif. Car l'outil

⁵⁰ Lorsqu'un produit est commercialisé, la valeur qui apporte est le jugement que le consommateur lui donne. L'exigence du client doit être prise en compte pour que la qualité fournie soit en adéquation avec son attente. Bien sûr la qualité perçue client est la résultante de la qualité produite, mais aussi de ce que reflète le produit une fois confronté à tous les produits concurrents. De ce fait la performance est aussi dépendante de facteurs exogènes qu'il faudra contrôler et piloter.

comptable n'est pas spécialement conçu pour le contexte. Comme le précise Berland N (2009)

- Ils n'informent pas de changement clés.
- Ils ne sont pas actionnable car ne fournissent pas des causes des phénomènes et surtout ils ne permettent pas de valoriser l'immatériel (par exemple précisons concurrentielle, moral des employé etc.)

Ainsi le contrôle de la performance sous forme d'évaluation économique et comptable de tout résultat de productivité, à évoluer pour intégrer des informations non financières (comme la satisfaction clients, pourcentage de défaut part des pannes, de réclamation etc.), le tout mesuré par des indicateurs de performance qui sont structurés dans des tableaux de bord. Les indicateurs de performance peuvent être considérés comme des instruments qui envoient donc des signaux, qui seront traduits par des actions en fonction des ressources disponibles pour l'atteinte des objectifs. Berrah (2002, p47), propose la définition suivante : « un indicateur de performance est une donnée quantifiée qui mesure l'efficacité et /ou l'efficience de tout ou partie d'un processus ou d'un système (réel ou simulé) , par rapport à une norme , un plan ou un objectif déterminé et accepté dans le cadre d'une stratégie d'entreprise. L'indicateur a pour vocation à être opérationnel et surtout il doit servir les finalités des entreprises. Selon le même auteur, le processus de pilotage peut être décliné en deux fonctions complémentaires :

- Une fonction prospective qui analyse l'évolution des processus, à travers des variables dites « d'actions » qui sont considérée comme des leviers qui influencent l'évolution des processus. Ils peuvent être internes au processus donc maitrisable par l'entité décisionnelle et d'autres externes qui sont justes observables par l'utilisateur
- Une fonction rétrospective qui analyse les résultats de l'évolution à travers des variables dites « d'état » qui sont des informations à observer pour s'assurer de la bonne réalisation de l'objectif assigné au processus.

Les indicateurs de performance sont donc des instruments de mesure complexe, qui permettent de juger si la performance est atteinte. Ces indicateurs doivent avant tout répondre à deux notions : la notion de mesure (recherche d'information) et la notion d'objectif (le but de l'évaluation). La conjugaison de

deux aspects donne à l'indicateur de performance le statut d'outils de gestion qui doit :

- être en adéquation avec l'objectif auquel il se rattache
- être bien défini par des formules et conventions de calculs acceptés par tous
- être capable de faire état d'un suivi dans le temps
- avoir une mesure valide permettant une évaluation efficace des objectifs.

En somme le jugement de la performance par les indicateurs est un système où les objectifs déclinés sur plusieurs niveaux de contrôle sont évalués, afin de servir la stratégie générale de l'entreprise

Par cela, la satisfaction clients au-delà de sa considération marketing comme concept théorique s'intègre dans une thématique de contrôle de gestion, où elle est la source d'un type d'indicateur de performance (non financier). Par ses qualités et son adéquation à une vision de contrôle tournée vers le marché (aspect clients), elle est utilisée comme un instrument de contrôle dont la mesure et l'évaluation sont les résultats des baromètres.

Ainsi la satisfaction clients comme indicateur se doit de suivre les exigences spécifiques au pilotage de la performance et les qualités imposées aux indicateurs de performance lui sont aussi attachés.

L'objet des développements qui vont suivre est axé sur la vérification de la pertinence des baromètres de satisfaction comme instrument de mesure, pour fournir une information fiable pour le pilotage de la performance. A travers les exigences des indicateurs nous examinerons l'utilisation de la satisfaction dans la pratique de comparaison entre magasins d'une même entreprise et le suivi dans le temps de ses derniers

Chapitre 2 : Facteurs de contingence face à l'évaluation des indicateurs

Dans ce chapitre, nous développons la procédure de diagnostic des indicateurs de satisfaction qui va guider les analyses. Suivant les précisions développées dans les chapitres précédents, concernant l'indicateur de satisfaction barométrique comme indice de performance, nous présentons plus en détail l'ensemble des éléments qui justifient la présence de certains facteurs d'influence.

Nous étudions le cadre de la mesure de satisfaction dans le cas des baromètres d'internes aux entreprises. Nous travaillons sur ces baromètres car ils sont les plus utilisés dans les entreprises pour juger de la satisfaction de leurs clients par points de ventes. Comme notre unité de travail est le magasin, ces baromètres sont adaptés.

Nous présentons les éléments qui peuvent perturber l'évaluation de l'indicateur de satisfaction lorsqu'il est utilisé comme un outil de comparaison. Ces éléments (facteurs de contingence) se manifestent comme des sources externes, qui suivant la situation d'évaluation de l'indicateur peuvent introduire des biais de mesures. Deux sortes de facteurs sont identifiées. Des facteurs relatifs aux caractéristiques des individus (âge, sexe, fréquence de fréquentation) et d'autres relatives à l'environnement (région, le pays).

Nous proposons une définition « mathématique » de l'action des facteurs de contingence. Cette définition permet d'attacher la notion de facteurs de contingence à des erreurs de mesure qui entachent les résultats lorsque l'indicateur est calculé statistiquement. Les facteurs de contingence engendrent des structures qui conditionnent le passage : données individuelles vers l'agrégat

Ces facteurs de contingence, vont guider l'analyse en formulant trois propositions qui visent à tester si l'indicateur répond aux exigences du contrôle de performance lorsqu'il est utilisé comme un outil de comparaison transversale et longitudinale.

Dans un premier temps nous donnons des illustrations, à partir des cas d'études, de la manifestation de ces facteurs de contingence. Suivant les facteurs étudiés des erreurs systématiques sont identifiées. Un autre point qui concerne la comparaison dans le temps des indicateurs est aussi discuté. Une stabilité apparente est constatée. Cette stabilité

paraît être conditionnée au niveau de calcul de l'indicateur : par magasin ou au global (sans distinguer les magasins)

Dans un second temps, nous revenons sur les rares travaux qui mettent en relation les facteurs sociodémographiques et la satisfaction. Cette revue d'articles permet de constater les effets des facteurs externes qui modulent la satisfaction. Toutefois les éléments rapportés ne permettent pas une généralisation de l'action de ces facteurs, mais assurent leurs existences.

Dans un troisième temps, nous discutons la dynamique de la satisfaction dans la littérature. La satisfaction étant un concept qui évolue dans le temps, ces études de permettent de confirmer le caractère contingent de la stabilité des indicateurs.

Nous concluons ce chapitre par le cadre théorique (d'ordre statistique) qui justifie l'existence des facteurs de contingence. Nous énonçons trois propositions qui traduisent le fonctionnement des indicateurs de performance : (1) **capacité à rendre compte des efforts marketing**, (2) **pouvoir discriminer entre les performances des magasins**, (3) **rendre compte des actions entreprises d'une vague à une autre**. Puis nous donnons le cadre pour tester, à l'aide de ces propositions, l'indicateur de satisfaction face à l'existence des facteurs de contingence.

Section 1 – **Présence des facteurs de contingence le constat empirique**

Dans cette section, nous illustrons des résultats constatés lors des études barométriques qui appuient la présence des problématiques de comparaison transversale : différence systématique dans le niveau de satisfaction entre certaines catégories ; et longitudinale : la stabilité des indicateurs dans le temps. L'idée est de présenter les résultats de plusieurs cas rencontrés lors du travail de cette thèse. Il s'agit de problématiques rencontrées lors de la comparaison entre deux groupes de clients, et lors de la comparaison dans le temps des indicateurs de satisfaction. Pour illustrer cela, nous sommes munis de plusieurs enquêtes qui concernent quatre secteurs d'activité : Téléphonie, Banque, Bricolage, Habillement, pour lesquelles nous avons étudié l'étendue des différences dans les niveaux de satisfaction constatés suivant l'âge des individus et le sexe de ces derniers. Le choix de ces deux facteurs (âge et sexe), est le résultat direct de

notre revue de la littérature⁵¹. De plus dans la pratique managériale, il n'est pas rare qu'on s'intéresse aux niveaux de satisfaction déclinés par tranche d'âge ou par sexe, afin de comparer ces segments⁵².

D'un autre côté, pour la perspective longitudinale, nous nous sommes munis d'un baromètre qui concerne une grande enseigne de bricolage. Les enquêtes sont réparties sur 16 vagues. Ce baromètre inclut un suivi de 14 critères de satisfaction pour des clients situés sur la France métropolitaine. Pour chaque critère un score résumant la satisfaction est présent, le tout pour 16 vagues d'enquêtes. Le baromètre par sa nature dynamique a servi à l'illustration de la présence d'une certaine stabilité des indicateurs dans le temps.

I – Différences systématiques : l'exemple de l'âge et le sexe des individus

Dans cette démarche d'illustration, nous voulons identifier les facteurs qui jouent un rôle perturbateur sur la mesure de satisfaction vue globalement sur la population des consommateurs. Dans cet objectif nous détaillons les effets de deux facteurs sociodémographiques à savoir l'âge et le sexe, en analysant un ensemble d'enquête. Dans cette perspective, nous donnons une analyse chiffrée résumant les différences dans le niveau de satisfaction conditionné par catégories d'âge et par sexe.

Toutes les enquêtes, sur lesquelles nous travaillons sont faites auprès de consommateurs français. Pour l'étude des deux facteurs âge et sexe il nous paraît pertinent d'avoir les mêmes catégories lors de l'analyse. Nous avons choisi pour le facteur âge les catégories suivantes :

Catégories d'âge	Moins de 20 ans	De 21 à 30 ans	De 31 à 40 ans	De 41 à 50	De 51 à 60	Plus de 60 ans
Modalités	1	2	3	4	5	6

Pour le facteur sexe la codification suivante :

⁵¹ Par rapport à d'autres facteurs sociodémographiques, l'âge et le sexe des répondants sont ceux les plus étudiés.

⁵² Cette distinction permet de connaître l'avis clients par catégorie afin d'adapter l'action marketing. Ceci est d'autant plus utilisé quand le produit/ou le service fourni est différent pour chaque catégorie, on pourra penser à l'exemple du prêt-à-porter.

Catégories de sexe	Hommes	Femmes
Modalités	1	2

Pour ce qui est de la nature des données disponibles pour l'analyse, les données pourraient être résumées par la table suivante :

Tableau 1 Informations disponibles par secteurs

Secteurs	Nombre Années disponibles	Nombre de données cumulées	Le critère de satisfaction retenue pour l'analyse	L'échelle de mesure
Bricolage	2 ans	145014	Satisfaction globale	Note de 1 à 10
Habillement	4 ans	49595	Satisfaction globale	Note de 1 à 10
Banque	4 ans	93052	Satisfaction globale	Note de 1 à 20
Téléphonie	1 an	7227	Satisfaction globale	Note de 1 à 10

Les effectifs observés par catégories dans les bases servant à implémenter l'analyse, sont les suivants :

Tableau 2 : Table des effectifs par secteurs

	AGE	<20 ans	21-30	31-40	41-50	51-60	>60 ans
Téléphonie	Effectifs	1213	2000	1579	1286	746	403
Bricolage	Effectifs	862	17677	32970	33149	34612	25744
Banque	Effectifs	1048	9319	14474	17896	18641	31674
Habillement	Effectifs	4930	10605	9132	9174	8928	6826

	SEXE	Homme	Femme
Assurance	Effectifs	4290	3260
Téléphonie	Effectifs	3493	3734
Bricolage	Effectifs	102512	42502
Banque	Effectifs	49679	43373
Habillement	Effectifs	8983	40612

Les deux derniers tableaux nous montrent que pour chaque catégorie un nombre de données important est présent. Ce qui permettra de conforter les résultats des analyses

statistiques qui vont suivre. La différence dans les effectifs (entre catégories) vient du fait que chaque enquête est effectuée sur des échantillons représentatifs des clients de l'entreprise en question. Toutefois étant le nombre important, la comparaison ne sera pas perturbée par le décalage dans l'effectif⁵³.

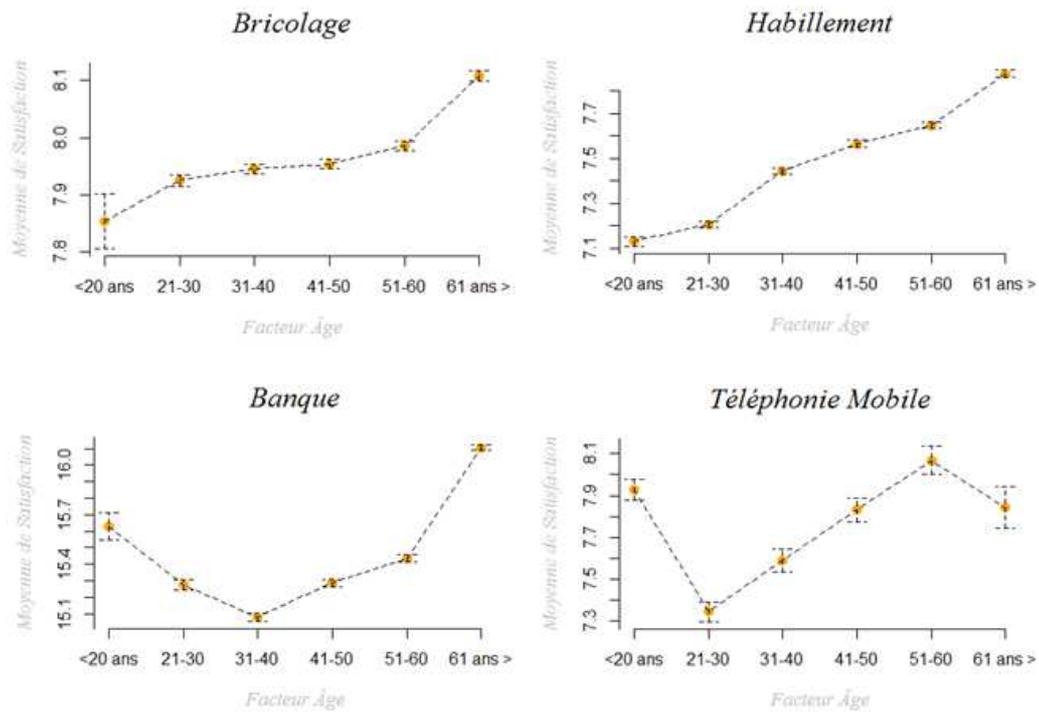


Figure 9: Différence dans les niveaux de satisfaction : âge des répondants

⁵³ Ce fait est appuyé par la loi des grand nombre qui assure un seuil minimum d'erreurs dans l'estimation lorsque le nombre de données est important.

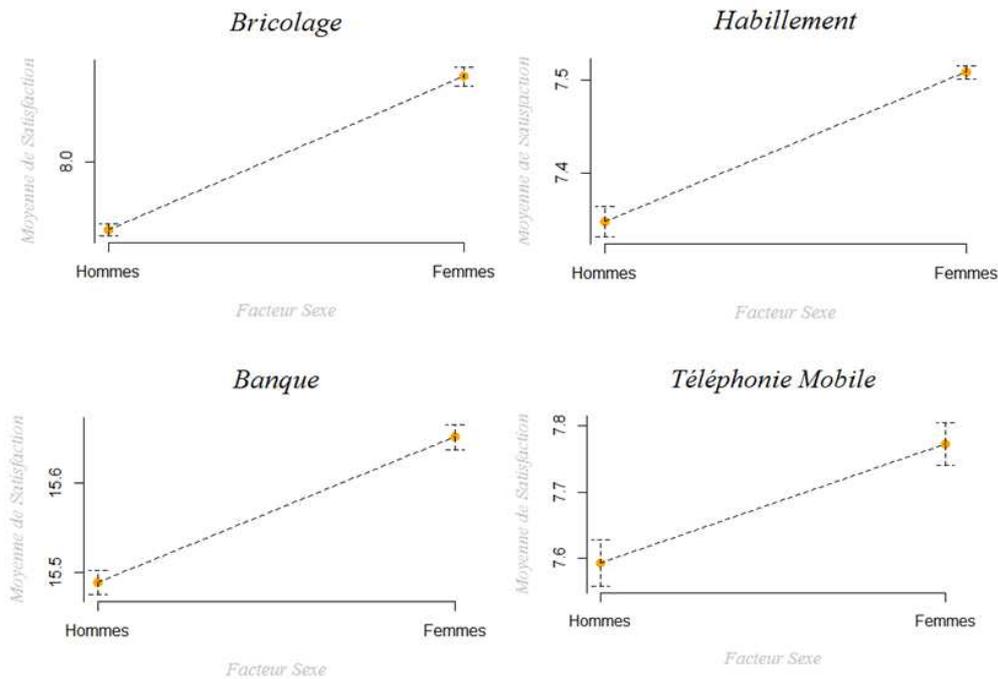


Figure 10 : Différence dans les niveaux de satisfaction : le sexe des répondants

Une analyse rapide de ces résultats (figure 9 et 10) de façon descriptive, nous informe qu'il existe une différence systématique entre les hommes et les femmes dans la note moyenne de satisfaction, le tout en faveur des femmes (les femmes semblent plus satisfaites que les hommes). Nous remarquons aussi une différence entre les différentes catégories d'âge avec la tendance d'une note moyenne supérieure pour plus de 60 ans (les personnes âgées). Toutefois le degré des différences pour le secteur de la téléphonie semble différer par rapport aux autres secteurs. Ce fait indique l'existence des caractères contingents des deux facteurs âge et sexe⁵⁴.

De ces graphiques, nous remarquons clairement une différence systématique entre les catégories d'âge ou de sexe des répondants, et cela est présent dans les quatre secteurs. Des tests statistiques basés sur des analyses de variance, nous confirment les résultats constatés. Les résultats de ces tests sont récapitulés dans les tableaux suivants :

⁵⁴ Les différences existent bien mais ne paraissent pas avoir la même intensité. Ce caractère explique le fait que les deux facteurs : âge et le sexe sont juste situationnels et ne peuvent être liés directement au concept de satisfaction, mais plutôt à sa mesure barométrique : les indicateurs. En reprenant le cas de la téléphonie, nous pouvons penser que les personnes âgées sont plus résistantes aux nouvelles technologies, ce qui fait que leurs satisfactions moyennes est plus basse. Ainsi il semble existait une interaction entre le secteur et l'âge.

Tableau 3: Résultats des analyses de variance par secteurs

Bricolage	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Âge	5		550.3321	110.066427	49.7023	1.27E-51
Residuals	145008		321122.2212	2.214514	49.7023	1.27E-51

Habillement	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Âge	5		2808.125	561.625048	278.5156	7.89E-295
Residuals	49589		99995.913	2.016494	278.5156	7.89E-295

Banque	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Âge	5		15235.46	3047.091447	345.3639	0.00E+00
Residuals	93046		820930.14	8.822842	345.3639	0.00E+00

Téléphonie Mobile	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Âge	5		466.1155	93.223094	23.04696	4.93E-23
Residuals	7221		29208.3627	4.044919	23.04696	4.93E-23

Bricolage	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Sexe	1		415.4808	415.480755	187.5436	1.16E-42
Residuals	145012		321257.0726	2.215383	187.5436	1.16E-42

Habillement	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Sexe	1		189.2661	189.266119	91.47099	1.18E-21
Residuals	49593		102614.7725	2.069138	91.47099	1.18E-21

Banque	Degré de liberté	de	somme des carrés	somme des carrés des Moyens	F	P value
Facteur Sexe	1		619.4935	619.49352	68.98946	1.00E-16
Residuals	93050		835546.0993	8.979539	68.98946	1.00E-16

Téléphonie Mobile	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Facteur Sexe	1	58.85669	58.856686	14.35862	1.52E-04
Residuals	7225	29615.62152	4.099048	14.35862	1.52E-04

De ces analyses supplémentaires (tableau 3), visant à qualifier les différences par test statistique, nous remarquons que les résultats confortent l'existence d'un effet de l'âge et le sexe sur le niveau de satisfaction, le tout avec des Pvalue très significatives. Ces conclusions confirment les différences systématiques constatées par segment d'âge et de sexe, quel que soit le secteur. L'amplitude de ces différences (plus forte, ou moins forte suivant le secteur) nous oriente plus vers une existence contingente⁵⁵ de ces facteurs. Ces calculs nous montrent que des facteurs (le cas de l'âge ou le sexe) peuvent jouer sur les différences entre les niveaux de satisfaction. Une question émerge de ces résultats, elle peut être formulée ainsi: est-ce que ce type de facteurs (liés aux caractéristiques et à l'environnement des individus sondés) sont assez importants pour que les indicateurs de satisfaction soit sensibles à leurs effets? La sensibilité dans ce cas se traduirait par le fait qu'un indicateur de satisfaction subit les effets de ces facteurs et ses fonctions d'indicateur de performance sont mises à défaut.

II – Comparaison longitudinale : l'invariance des baromètres

Comme pour les problématiques transversales, nous illustrerons dans cette partie les résultats empiriques constatés dans les baromètres de satisfaction. Il s'agit de données relatives à 14 critères de satisfaction qui sont récoltées par trimestre, sur un échantillon représentatif de l'activité de la firme (au niveau national)⁵⁶. Pour chaque critère, un score est calculé pour 17 trimestres de 2006 à 2011. Le score (de 0 à 60) est construit comme une somme pondérée de la distribution de notes récoltées à partir d'une échelle de satisfaction à 6 points.

⁵⁵ Suivant la situation, le nombre de donnée et le secteur, les effets apparaissent avec plus ou moins d'intensité.

⁵⁶ Ces données qui vont servir pour les analyses futures seront détaillées dans le chapitre méthodologie.

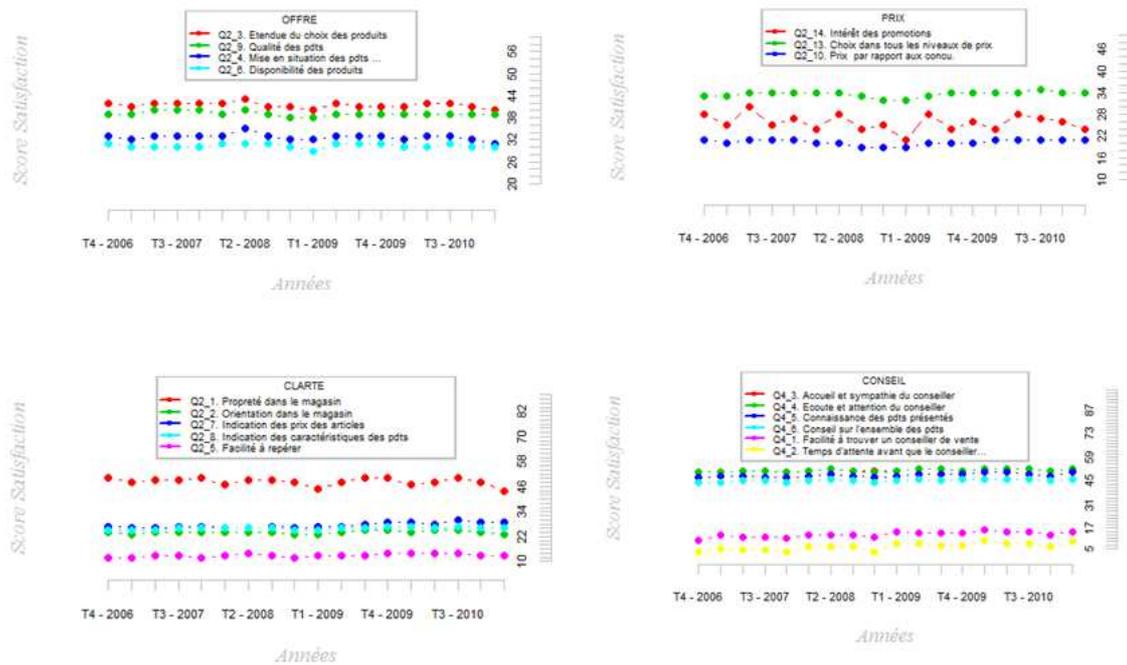


Figure 11: Sensibilité des critères de satisfaction sur 16 vagues d'enquêtes

La figure 11, nous montre une stabilité temporelle dans les scores au niveau national, à part pour le critère « intérêt des promotions » qui lui paraît fluctuer plus que les autres (car les promotions, autant que sujet, changent avec le temps). Les scores des autres critères présentent des évolutions dans le temps très proches. Un tel résultat est compliqué à appréhender pour la prise de décision. Un gain d'un point ou deux sur un score variant de 0 à 60, est-il significatif pour inférer une amélioration ? De même une perte de un ou deux points, signifie-t-elle une détérioration. Ces questions posées quotidiennement par les managers restent aujourd'hui sans réponse et sans explication. La stabilité est-elle un fait qu'il faudra accepter ou cette stabilité est juste contingente ? Le caractère contingent est-il le seul fruit d'une résultante globale (le cas des scores nationaux : ensemble de magasin) ? Est-ce que l'agrégation vue comme un mélange de population fait que les résultantes sont entachées d'erreurs susceptibles de causer de la stabilité dans le temps ? Est-ce que le cycle clients (entrants/sortant) induit cette forme de stabilité. Voilà autant de questions sur lesquelles il faut apporter des explications.

Dans la suite avec la considération des magasins comme unité de contrôle, nous avons voulu décrypter cette stabilité et savoir si des variables externes au concept de

satisfaction (facteurs de contingence) sont responsables de cette invariance qui est constaté dans les baromètres. Nous attirons l'attention sur le fait que cette stabilité n'est pas propre aux baromètres internes, mais touche également les baromètres nationaux comme ceux de l'ACSI⁵⁷.

En somme par ces analyses nous constatons deux sources de problématiques : transversale et longitudinale qui coexistent avec l'effet de certains facteurs. Ces constats qui accompagnent le manager dans la pratique, posent la question de leurs sources et de leurs importances.

Section 2 – La différence entre unité sociodémographiques et géographique

Si au cours des dernières années, de nombreuses études touchant à la satisfaction ont été consacrées aux déterminants, aux modérateurs et aux conséquences de la satisfaction, les résultats quant à eux sont contrastés. Il apparaît dans ces recherches que les relations entre la satisfaction et ses conséquences, varient fortement selon les secteurs, les types de consommateurs, les produits. Dans une étude récente Cooil et Al (2007), Les auteurs passent en revue un certain nombre de facteurs de modération tel que l'âge, le revenu, l'éducation, l'expertise, et recensent les études qui les ont étudiés. Cependant, leurs études et celles qui les ont précédées (Homburg et Giering; 2001 ou encore Ueltschy et Krampf, 2001) s'appuient généralement sur des données empiriques qui concernent une seule firme, une seule clientèle, un seul échantillon ce qui limite la portée générale de ces propositions, d'autant plus que d'une étude à l'autre, des résultats contradictoires sont enregistrés.

La littérature marketing s'est déjà intéressée à des aspects de contraste entre catégories de la population, vus comme des facteurs influents sur la mesure de satisfaction. Cependant ces études restent rares et surtout aucun consensus sur la nature de l'effet n'a pu être formellement identifiée pour chaque groupe défini selon les caractéristiques sociodémographiques. Dans une étude sur le parc automobile menée par Peterson et Wilson (1992), il ressort que les personnes âgées (plus de 55 ans) sont plus satisfaites que les plus jeunes (moins de 35 ans). Moutinho et Goode (1996) dans leur

⁵⁷ Nous avons fait référence à la stabilité des scores ACSI, dès l'introduction de notre thèse. Les scores ACSI par firme ou au niveau national bougent peu. La variation est au maximum de l'ordre de +5 -5 points, pour un score qui lui calculé sur un support de 0 à 100.

étude sur les distributeurs automatiques de billets, révèlent que l'âge a une influence sur la satisfaction clients. D'autres recherches ont montré un effet du sexe sur la satisfaction (Goode et Moutinho 1995 ; PhilemonOyewole, 2001) quand d'autres ont montré une influence négative de l'éducation (Pfaff et blivice, 1977 ; Parasuraman et Futrell, 1983) et du revenu (Soley et Reid 1983, PhilemonOyewole, 2001) sur la satisfaction. Et enfin des recherches se sont penchées sur la différence entre les pays et les catégories du produit (Johnoson et al 2002). Leurs résultats basés sur les études des baromètres nationaux de trois pays (États-Unis, Suède et Allemagne) montrent l'existence de différences dans le niveau de satisfaction entre la catégorie du produit, des services, et une certaine interaction avec le pays. Cette étude met en lumière des différences dans les niveaux de satisfaction et complète une recherche antécédente d'Anderson E.W (1994) dans laquelle l'auteur pointait déjà la variation du niveau de satisfaction et l'intensité de sa liaison avec le ré-achat selon des conditions extérieures comme la concurrence et l'expérience des consommateurs. Dans la même idée récemment un article de Songting Dong, et al (2009), a pointé les effets des variables économiques et démographiques sur la forme fonctionnelle qui peut lier la satisfaction et le ré-achat. Leur étude établit les effets sans autant donner un cadre général expliquant cette influence.

De ces recherches, il ressort qu'il existe bien une influence de certains facteurs sur la mesure de satisfaction, cependant aucun élément théorique n'est avancé pour expliquer leurs effets. Même si à chaque fois, des éléments de réponses sont avancés, ils restent contingents aux types d'études et du secteur. Les expériences relevées concernent les caractéristiques individuelles et leur environnement. Par définition, ces facteurs ne sont identifiables qu'en présence d'une population, ce qui reste différent des caractères liés au niveau psychologique de l'individu. A partir de cette considération les facteurs paraissent comme des quantités attachées à la mesure de satisfaction (au niveau de la population), qui conditionnent les individus en groupes. Ils sont attachés uniquement à la mesure car il nous semble contradictoire de les rattacher au concept de satisfaction car leur influence n'est pas régulière ni dans son sens (ex : les femmes, n'ont pas toujours un niveau de satisfaction élevée par rapport aux hommes) ni dans son intensité. Ainsi leur caractère contingents ne peut justifier leur influence sur le concept lui-même mais plutôt sur sa mesure de façon globale (construction des échelles, calcul d'indicateurs, technique de sondage).

C'est dans la perspective de mieux comprendre l'influence de ces facteurs sur la mesure de satisfaction que le besoin de formaliser leurs effets est devenu urgent. Leur action a déjà été identifiée dans la littérature marketing mais l'explication de leur influence reste toujours en suspens. Nous attirons l'attention sur le fait que l'explication de leurs effets dans les recherches qui ont travaillé sur le sujet, concernait le concept psychologique c'est-à-dire au niveau de la formation de la satisfaction, et non sur l'évaluation de l'indicateur agrégé (moyenne, taux de satisfaction)

Sur ce principe de dissociation entre le concept de satisfaction et sa mesure, nous donnerons le statut de facteurs contingents, à tout élément attaché à la procédure de mesure au niveau de l'agrégat (la population). Ce positionnement implique que l'étude des facteurs de contingence, se situe au niveau de l'indicateur (agrégat des mesures clients) et non au niveau de l'individu (concept psychologique).

Ainsi dans les développements qui vont suivre, nous considérons ces facteurs qui sont qualifiés de contingents comme des causes qui perturbent la mesure barométrique. Leurs effets produisent des sources d'erreurs lors de la phase de comparaison par indicateurs (agrégats)

Section 3 – **La dynamique de la satisfaction dans la littérature**

Toutes les recherches qui se sont intéressées au concept de satisfaction, témoignent de l'importance accordée à ce dernier. Cette diversité de recherche a malheureusement provoquée une multiplicité des définitions et des catégorisations de ce qu'est la satisfaction clients. Cependant suivant les conclusions avancées, toutes les recherches se mettent d'accord sur le fait que la satisfaction est un résultat de comparaison et qu'elle est considérée comme un construit dual cognitif et affectif (Vanhamme 2002)

La satisfaction est souvent considérée sous deux aspects : une forme transactionnelle et une forme relationnelle. Dans l'approche transactionnelle la satisfaction est définie comme un état postérieur et inhérent à une transaction spécifique et bornée dans le temps (Audrain-Pontevia, 2003). C'est cette considération qui a été la plus étudiée et dans laquelle les antécédents et conséquences de la satisfaction ont été examinés. Dans l'approche relationnelle, la satisfaction est définie comme étant « *le niveau des satisfactions cumulées consécutives à plusieurs expériences de consommation d'un objet (produit ou service)* » (Audrain-Evrard, 2001) ou encore comme « *une*

évaluation globale continue de l'aptitude de l'entreprise ou de la marque à fournir les bénéfices recherchés par le client. » (Ngobo, 1997). Ces définitions renvoient avant tout aux caractères temporels de la satisfaction qui a été qualifiée de satisfaction cumulée. Ainsi dans une étude dynamique de la satisfaction, il paraît que c'est la satisfaction cumulée qui est présente le long des processus de mesure barométrique d'un point de vue individuel.

Pour étudier la dynamique de la satisfaction, il paraît judicieux de distinguer les deux formes de satisfaction, le tout en s'appuyant sur l'aspect relationnel⁵⁸ car par définition il est lié à une dimension temporelle.

Comme la satisfaction cumulée ou relationnelle a été faiblement étudiée⁵⁹, les recherches ont été peu nombreuses sur la dynamique de la satisfaction. L'étude de cette dynamique est un sujet important étant donné les attentes managériales à ce sujet. Toutefois les études de l'effet du temps sur la satisfaction ont permis d'appréhender l'évolution conjointe des variables qui lui sont subséquentes, comme le précise Audrain-Evrard, (2001). Comprendre mieux comment la satisfaction interagit avec les concepts qui lui sont proches (ex la qualité perçue, l'attitude) permet d'une part de dissocier ce qui est vraiment de la satisfaction et d'autre part d'établir les liens de causalité entre ces antécédents et ces conséquences.

Les premières recherches étudiant la satisfaction clients se sont penchées avant tout sur les relations entre la satisfaction et l'attitude ou la fidélité dans le temps. C'est à partir de ces points que les premiers enseignements ont vu le jour sur la liaison entre la satisfaction et le processus post-consommatoire comme la fidélité des consommateurs. Oliver (1980) observe que la satisfaction est déterminée par la disconfirmation des attentes, et qu'elle est un antécédent de l'attitude, qui est elle-même antécédente de l'intention d'achat. Ces résultats sont confirmés par d'autres auteurs comme Trawick (1981) et Bearden et Teel (1983). Ces conclusions prouvent que la satisfaction a été depuis son début rapprochée à la dimension temporelle ce qui valide son approche dynamique.

D'autres recherches comme celles de Bolton (1998) montrent, pour un service, que la satisfaction détermine la durée de la relation sous l'hypothèse que la satisfaction

⁵⁸L'aspect relationnel de la satisfaction inclue une suite d'expériences dans le temps et toutes autres interactions entre les clients et l'entreprise.

⁵⁹ En comparaison à la satisfaction transactionnelle

s'actualise au fur et à mesure du développement de la relation. La théorie sur laquelle il se base, qui a été présentée par Hogart et Einhorn (1992), stipule que chaque consommateur réajuste ses attentes en intégrant des informations. Dans la même perspective, les deux auteurs Bolton et Lemmon (1996) montrent que la satisfaction est liée à l'utilisation à venir des services et stipulent que la satisfaction est plus élevée chez les consommateurs fidèles par rapport aux autres. D'autres recherches soulignent aussi que la satisfaction dans sa dynamique est un déterminant de l'engagement et de la confiance (Tax, Brown et Chrandrashekar, 1998). Ces résultats montrent l'intérêt de la satisfaction dans les processus post-consommateurs et que suivre sa dynamique devient déterminant. Dans cette logique les recherches sur la dynamique comme celles de Bolton, induisent qu'un consommateur donne autant de poids à la satisfaction cumulée et qu'une nouvelle information a d'autant moins de poids sur sa relation avec l'entreprise que si celle-ci est avancée dans le temps. Le même auteur, rajoute que si l'information nouvelle diffère par rapport à des informations antérieures, alors elle amènera une modification importante de la satisfaction. Mais si les nouvelles informations diffèrent faiblement des informations antérieures, alors cette dernière sera assimilée aux informations précédentes et le niveau de satisfaction ne sera plus modifié.

Dans la même lignée, des recherches comme celles d'Inman, Dyer et Jia (1997) montrent que la satisfaction est influencée par les expériences passées et que les croyances antérieures influencent les croyances actuelles. Ces recherches montrent aussi que le consommateur réajuste ces croyances et sa perception de la qualité dans le temps. Ces résultats sont rejoints dans une recherche de Boulding, Klara, Staeling (1999) sur la qualité perçue et sa dynamique, dont les conclusions expliquent que les attentes prédictives (ce que les consommateurs prévoient en termes de performance) font l'objet de réajustement a posteriori de l'expérience de consommation. Ces enseignements comme ceux tirés des recherches de Bolton montrent que la satisfaction reste étroitement liée à la qualité perçue et que sa dynamique reste liée à l'évolution de ses déterminants.

Un autre point soulevé dans les études longitudinales de la satisfaction clients est celui du sens de causalité avec la qualité perçue. Il faut dire que par leurs définitions, les deux concepts peuvent être confondus. Depuis les travaux de Parasuraman, Berry et Zeithaml (1985) la qualité perçue d'un service a été définie comme une comparaison de l'écart entre les attentes et la perception des niveaux de performance réelle du service.

Ainsi la satisfaction se trouve confrontée à une confusion avec la qualité perçue, car elle-même a été définie selon un modèle de confirmation /disconfirmation des attentes (Oliver 1980). C'est à partir de cela que la différence entre les deux concepts n'est plus bien perçue (Ngobo, 1997).

Différentes recherches se sont intéressées à la liaison et à la conceptualisation des deux dimensions : qualité perçue et satisfaction. Certains auteurs comme (Parasuraman et al, 1994 ; Zeithaml et al, 1993) ont essayé de dégager une structure explicative basée sur les standards de comparaison en notant que la qualité perçue est relative à des croyances normatives du client (ce que devrait offrir l'entreprise en matière de service) quant à la satisfaction elle serait relative à des anticipations (ce que les clients croient qu'une entreprise de service lui offrira). Cependant comme le précise Ngobo 1997, ces distinctions basées sur les standards de comparaison ne permettent pas de distinguer entre la qualité perçue et la satisfaction clients. D'autres explications sont avancées, ils placent la satisfaction dans une vision transactionnelle, comme celle de Bitner (1990), Bolton Drew (1991) et Parasuraman et al (1998). Ces auteurs suggèrent que la qualité perçue est une attitude globale alors que la satisfaction est spécifique. Cette conceptualisation situe la satisfaction d'un point de vue transactionnelle et la qualité perçue d'un point de vue globale ou relationnelle.

Iacobucci, Grayson et Ostrom (1994) pensent que la distinction entre la qualité et la satisfaction peut se faire en considérant que la qualité perçue comme un jugement purement cognitif essentiellement tournée vers l'extérieur (ex : est-ce que la qualité est bonne) tandis que la satisfaction est un dual cognitif/affectif qui serait interne, personnel et impliquant (ex : l'avez-vous aimé) (Ngobo 1997). Cependant comme le souligne le dernier auteur, que dans une perspective temporelle la part de l'émotion diminue et l'évaluation satisfaction devient de plus en plus cognitive, de ce fait la qualité perçue et la satisfaction peuvent converger (Dabholkar, 1993, 1995)

Ainsi dans cette perspective temporelle la qualité perçue et la satisfaction paraissent liées et leurs distinctions se situent dans leurs liens de causalité dans le temps. Ngobo (1997) préconise que la distinction entre concepts ne peut être qu'au niveau transactionnel, et qu'au niveau relationnel les deux concepts ont tendance à converger et à être assimilés à une attitude globale à l'égard de l'entreprise et de ses prestations.

D'autre part les conséquences de la satisfaction comme la fidélité restent confirmés dans une dimension temporelle. Cependant les autres subséquents de la satisfaction comme la recommandation ne sont à notre connaissance que rarement étudiés dans cette perspective temporelle à l'exception de l'article de White C.J.(2010) qui a étudié la dynamique positive entre la dimension affectif de la satisfaction et le bouche-à-oreille positif. La satisfaction a ainsi été étudiée comme composante psychologique et rarement comme indicateur même si dans le développement précèdent la satisfaction se confond avec d'autres concepts. L'autre élément qui ressort de ces recherches est que la satisfaction n'est pas stable et qu'elle est sujette à des réajustements au fil du temps. Il est clair que la satisfaction dans sa vision barométrique subit ces constats théoriques mais aussi des effets de mesures qui sont rarement pris en compte dans la littérature académique.

En résumé, toute forme de stabilité dans la satisfaction est rarement présente dans la littérature, ce qui nous conforte de considérer qu'une stabilité des indicateurs de satisfaction ne peut être que contingente, liée à la situation, à l'évaluation de l'indicateur.⁶⁰

La stabilité des indicateurs de satisfaction, trouve des débuts de réponse en se basant sur les résultats des travaux qui supposent la réactualisation des attentes individuelles. Cette réactualisation est soutenue par l'hypothèse que la satisfaction s'actualise au fur et à mesure du développement de la relation. En effet, on admettant que les clients d'une entreprise ne sont pas éternels et qui ont un cycle de vie, le double mouvement de la réévaluation des attentes et du renouvellement du portefeuille produit une stabilité des indicateurs.

En n'admettant que H1 : *les attentes des clients sont continuellement réévaluées, au même titre que l'amélioration du produit ou du service* et que H2 : *Les clients recrutés (satisfaits) remplacent les partants (insatisfait)*

L'augmentation de la qualité ou du niveau de service (donc de satisfaction) conduit à une augmentation des attentes. A mesure que les entreprises répondent aux attentes des clients pour mieux les satisfaire, elles encouragent aussi la réévaluation des attentes, se condamnant à de nouvelles innovations et amélioration pour simplement maintenir le

⁶⁰ Par format nous entendons le mode de calcul de l'indicateur et le support sur lequel il est évalué. Un support qui renvoi au niveau de la mesure est effectuée : par exemple l'ensemble des magasins de la firme, les magasins d'une région etc.

niveau de satisfaction. Une sorte de tapis roulant qui peut créer un état d'équilibre permanent. Les entreprises fidélisent mais recrutent aussi des nouveaux clients car les départs sont inéluctables. On peut supposer que les clients partants sont ceux qui ont la plus basse satisfaction et des attentes assouvies face à des attentes élevées. En revanche les nouveaux clients n'ont pas un niveau prononcé d'attentes et peuvent être satisfait facilement et sans trop d'effort.

A partir de cette double hypothèse, deux configurations deviennent possibles :

-Une progression de la satisfaction par rapport à la vague précédente : c'est le cas où les nouveaux arrivants sont plus nombreux que les partants et en même temps, les restants sont constants au niveau de la satisfaction

-Une diminution de la satisfaction par rapport à la vague précédente : c'est le cas où les recrutés sont moins nombreux que les partants et que les restants sont en phase de devenir des futures partant (niveau de satisfaction baisse).

Les deux hypothèses émises précédemment, sont centrées sur l'augmentation des attentes clients vue globalement sur la population des consommateurs. Si cette augmentation est comblée par l'évolution du produit, il y a équilibre (H1). Si ce n'est pas le cas, un déséquilibre négatif peut être corrigé par le recrutement des nouveaux clients (H2). Si le déséquilibre est positif, les améliorations vont au-delà des nouvelles attentes, une stabilité des indicateurs marquerait simplement un équilibre. On voit ici que l'invariance des mesures de satisfaction est conditionnée à deux cas qui se complètent et qui sont maintenus comme tels par les entreprises. La stabilité des indicateurs de satisfaction masque en fait une dynamique plus forte. L'équilibre se situe dans le rapport entre les attentes et la performance perçue par les clients, et entre le renouvellement de ces derniers.

Ici l'explication de la stabilité des indicateurs est le reflet d'un processus qui est sous-jacent à la satisfaction et qui est visible lorsque cette dernière est mesurée au niveau d'une population. Toutefois la nature des niveaux de calcul considérés reste en suspens. En effet, les indicateurs étant des quantités qui peuvent être distinguées suivant un

magasin, une région ou encore sur l'ensemble des clients d'une entreprise. La question de la stabilité supposée doit être testée à ces différents niveaux, pour qu'elle soit considérée comme existante. Ce dernier point sera développé dans cette recherche. Nous verrons ainsi que la stabilité des indicateurs est aussi conditionnée à certains facteurs (notamment géographiques) qui module cette dernière.

Le but de notre travail est de faire un diagnostic de cette stabilité par niveaux de calcul, pour voir dans quelles conditions les indices de satisfaction gardent leurs propriétés d'indicateurs de performance. Par cela, nous travaillerons non pas sur la dynamique du concept de satisfaction mais plutôt sur l'indicateur vu comme une résultante agrégée d'un ensemble de mesures individuelles.

Section 4 – **Vers une définition des facteurs de contingence**

En considérant les baromètres de satisfaction comme des outils de contraste entre les résultats de satisfaction d'une entreprise ou comme un moyen de juger l'évolution dans le temps d'une société, une question concernant l'usage émerge. Cette interrogation concerne la validité d'une telle comparaison et son efficacité à répondre aux objectifs voulus. En effet, étant donné que les enquêtes effectuées sur plusieurs vagues, ne concernent que rarement les mêmes individus, le rapprochement des résultats peut être biaisé par la considération des facteurs de contingence c'est-à-dire les facteurs liés aux caractéristiques des individus et à leur environnement.

Dans la perspective de mieux contrôler la mesure de satisfaction barométrique, nous nous sommes intéressés à un certain nombre de facteurs que nous avons qualifié de contingents ; ces derniers se distinguent en deux classes: une partie est liée aux caractéristiques des individus : variables sociodémographiques et une autre partie est attachée à l'environnement où la satisfaction est mesurée : la région ou encore le pays.

Dans notre démarche nous voulons identifier les facteurs qui jouent un rôle perturbateur sur la mesure de satisfaction vu globalement sur la population des consommateurs. Dans cet objectif nous avons voulu détailler les effets de certains facteurs sociodémographiques et géographiques sur l'indicateur de satisfaction utilisé.

I – Mesure barométrique comme agrégat

La mesure de satisfaction dans le cadre barométrique vise à déterminer et à hiérarchiser les leviers d'actions opérationnels, à partir de la mesure du niveau de satisfaction globale d'une population de consommateurs. Selon le modèle employé, le questionnaire sera conçu de façon à approcher les différents attributs de la satisfaction (exemple : la satisfaction envers l'accueil, la satisfaction du service clients, ou encore la propreté d'un magasin), ou de façon à mettre en relation plusieurs concepts psychologiques liés à la satisfaction globale par des items explicatifs (qualité, performance, valeur, image de marque). Dans les deux cas la mesure exige l'interrogation d'un échantillon représentatif des clients de l'entreprise.

En règle générale, la satisfaction dans les baromètres se mesure directement en une seule question : quel est votre niveau de satisfaction envers le produit/service ?; soit indirectement, à travers les attentes, les problèmes rencontrés ou les manifestations à suggérer. Nous trouvons aussi une mesure globale de la performance et des évaluations concernant chaque élément de l'offre. Il en vient donc « une satisfaction générale » et des satisfactions par attributs sur lesquels les individus sont interviewés⁶¹.

Les enquêtes de satisfaction dans le cadre barométrique, intègrent aussi différents éléments relatifs « l'intention du ré-achat, la recommandation, l'image ou encore la confiance » qui normalement devraient être liés à la satisfaction. Aussi on trouve des éléments opérationnels liés à la satisfaction comme le traitement des réclamations ou encore la qualité/prix fournis. Une enquête de satisfaction produit donc des données issues de la récolte des avis des clients à un instant donné et relativement aux différents éléments supposés influencer la satisfaction clients. Ces sondages cherchent avant tout à déterminer les critères pour lesquels une amélioration doit être réalisée en priorité et un niveau de satisfaction doit être atteint, pour que la satisfaction soit optimale.

Toutefois les avis récoltés sont des informations relatives à un ensemble d'individus qui présentent un premier point commun qui est d'être client de l'entreprise. La mesure obtenue sur cet ensemble est une mesure cumulée ou agrégée par rapport à la satisfaction de chaque consommateur, pour fournir une satisfaction microéconomique. Il s'agit en effet d'un indicateur de la satisfaction d'une population. Toute procédure de mesure dans les baromètres est donc une « résultante agrégée », ceci suppose aussi

⁶¹ Les satisfactions mesurées pour chaque attribut sont utilisées pour expliquer la satisfaction générale. Par exemple, on s'intéresse à la satisfaction ressentie des consommateurs sur les différentes composantes du service/et ou le produit, pour expliquer le jugement global. Ce type de procédure est souvent nommée le modèle multi-attributs de la satisfaction. Ce modèle consiste à mesurer la satisfaction par ces attributs.

qu'elle est calculée de manière identique (même questionnaire) et de façon continue dans le temps (plusieurs vagues) et sur un ensemble de consommateurs. Cette mesure cumulée présente un caractère contradictoire par rapport à la définition même de la satisfaction, elle ne fait pas de référence aux attentes, normes ou autre standard de référence comme les dimensions cognitives et émotionnelles. La seule mesure disponible est le résultat sur le **niveau de la satisfaction** (agrégée) et surtout cumulé (plusieurs transactions)⁶² du fait du temps (plusieurs vagues). Donc loin d'être une satisfaction transactionnelle, la résultante de cette mesure est une combinaison de la satisfaction et de ces concepts voisins, dépendante du temps et conditionnée au niveau de la population générale. Ce fait est lié à la pratique barométrique relative aux formulations des questions qui ne concernent que le niveau de satisfaction par rapport à un attribut. En ne formulant qu'une seule question l'avis recueilli englobe la performance perçue, la satisfaction, la qualité, l'image et tous autres concepts proches de la satisfaction. De plus, comme la mesure concerne un ensemble d'individus, elle est plutôt une mesure microéconomique (indicateurs de satisfaction au niveau de la population) qui reste très différente de celle mesurée dans le cadre académique⁶³. Ce cas de figure crée un décalage entre la définition de la satisfaction et sa mesure dans le cadre barométrique. Cette séparation se justifie d'une part par les objectifs du praticien qui cherche à connaître le niveau de satisfaction (exprimé par des indicateurs) et non à comprendre sa formation (chercher les liens) ; et d'autre part par le fait que le sondeur suppose que le paradigme de disconfirmation, se réalise chez l'individu qui intègre naturellement le rapport entre attentes et performance perçue. Ainsi le praticien suppose que le niveau de satisfaction est une résultante qu'on peut capter directement en posant la question directement⁶⁴.

En somme, la mesure de satisfaction barométrique est une mesure relative au temps et concerne une population. Cet aspect de la mesure de satisfaction inclut dans sa construction une dimension sociale (interaction entre individus) qui est présente que dans ce type de mesure microéconomique.

⁶² La littérature a distinguée deux formes : la satisfaction cumulée et transactionnelle. Pour les baromètres nationaux c'est la satisfaction cumulée qui est considérée (Fornell, 1996).

⁶³ La mesure dans le cadre académique est construite à la lumière de tous les éléments influents (déterminants et conséquences)

⁶⁴Voir l'article de D. Ray : mesure de la satisfaction clients : pratique et besoin actuels des entreprise, revue française de Marketing, N°194-2/5, 2004

II – Comparer des agrégats : sources des différences et stabilité des indicateurs

Dans la pratique, les indicateurs produits sont des instruments de comparaison. Nous nous entendons par comparaison toutes procédures de rapprochement entre deux indicateurs ou plus, calculés par unité dans un but de confronter les résultantes. La comparaison se distingue ainsi en deux formes : transversale et temporelle. L'objet des comparaisons qui est le point d'ancrage de notre travail, est l'unité de gestion de l'activité de l'entreprise : différents points de ventes. Cette unité que nous qualifierons de naturelle, est la seule sensée être contrôlée par les firmes. Ainsi les comparaisons qui vont nous intéresser seront celles effectuées entre magasins (transversale et longitudinale) et entre les agrégats de tous les magasins d'une vague à une autre.

Dans une enquête de satisfaction les indicateurs peuvent prendre la forme d'une moyenne, ou un taux de satisfaction. Le fait de comparer des taux ou moyennes par unité, sous-entend que les unités sont déjà comparables. Être comparable dans le cas de la satisfaction, signifie que la population de consommateurs restreinte à chaque unité de mesure produit un avis (ici la satisfaction) sans l'influence des caractéristiques propres à chaque unité. D'un autre côté dans le cas de la comparaison temporelle, les indicateurs doivent à tout prix faire état de l'évolution, car par construction, le temps a un effet qui a pour cause les actions entreprises dans chaque période. Des actions qui sont par construction évolutives.

Ainsi pour les comparaisons transversales, les différences doivent être le fruit du seul fait de l'action marketing. Dans la comparaison longitudinale, la différence doit être visible et une stabilité ne peut être que contingente.

Dans la pratique ces cas de figure sont loin d'être réalisés car une population est naturellement organisée par catégories sous forme d'unités qui peuvent prendre plusieurs formes :

- Unité caractérisant l'individu : toutes les variables sociodémographiques
- Unité relative à l'environnement de l'individu : les variables géographiques, région et pays etc.

Ces unités déclinées ainsi deviennent des modalités sur lesquelles les comparaisons sont sous-jacentes. Les comparaisons sont effectuées directement sur ces

unités ou sur d'autres unités incluant ces dernières.⁶⁵ La firme cherche toujours à partir des résultats des enquêtes, à savoir si telle ou telle unité possède des scores de satisfaction supérieurs aux autres. Puis à l'issue des résultats, les entreprises hiérarchisent les unités pour inférer les améliorations futures de leurs activités. Elles procèdent aussi, conditionnellement à chaque unité, à une comparaison dans le temps pour juger si les actions passées ont eu les effets escomptés. Ici, on retrouve clairement l'intérêt des indicateurs comme outils de contrôle et de pilotage.

Cependant, en générale, ces unités sont traitées de façon indépendante par comparaison sommaire des indicateurs à l'intérieur de chaque modalité. Dans la pratique, ces indicateurs de satisfaction (moyenne, taux ou score) sont calculés sur chaque unité puis comparés sans se soucier de la liaison qui peut y avoir entre l'ensemble de ces unités. Ainsi en agissant de la sorte, les différences constatées dans les scores de satisfaction sont attachés au seul fait de ce qu'on appellera « l'effort Marketing ». Ce terme inclus toutes les dimensions de services/ou de produits qui agissent sur la satisfaction. Par exemple un taux de satisfaction élevé dans un Magasin A par rapport à un magasin B, signifie que les services /produits dans le magasin A sont plus satisfaisants. D'un autre côté, une stabilité dans le niveau de satisfaction d'une période à une autre, sera interprétée comme une stationnarité dans la satisfaction clients. L'effort fourni précédemment n'aura pas permis de la faire évoluer. Ce type de constatations nous paraît réducteur, étant donné la présence des deux points suivants :

Une interaction peut exister entre unités, par le fait que l'indicateur n'est qu'une représentation imparfaite, calculée sur une population, qui est par essence hétérogène dans sa composition.

A – L'interaction naturelle entre unité dans la mesure barométrique

Nous avons mis en lumière plusieurs types d'unités de comparaison qui peuvent exister lors de l'utilisation des indicateurs de satisfaction. Ces unités présentent par nature une structure hiérarchique qui n'est jamais prise en compte comme telle. En effet, un magasin est niché dans une région qui elle-même appartient à un pays. Ceci signifie qu'un individu fréquentant le magasin est conditionné par son appartenance à une catégorie socio-économique, de sa région d'habitation et de son pays. Cette structure

⁶⁵ Nous pensons par exemple au cas du magasin qui est forcément niché dans une région. Une comparaison à partir des magasins est donc dépendante de la région.

hiérarchique est donc susceptible d'apparaître lorsque l'indicateur est évalué statistiquement au niveau de la population. Récolter des avis de consommateurs dans un magasin donné, risque donc d'être dépendant de la région et du pays. Sous ces conditions, une différence constatée dans les scores de satisfaction (moyenne ou taux de satisfait) entre deux magasins dans deux régions et ou deux pays différents ne peut pas être attaché au seul fait de l'activité de la firme. Dans ce cas, les unités sont par nature dépendantes, la non prise de cette interaction peut causer des biais lors de la comparaison transversale et temporelle.

Dans le cas de la comparaison transversale, la présence d'une structure de la sorte peut amener à des fausses conclusions dues à l'hétérogénéité des unités (chaque unité appartient à une division) qui ne sont pas comparables par substrat. Ainsi identifier une supériorité d'une unité, en terme de niveau de satisfaction, peut être fortement altérée par la structure hiérarchique. Ce problème est géré en partie dans la procédure de sondage (sondage par grappe) mais jamais quand il s'agit d'une comparaison entre les unités. Les sondages rendent fidèles les échantillons en représentant la population, mais ne permettent pas de contourner les biais de comparaison comme ceux indiqués auparavant. Car si la structure hiérarchique est respectée lors du sondage l'effet de la structure quant à lui n'est pas spécifié mais reproduit à l'identique dans l'échantillon.

Dans le cas de la comparaison temporelle ce chevauchement des unités qui se répète dans chaque vague peut induire une certaine stabilité vue globalement, c'est à dire en agrégeant cette fois les unités entre elles, ex : ensemble de magasins. Ainsi il est légitime de se poser la question que dans un cas de contraste entre deux périodes, la cause peut être autre que celle liée à l'activité de l'entreprise et peut être liée à des causes plus profondes qui dépassent l'action des entreprises pour satisfaire leurs clients. Notamment par l'effet supposé de la structure hiérarchique des unités qui peut perdurer dans le temps et causer ainsi une certaine stabilité temporelle des indicateurs de satisfaction.

B – Hétérogénéité de la population sur laquelle la mesure est basée

Jusqu'à présent la satisfaction a été souvent expliquée par des dimensions psychologiques liées, aux attentes de l'individu et à la performance du produit/service qu'il perçoit (disconfirmation des attentes), à sa perception des résultats d'un produit/ou d'un service compte tenu des coûts engagés (théorie de l'équité) et à sa tendance à attribuer le succès ou l'échec au producteur du produit/ ou service (théorie de

l'attribution). Dans tous ces paradigmes, le concept de satisfaction a été considéré comme une résultante individuelle qui s'agrège au niveau de la population pour produire des indicateurs de niveau de satisfaction, relatifs à une somme d'individus isolés.

Toutefois, la considération « individuel » est rarement vérifiée comme telle au niveau de la population des consommateurs. Ce fait est dû à la négligence d'une activité collective présente au niveau de la population des consommateurs. D'autant plus qu'un individu n'est jamais isolé et ne vit qu'en relation avec un groupe même lorsque il s'oppose à ce dernier. Ainsi il existe des « groupes sociaux » qui peuvent présenter des traits communs où chaque individu est conscient de son appartenance. Ces groupes peuvent influencer le comportement du réachat donc de la satisfaction en amont. De plus il existe déjà un mal entendu dans la mesure de satisfaction barométrique, elle est considérée comme une composante latente psychologique liée à l'individu mais la mesure est effectuée au niveau de la population. Comme cette population est organisée naturellement en catégorie (âge, sexe, CSP, niveau d'éducation etc.), la mesure finale comme une résultante moyenne de tous les individus paraît comme une mesure pondérée par un ensemble de catégories qu'il faudra identifier.

Cette segmentation en catégorie naturelle de la population de consommateurs peut agir au niveau de l'indicateur en surévaluant et sous évaluant les résultats des contrastes entre niveau de satisfaction, sans pour autant que les faits observés soient liés à l'action marketing de l'entreprise.

III – Indicateurs de satisfaction face aux facteurs de contingence

L'indicateur de satisfaction est le résultat d'un calcul statistique, à partir d'un ensemble de mesures individuelles. Tout l'intérêt de notre démarche théorique qui suivra est axé sur l'évaluation de l'agrégation des mesures individuelles (ensemble de clients) pour obtenir des indicateurs répondants aux exigences du contrôle de la performance. L'existence supposée de problématiques de comparaison transversale et longitudinale expliquée par la présence hypothétique des facteurs de contingence, nous a orienté à proposer un cadre théorique de leurs action. La procédure que nous proposons est d'ordre statistique. Elle vise à expliquer l'existence des facteurs de contingence lors de l'évaluation de l'indicateur de satisfaction. Dans cette perspective, l'existence des facteurs de contingence est confrontée au fonctionnement de l'indicateur de satisfaction comme indice de performance afin d'étudier sa robustesse. Ce travail a permis de définir

trois propositions, résumant l'action des problématiques rencontrées (transversale et longitudinale) que nous testerons empiriquement afin de savoir si ces propositions sont fiables devant l'action des facteurs de contingence.

A – Facteurs de contingence : une source d'erreurs systématiques

Les facteurs contingents, dans la mesure de satisfaction, sont des quantités qui agissent sur le processus de mesure. Ils engendrent des différences systématiques dans les résultats en se restreignant à des unités (ex : niveau de satisfaction élevé pour les femmes par rapport aux hommes). Ils peuvent produire une stabilité lorsque la mesure est globale (agrégation des unités). Ces facteurs sont des sources de biais dans la procédure de mesure qui peuvent être attachés à deux sources d'erreurs qui sont : l'erreur aléatoire et l'erreur systématique.

- Une erreur aléatoire pourrait être vue comme une erreur due à un aléa dans la mesure. Cette erreur change d'une expérience à une autre, elle ne peut pas être déterminée avec précision, mais sa variance peut être estimée grâce aux lois de probabilité. Cette erreur est vue comme la résultante de la fluctuation dans la mesure au tour de la vraie valeur qu'on veut déterminer. De manière plus générale, ce type d'erreur, se pose chaque fois que l'on considère un *échantillon aléatoire* d'individus ou d'objets.
- Une erreur systématique quant à elle est directement liée à l'influence extérieure qui se conjugue au phénomène. Elle cause soit, une surévaluation dans la mesure soit une sous-évaluation. La prise en compte de ces imprécisions (facteur) demande plus d'informations qui sont souvent non directement liées aux problèmes considérés. Contrairement à l'erreur aléatoire, l'erreur systématique ne sera pas diminuée par une série de mesures (cumul de données). Elle doit être repérée par l'expérimentateur et éliminée dans le meilleur des cas sinon elle doit être au moins contrôlée. D'où l'intérêt d'une part d'identifier les effets de toutes sources contingentes au phénomène en corrigeant leurs influences et d'autre part de contrôler leurs effets, c'est-à-dire raisonner conditionnellement à chaque facteur.

L'erreur distinguée sous ces deux formes : aléatoire ou systématique est bien connue de l'expérimentateur qui fait tout son possible pour l'éviter. Un des moyens qu'il possède est de prendre un maximum d'informations (une grande source de données) pour atténuer

l'erreur dans la mesure. Toutefois cette considération n'est pas très efficace pour corriger l'erreur systématique. Cela est dû majoritairement à la difficulté d'identifier cette erreur soit par ignorance de sa présence, soit par une incapacité de la corriger.

Afin d'illustrer le rôle respectif de chaque source d'erreur, nous proposons de comparer leurs fonctions en faisant une analogie avec le tir sur cible (figure).

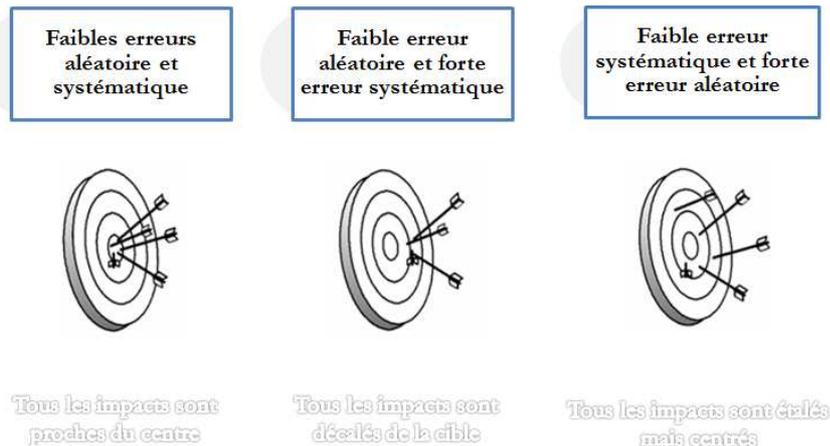


Figure 12:Le rôle des deux erreurs systématique et aléatoire

Ainsi comme le montre le graphique, les erreurs peuvent avoir des effets non négligeables suivant le degré d'importance de l'une ou l'autre. Dans notre cas les facteurs contingents à la mesure de satisfaction, sont liés à ces deux formes d'erreur, toutefois leurs actions sont plus attachées à l'erreur systématique car leurs effets (surévaluation, sous-évaluation, stabilité) ont un caractère systématique.

Par conséquent nous donnerons le statut de **facteurs contingents à la mesure de satisfaction** à toutes sources, mesurables ou non, responsables d'une erreur systématique dans la mesure (au niveau de la population) et qui est non redressée par le cumul des données⁶⁶. D'un point de vue mathématique : En appelant **S** la mesure de satisfaction d'une population, **f** la fonction explicative de sa réalisation et **les attributs** ou variables explicatives de sa formation. Le modèle peut s'écrire :

⁶⁶ Un nombre de données important nous sommes assurés que l'erreur aléatoire lors de l'évaluation d'un indicateur (agrégé à partir d'un ensemble de mesures individuelles) est minime.

$$\text{Satisfaction} = F(\text{attributs}) + \text{Erreur}_{\text{systématique}} + \text{Erreur}_{\text{aléatoire}}$$

Ces deux erreurs apparaissent lors de la considération agrégée (au niveau de la population) et leurs influences touchent le niveau de satisfaction obtenu via les indicateurs, sur l'ensemble des individus.

B – Facteurs de contingences : considération statistique

Selon la définition des facteurs de contingences, nous avons énoncé deux postulats :

- a) Un élément prend le statut de facteurs contingents, dès lors qu'il engendre un biais systématique dans la mesure. En d'autre terme les facteurs contingents conditionnent la mesure. Dans chaque niveau du facteur la mesure obtenue diffère.
- b) Ces facteurs contingents ne sont présents que dans la situation des mesures agrégées sur la population. C'est à dire qu'ils concernent seulement les mesures effectuées sur un ensemble d'observation $(x_i \in I)$ où I est un support fini indexant les observations. Ainsi par définition, les facteurs contingents sont des quantités qui provoquent des problèmes de non échangeabilité des observations.

➤ **Le principe d'échangeabilité, le théorème de Finetti**

Une famille $(x_i \in I)$ est dite échangeable si pour toute permutation de son support I , la loi jointe de ces familles est identique (Finetti 1937). Ce concept très essentiel en statistique sert à justifier l'idée d'attribuer les mêmes lois de probabilité aux observations lors de l'inférence statistique. De plus l'importance du théorème de Finetti est d'établir une liaison entre la notion d'échangeabilité et la notion d'indépendance. Ainsi le concept d'échangeabilité est fondamental pour poursuivre le raisonnement statistique suivant les hypothèses « **d'indépendance des observations et identiquement distribuées** » i.i.d.

Le principe d'échangeabilité des observations assure la vraisemblance de la notion « i.i.d ». Sans cette dernière aucune inférence statistique ne doit être envisagée⁶⁷. Donc aucune généralisation de résultats ne peut être admise comme juste.

Dans la pratique, la notion d'échangeabilité revient à traiter symétriquement les observations. Ceci veut dire que l'observation n°1 peut être échangée avec toutes autres observations sans risquer de perturber les résultats. Cela veut dire que le numéro (position de l'observation dans l'échantillon) est indépendant de tout ordre. Un exemple où les données sont ordonnées est le cas des données indexés par le temps. Dans ce cas on ne peut pas mélanger les valeurs des observations au risque de fausser les résultats.

Ainsi le principe d'échangeabilité, permet de garantir que les observations prises indépendamment de l'indice « i » apporterai la même information lors de l'inférence statistique. Ce qui se traduit dans une considération purement statistique par le fait que si deux observations x_1 , x_2 dépendent du même paramètre, elles doivent apporter la même information et doivent conduire à la même inférence, principe de vraisemblance, (Robert, 2005). Ce principe garanti la cohérence des étapes d'inférences statistiques.

Dans notre cas, si un facteur contingent, supposé influent, conditionne les observations sous forme de groupe, cela implique que la possibilité d'échanger des observations entre groupes n'est plus envisageable. Or si l'analyse statistique, la mesure agrégée, est faite indépendamment du facteur, le principe d'échangeabilité risque de ne plus être respecté, donc la computation statistique qui suivra présentera une erreur dans la mesure.

Ainsi il paraît primordial de raisonner conditionnellement à des facteurs contingents, s'ils sont influents, cela va garantir l'échangeabilité par niveau de facteur, donc le respect du principe. Le fait d'ignorer leurs existences lorsqu'on veut expliquer statistiquement une variation de la satisfaction par attributs, peut produire des résultats erronés (sur paramètre du modèle) du fait de non-respect des niveaux des facteurs supposés influents.

De ce fait, les facteurs contingents agissent comme *des niveaux conditionnant* les observations et induisent des erreurs systématiques lorsqu'ils ne sont pas pris en compte. Ce raisonnement nous conduit à donner la définition suivante des facteurs contingents :

⁶⁷ Cette condition est au cœur des tests paramétriques. Elle doit être vraie pour que les outils probabilistes pour mener l'inférence (test statistique), soit utilisés.

Les facteurs contingents sont des quantités qui causent des erreurs systématiques lors de la mesure statistique des baromètres de satisfaction considérée d'un point de vue agrégé. Leurs biais sont présents lors de la mesure agrégée (au niveau d'une population). Ils conditionnent les observations et causent un problème d'échangeabilité des observations lorsqu'ils ne sont pas pris en compte. Ils imposent par leur présence une structure des données qui lorsque elle est ignorée, la mesure agrégée est évaluée avec erreur.

C – Différents facteurs contingents liés à la mesure de satisfaction

Ce sont des facteurs qui provoquent des erreurs systématiques. Il s'agit des facteurs attachés aux caractéristiques des individus et de leur environnement. Ces facteurs présentent la faculté de conditionner les individus dans des groupes qui présentent des particularités communes. De plus ces facteurs de contingence ne sont détectables qu'au niveau d'une population d'études. Ils imposent une structure hiérarchique qui conditionne la mesure en créant une différence entre les modalités de calcul ou en induisant une stabilité temporelle lorsque ces derniers ne sont pas pris en compte ; c'est à dire lorsqu'il y a agrégation de toutes les unités.

➤ **Facteurs relatifs aux caractéristiques des individus**

Nous pensons notamment au sexe, à l'âge et au niveau de l'éducation ou encore à la catégorie socioprofessionnelle. Ces facteurs spécifiques aux individus sont liés à l'aspect même de ces derniers dans la société. Ils deviennent des éléments qui d'une part segmentent la population générale et d'autre part créent des groupes qui peuvent être vus comme des unités qui découpent la population et engendrent une interaction intra-groupe.

Ces catégories de facteurs, par leur qualité de contingence peuvent être plus ou moins importantes selon la situation considérée⁶⁸ (ex : le secteur, le type de produit ou service). Ils sont tous hypothétiquement influents car ils sont attachés à la même structure, ici c'est l'individu lui-même. Quand ces derniers provoquent des différences systématiques, ils sont à l'origine d'une interaction, qui quand elle n'est pas prise en

⁶⁸ Dans un secteur de l'habillement, les produits diffèrent suivant les hommes et les femmes. Donc le jugement d'un client appartenant à l'un des deux sexes, donne un avis sur des choses qui sont différentes par essence. Ainsi le facteur sexe, joue ici un rôle qui conditionne les niveaux de satisfaction et devient important par nature. Par cela, il en vient que si les effectifs des deux catégories ne sont pas représentatifs de la population, la considération agrégée (indicateur sur l'ensemble de l'échantillon) peut être sensiblement différente lorsqu'on questionne beaucoup d'hommes ou inversement beaucoup de femmes.

compte elle peut biaiser les résultats finaux en déséquilibrant les comparaisons. Par exemple les résultats d'une comparaison entre les niveaux de satisfaction de deux magasins, calculés sur des populations hétérogènes suivant les sexes ou l'âge, ne peut pas être considérée comme robuste.

D'une façon générale, ces facteurs provoquent des interactions qui influencent indirectement la satisfaction en modulant ses antécédents, ou en ayant des effets directs touchant la mesure de satisfaction elle-même, du fait de l'agrégat.

➤ **Facteurs relatifs à l'univers des individus le cas des facteurs géographiques**

Ces facteurs représentent des éléments relatifs à l'univers où les individus évoluent, habitants de la même région, même pays. Ces facteurs agissent lorsque le besoin de comparaison entre les niveaux de satisfaction devient nécessaire. Ce sont des facteurs contingents qui causent des emboîtements dits géographiques, causant ainsi par essence une différenciation dans la population des individus. Par exemple, deux magasins implantés dans deux endroits n'ont pas les mêmes clients. Ces clients provenant de zones géographiques différentes et ayant des profils (variables sociodémographique), ils n'ont pas forcément les mêmes comportements⁶⁹

Ces facteurs agissent comme des classes dans lesquelles la population est caractérisée par une certaine homogénéité. Ils peuvent fausser l'interprétation des résultats en introduisant l'erreur systématique dite « écologique ». Cette erreur est présente lors du calcul de l'indicateur et cause des problèmes de compromis entre les unités géographiques. Elle se manifeste comme des écarts systématiques lors de la comparaison transversale (entre magasin) ou comme une stabilité apparente lorsqu'il s'agit d'une comparaison temporelle sur une mesure globale, agrégation des unités, ici agrégation sur l'ensemble des magasins.

Ils présentent les mêmes actions que les facteurs liés aux caractéristiques des individus, mais à la différence des premiers ils les englobent. Les facteurs géographiques peuvent être vus comme des quantités **macro** attachées à l'unité de contrôle et aux individus qu'ils la constituent. Les facteurs liés aux caractéristiques des individus peuvent être considérés

⁶⁹ Ce caractère relève du géomarketing. Les points de ventes étant dans des zones géographiques différentes, ils n'influencent pas les mêmes individus. Les points de ventes sont forcément affectés par les profils de la zone où ils se trouvent. Deux magasins identiques en termes de services, situés dans des zones de profils sociodémographiques différents, n'ont pas forcément les mêmes performances.

comme des quantités **micro** relatives uniquement aux individus. Par exemple, quand le magasin est l'unité de contrôle, les facteurs **micro** ne sont pas stables dans les magasins (les clients évoluent naturellement sur ces facteurs), quant aux facteurs de type **macro** (région, pays) l'ensemble des clients n'évoluent pas forcément sur ces facteurs. De plus comme un magasin sera toujours dans la même situation géographique, ils attirent plus de 80% de ces clients d'une zone bien spécifique (zone attraction). Les facteurs **macro** deviennent des caractéristiques attachées aux clients et aux magasins. Pour la partie « englobante » des facteurs **macro** (ici géographique), les clients appartiennent avant tout à une région puis à une catégorie sociodémographique. Ainsi la conjugaison de ces facteurs induit l'existence d'une structure hiérarchique qui conditionne les avis clients. Un conditionnement qui devient visible dans les comparaisons à partir des considérations agrégées, c'est-à-dire lorsque la structure hiérarchique n'est pas prise en compte.



Figure 13: Actions des facteurs de contingence

Section 5 – Propositions de test de représentativité et de sensibilité des indicateurs de satisfaction

Dans la perspective de mieux connaître leurs clients, les professionnels des études font déjà la distinction entre les catégories socioprofessionnelles, la région ou le pays, en traitant parfois ces groupes différemment. Mais lorsque le besoin de comparaison apparaît, la procédure n'est jamais accompagnée par une analyse prenant en compte la structure qui peut être sous-jacente à ces variables. Souvent ces facteurs sont utilisés comme des variables discriminantes sur lesquelles des typologies d'individus sont construites. Cette classification via les variables sociodémographiques et environnementales, est faite *a priori* en supposant une influence contingente. Il faut dire

que ces constatations ne sont pas spécifiques à la satisfaction et touchent majoritairement tous les aspects du comportement des consommateurs lorsque l'on veut le mesurer à travers des questionnaires.

Néanmoins, dans cette optique, une question pertinente surgit dans la procédure de comparaison entre différents magasins d'une même entreprise. Elle concerne la validité de cette dernière. Il est clair que chaque entreprise possède un type de clientèle souvent caractérisé par une répartition suivant les données sociodémographiques et environnementales. Si cette hétérogénéité est présente dans la population un rapprochement entre les résultats de satisfaction et impliquant deux populations différentes par leurs caractéristiques individuelles (âge, sexe, revenu, niveau d'éducation) est déraisonnable. Ce problème apparent d'hétérogénéité des populations fait qu'une comparaison brute est sujette à erreurs, surtout quand on soutient l'existence d'un effet de ces facteurs.

Dans nos précédents développements nous avons pointé l'existence de problématique de comparaison : différences systématiques et stabilité. Les hypothèses que nous soutenons, indique que les différences systématiques lors des comparaisons, ont des sources contingentes, la stabilité des indicateurs est conditionnée au niveau du calcul. Les propositions qui vont être formulées sont relatives à ces hypothèses. On veut voir si l'indicateur est touché dans ses fonctions, c'est-à-dire dans le contrôle de performance des unités, ici c'est l'unité magasin. Nous voulons connaître si les facteurs qui causent des différences systématiques dans le niveau de satisfaction, sont préjudiciables à l'indicateur lui-même. Car il se peut que les différences systématiques soient justes naturelles. C'est pourquoi nous avons choisi de confronter les indicateurs à la présence de ces facteurs de contingence en étudiant trois propositions relatives au fonctionnement de ces derniers comme instrument de pilotage.

Ainsi, dans cette conception pratique, il paraît nécessaire, d'une part de détecter si ces facteurs ont un effet important sur l'indicateur de satisfaction comme instrument de pilotage et d'autre part de tester la validité de ces indicateurs en étudiant la représentativité et la sensibilité de ces derniers dans leurs fonctions de comparaison⁷⁰.

En ce qui concerne la comparaison inter-entreprises, ces facteurs peuvent jouer un rôle aussi important qu'en interne, et une comparaison directe n'est pas vraiment

⁷⁰ Pour que les comparaisons soit valides, l'indicateur doit apporter une conformation fiable indépendante de tout aspect étranger à ce que l'indicateur doit mesurer.

recommandée. Ce cas ne va pas être traité dans notre recherche à défaut de moyens. Néanmoins il faut nécessairement contrôler ces facteurs spontanément, pour que la comparaison soit valide. En somme les facteurs contingents, stabilisateurs et ceux causant une différence lors des comparaisons, doivent être identifiés en explicitant leurs effets, puis de les contrôler ou de les redresser afin de fidéliser la mesure de satisfaction barométrique.

Suivant les constatations de présence de certains facteurs de contingence, nous soutenons le fait que l'intégrité d'un indicateur de satisfaction peut être mise en défaut. Le but dans cette démarche est de pouvoir mettre en compétition l'utilisation des indicateurs face à l'existence des effets des facteurs de contingence en énonçant des propositions qui vont former le socle de notre recherche.

Les conditions de formulation des propositions viennent de deux considérations : la première est relative à l'indicateur comme un instrument de pilotage et la deuxième est attachée à l'influence supposée des facteurs de contingence. Dans cette vision la procédure de test vise à formuler des propositions relatives aux fonctions des indicateurs de satisfaction dans la pratique. Il faut rappeler qu'un indicateur peut avoir plusieurs supports (questions) soit une satisfaction globale soit une satisfaction par attribut de service ou de produits. Les attributs reflètent les perceptions des efforts marketings par les clients et sont censés expliquer la satisfaction globale des interviewés. Nous nous intéressons ici à l'étude des caractéristiques d'un indicateur (satisfaction globale/ou par attributs) dans son rôle comme instrument de contrôle entre des unités de ventes (magasins).

La fonction principale des indicateurs de performance est de permettre des comparaisons et d'attribuer les différences à des actions ou des événements identifiables. Lorsque l'on compare deux unités et que l'on obtient une valeur pour un indicateur (satisfaction globale/ ou attributs) donné S_1 est supérieure à S_2 , elle doit être attachée simplement à l'effort marketing (représenté par les attributs) ou à une différence d'efficacité managériale. Un indicateur doit aussi permettre de détecter la performance quelle que soit sa nature: moyenne ou taux de satisfaction et doit représenter fidèlement les informations issues du sondage. Dans la dimension temporelle, l'indicateur doit être sensible par rapport aux actions entreprises à temps $(t-1)$ et cela quelle que soit l'évaluation: au global (agrégat de tous les magasins) ou conditionnellement à chaque magasin.

Comme la satisfaction est considérée comme un indicateur de performance, sa mesure via les indicateurs doit être fiable. Cette supposition sera la base de notre procédure de test. L'indicateur doit à tout prix être représentatif et sensible pour que les informations fournies par ce dernier soient justes. Dans le développement précédent nous avons vu qu'il pouvait y avoir un risque que les facteurs de contingence agissent sur l'indicateur, donc sur la prise de décision. Ainsi nous nous sommes focalisés sur les propriétés de l'indicateur de performance et sur ses fonctions de comparaisons pour formuler nos propositions. C'est sur ce principe de respect des exigences qu'un indicateur de performance doit respecter, que nous avons formulé trois propositions énoncées comme suit :

***P1** : l'indicateur de satisfaction doit être corrélé de façon homogène aux facteurs d'actions (attributs de satisfaction) quelle que soit l'unité magasin considérée.*

***P2** : Si un indicateur détecte une différence dans une satisfaction globale/ ou par attributs entre magasins, cette différence doit être liée au seul effet de l'effort de Marketing*

***P3** : Un indicateur doit être sensible au temps quelle que soit la dimension du calcul au global (ensembles des magasins) ou par magasin.*

Ainsi ces propositions visent à examiner l'indicateur dans sa structure avant qu'il ne soit utilisé comme instrument de pilotage sur les unités de gestion⁷¹. A la lumière des problématiques de comparaisons, pointées dans le début du chapitre, les propositions visent à vérifier si les qualités d'un indice de performance sont toujours valables pour les indicateurs de satisfaction barométriques.

Ces propositions vont être le centre de la partie axée sur les tests empiriques. La procédure empirique sera basée sur des études de cas. Les trois propositions qui traduisent la robustesse d'un indicateur de performance, ciblent à coller aux propriétés de ce dernier, à savoir : corrélés aux leviers d'action, objectivité, précision et sensibilité.

⁷¹ L'unité de gestion dans notre cas est le magasin, c'est cette dernière qui est l'unité d'intérêt à partir de laquelle l'indicateur est utilisé comme instrument de contrôle. L'idée est de savoir si l'indicateur calculé par cette unité, est toujours fiable devant la considération des facteurs de contingence.

Pour arriver à tester cela nous avons choisi l'unité magasin car les comparaisons par magasin sont plus importantes dans la phase de pilotage procédée par les firmes. Notre objectif est d'une part de vérifier si les propositions sont valables dans le cadre de ces comparaisons et de spécifier l'importance de l'action des facteurs de contingence. Nous raisonnerons par une logique de contre-exemple. Étant donné le fait que les analyses seront axées sur un cas particulier, ici c'est juste la mise en défaut des propositions qui sera envisagée.

En se basant sur l'unité magasin, chaque proposition est traduite par des quantités statistiques⁷² (indices). Un échantillon statistique est donc obtenu (un indice par magasin). A la base de cet échantillon, nous testons la vraisemblance de la proposition en analysant les distributions des indices avec la présence des facteurs de contingence. Les résultats des analyses de ces indices nous permettent de juger des effets de facteurs de contingence et de statuer sur le contenu de la proposition.

Pour la proposition P1, les actions marketing (attributs) doivent être corrélées de façon homogène à l'indicateur de satisfaction quel que soit le magasin. Si une disparité entre cette liaison apparaît sur l'échantillon (représentés par les d'indices), on peut penser que cette disparité est liée à des sources extérieures d'où le rapprochement avec l'existence des facteurs de contingence.

Pour la proposition P2, les différences de niveau de satisfaction entre magasins ne peuvent être que le fruit des actions marketing. Si une différence est détectée par un type d'indicateur, moyenne des notes, mais pas par un autre, taux de satisfaction, il y a un manquement d'objectivité dans l'évaluation de l'indicateur qui est peut être due à des sources externes.

Pour la proposition P3 : la sensibilité des indicateurs doit être visible quel que soit le niveau de calcul par magasin ou au global (quel que soit le magasin). L'étendue de la sensibilité dans le temps ne doit pas être différente dans chaque niveau. Si c'est le cas, des sources externes peuvent être les causes.

Les études des cas qui servent au test des propositions, concernent deux entreprises de la grande distribution. Pour le premiers cas, nous disposons de plusieurs milliers de données récoltées sur pour 221 magasins recouvrant plusieurs régions et deux

⁷² Comme nous le verrons dans la suite, chaque proposition sera représentée par un indice statistique calculé pour chaque magasin.

pays qui sont la France et Allemagne. Le second cas, exclusivement consacré au test de la proposition P3 (sensibilité dans le temps), compte un nombre important de données et concerne plusieurs régions situées en France sur 16 vagues d'enquête.

Chapitre 3 : Méthodologie de test des propositions

Les propositions mises en jeux dans notre recherche visent à tester l'indicateur de satisfaction à remplir son rôle d'indice de performance. La procédure vise à examiner ses fonctions de comparaisons transversales et longitudinales par unité magasin⁷³. La démarche a pour ambition de déterminer si la présence des facteurs de contingence diminue les qualités de l'indicateur.

Dès le départ, nous nous sommes confrontés à un problème méthodologique, car cet examen s'inscrit dans une vision qui s'éloigne du cadre de ce qui est commun dans une recherche en marketing. Notre premier problème concerne la possibilité de traiter un grand nombre de données et le deuxième se rapporte à la possibilité d'analyser des données qui ne concernent pas forcément les individus directement⁷⁴.

Pour la première difficulté (le volume des données), nous avons adopté une méthodologie qui évite le biais des données volumineuses. En effet, la majorité des tests statistiques classiques, notamment le test **F de Fisher (Anova)**, est sensible aux données. Plus les données sont nombreuses plus le test a tendance à révéler un résultat significatif (effet non nul). Pour remédier à ce problème, nous nous sommes axés sur une méthodologie qui s'affranchit du nombre de données en travaillant directement sur la taille de l'effet⁷⁵. Cette considération nous a amenée aux méthodes de comparaisons bayésiennes, initié par Rouanet et Lépine (1976). Les méthodes de comparaisons bayésiennes sont des techniques statistiques inspirées de la refonte de l'analyse de la variance, qui elle-même est une procédure de comparaison de moyennes. Ces deux auteurs voulant allier les pratiques des chercheurs à la théorie statistique, ont permis de faire émerger un cadre mathématique justifiant les méthodes de comparaisons, dont l'analyse de variance est une composante. Par leurs travaux et ceux qui les ont suivis, ils ont permis de donner un cadre mathématique pour l'analyse des données expérimentales via la notion de l'importance de

⁷³ Nous chercherons à savoir si l'indicateur de satisfaction permet de juger de performance des magasins entre eux et dans le temps

⁷⁴ L'individu correspond ici au consommateur sondé. Dans une littérature marketing axée sur la satisfaction clients, le consommateur est souvent le centre de questionnement et de recherche. Le consommateur constitue l'individu statistique à partir duquel les analyses sont effectuées. Il s'agit en effet d'inférer les conclusions au niveau du consommateur. Dans notre recherche, la problématique concerne l'indicateur (un résumé de notes de satisfaction) au niveau d'un magasin. C'est sur ce niveau que les analyses doivent se porter en s'adaptant à cette considération.

⁷⁵ Cette notion sera définie dans la suite de ce texte, toutefois pour fixer les idées on pourra prendre l'exemple de deux facteurs qui influencent un phénomène. L'intérêt ici est de connaître la force de leur action pour statuer sur leur importance. D'où la notion de taille de l'effet.

l'effet. Par données expérimentales, ils sous entendent toutes procédures de récolte d'informations qui s'oppose à celle des données issues de la simple observation. Comme les baromètres de satisfaction sont des enquêtes d'opinion, donc des plans d'expérience mettant en jeux des questions et des conditions structurant la récolte de l'information, elles s'inscrivent dans le cadre des données expérimentales. De ce fait, cette méthodologie qui sera présentée dans la suite, nous a donné les moyens d'une part de gérer le problème des nombres de données via la considération de l'importance des effets, et d'avoir une méthodologie complète (descriptive et inférentielle) qui peut être dérivée des analyses classiques comme de l'analyse de variance. Ce qui facilite l'interprétation de ses résultats⁷⁶.

Pour la deuxième complication, elle concerne la notion d'indicateur qui n'est pas attachée à un individu mais à un groupe d'individus assemblés sous forme d'unité (dans notre cas le magasin). Cette difficulté a été contournée par l'approche de la méta-analyse. La méta-analyse a servi dans notre recherche à justifier notre approche d'agrégation⁷⁷ (l'analyse sur les données agrégées). La méta-analyse nous a permis de produire directement par le procédé de synthèse de l'information, des statistiques qui relèvent d'un niveau supérieur qui est dans notre cas le magasin et non le consommateur sondé.

La jonction entre les deux techniques : comparaisons bayésiennes et méta-analyse, s'est située au niveau du traitement statistique. Dans une approche classique, la méta-analyse est une méthode qui à partir d'une information agrégée (résultats sur un ensemble d'individus) permet de combiner les informations (plusieurs sources de données) afin d'affirmer ou pas la présence d'un effet d'une variable. Cette approche sert avant tout à gagner dans la puissance statistique. Dans notre cas, nous avons utilisé l'esprit de cette technique non pas pour prouver l'effet de la variable d'intérêt, mais plutôt pour expliquer sa variation avec d'autres variables externes situées au niveau agrégé. La méta-analyse est là pour justifier l'analyse sur des données agrégées. Elle permet en outre de légitimer les analyses et de garantir la démarche d'agrégation à des niveaux supérieurs à des fins d'analyse.

⁷⁶ La force majeure de cette technique est qu'à partir d'une table de variance (sortie d'Anova, F de Fisher, degré de liberté), nous pouvons obtenir la taille de l'effet et faire de l'inférence statistique. Tous ces éléments seront détaillés dans les sections suivantes.

⁷⁷ Par agrégation, nous sous entendons le fait de passer du niveau d'individus au groupe d'individus. Ce passage est accompagné par le calcul d'une statistique résumant le groupe dont l'indicateur de satisfaction est un exemple.

Pour mieux présenter la méthodologie qui sera adoptée, nous présenterons la méta-analyse et son fonctionnement habituel. Nous exposerons la méthodologie de l'analyse des comparaisons bayésiennes, puis nous détaillerons la méthode qui sera utilisée dans cette recherche en explicitant le pont méta-analyse et analyse de comparaison bayésienne. Toutefois avant ces développements, nous consacrerons une section à définir les données sur lesquelles les analyses seront effectuées.

Section 1 – Les données d'analyse : sources et structures

Nos travaux thèse ont été menés au sein de la société partenaire Init-Satisfaction. Cette société est un cabinet d'études marketing spécialiste dans la mesure de la satisfaction et de la fidélité clients. Créée en 1995, INIT est historiquement la première société à être clairement positionnée sur la mesure de la satisfaction clients. INIT a pu collecter un très grand nombre de données par des instruments variés, sur des objets différents, à des multiples moments, dans des lieux différents. Elle dispose ainsi de plusieurs centaines voir plusieurs milliers d'études effectuées auprès de nombreuses sociétés françaises et internationales. L'apport essentiel de l'entreprise à notre projet a été d'apporter un corpus de données uniques dont la volumétrie considérable permet d'envisager des approches méthodologiques rarement employées dans ce domaine.

Les données mises à nos dispositions sont des enquêtes de satisfaction. Ces dernières sont soumises aux consommateurs à la sortie de caisse de magasins. Nous disposons de deux baromètres, un qui concerne une entreprise de prêt-à-porter et l'autre une grande enseigne de bricolage. Le critère central de ces enquêtes est le magasin, en tant qu'objet d'études. Ces enquêtes permettent de sonder l'avis client relativement à la satisfaction globale des magasins et sur un ensemble de critères spécifiques au fonctionnement de ces derniers. Les critères sur lesquels les consommateurs donnent leurs avis servent d'éléments explicatifs de la satisfaction globale. Ce modèle suppose que l'avis de satisfaction exprimé par les clients est fonction d'un ensemble de critères comme : le fonctionnement du magasin, son personnel, le produit commercialisé, les prix pratiqués. Le but de ces questionnaires est de pouvoir analyser les leviers de la satisfaction (les attributs) et de pouvoir réajuster l'offre pour améliorer la satisfaction. Une fois les avis clients récoltés, l'analyse des questionnaires servira à :

1. Établir les critères les plus explicatifs de la satisfaction⁷⁸
2. Produire un tableau de bord pour rendre compte de la performance des unités, dans le temps, et dans l'espace.
3. Synthétiser les informations pour la communication des résultats en interne. Sur ce dernier point, il s'agit de produire des indicateurs simples compréhensibles par tous, pour que les managers (des magasins) soient informés de la satisfaction de leurs clients. Dans la majorité des cas c'est le service des études (situé au siège de l'entreprise) qui est le commanditaire de ces enquêtes.

Ainsi le but de ces enquêtes qui ont servi pour nos travaux de thèse, est de produire des indicateurs pour faire état de la satisfaction clients des entreprises. Ils sont très différents de ce qu'on peut rencontrer dans les recherches en marketing⁷⁹. Cependant ayant un sujet qui concerne les problématiques des indicateurs, ce type d'instrument est bien adapté à nos propos car notre volonté au final est de tester la robustesse de l'indicateur de satisfaction sous ses différentes formes (taux, moyen ou score) dans son rôle d'indicateur de performance des magasins, lorsqu'on les compare entre eux et dans le temps.

Pour les deux baromètres, les questionnaires utilisés ont les structures suivantes :

I – Baromètre B1 : concernant une enseigne de prêt-à-porter :

Les informations issues de ce baromètre vont nous servir à répondre aux deux propositions P1 et P2. La structure du questionnaire peut être résumée par les deux tableaux suivants :

Tableau 4 : Items de satisfactions baromètre B1

Notes de satisfaction sur	Echelle
Le magasin en général (satisfaction globale)	1 à 10
Le caractère attractif des étalages	1 à 10
L'aménagement accueillant de l'entrée	1 à 10
L'espace permettant de circuler entre les tringles	1 à 10
La facilité à s'orienter dans le magasin	1 à 10
Le caractère attractif de la présentation des vêtements	1 à 10

⁷⁸ Le terme « explicatifs » renvoie au fait de déterminer les critères qui engendrent la satisfaction/et ou l'insatisfaction. Dans le cadre professionnel on parle souvent de contribution du critère à la satisfaction globale

⁷⁹ Dans la recherche en Marketing, il est très rare de partir de données existantes. En général, les questionnaires sont produits par les chercheurs puis les données sont récoltées. Dans la recherche en Marketing, la construction des questionnaires représentant bien le phénomène est une étape très importante.

La propreté et l'ordre du magasin	1 à 10
La disponibilité du personnel de vente à vous servir	1 à 10
La qualité des conseils fournis par le personnel de vente	1 à 10
L'amabilité du personnel de vente	1 à 10
La facilité à trouver les cabines d'essayage	1 à 10
La disponibilité du personnel aux cabines d'essayage	1 à 10
La facilité à trouver les caisses	1 à 10
La rapidité du passage à la caisse	1 à 10
L'amabilité et le sourire du personnel aux caisses	1 à 10
La finition des vêtements	1 à 10
La coupe des vêtements	1 à 10
Le style des vêtements	1 à 10
Le choix des modèles	1 à 10
La compétitivité des prix	1 à 10
Le renouvellement des vêtements à la mode	1 à 10

Tableau 5 : Ensemble des informations or notes de satisfaction

Renseignements sur
L'âge des répondants
Le sexe des répondants
La fréquence de fréquentation du magasin
Les adresses des magasins : d'où les variables régions et pays

L'ensemble des variables renseignées dans le tableau (1 et 2), constituent les facteurs de contingence qui seront confronté à l'indicateur dans son utilisation comme indicateurs de performance. Dans la proposition P1, ces facteurs serviront à tester si la liaison entre indicateurs (ici la satisfaction globale) et les critères censés l'expliquer (attributs), est contrastée par la variation de ces facteurs. Pour P2, ces mêmes facteurs seront confrontés à l'indicateur, cette fois pour juger de sa capacité à être discriminant dans ses fonctions de comparaison. C'est-à-dire juger des différences entre les performances des magasins.

Pour ce baromètre, nous avons au total, 221 magasins répartis sur deux pays : **France et Allemagne**. Leur répartition géographique est visible dans la carte suivante :

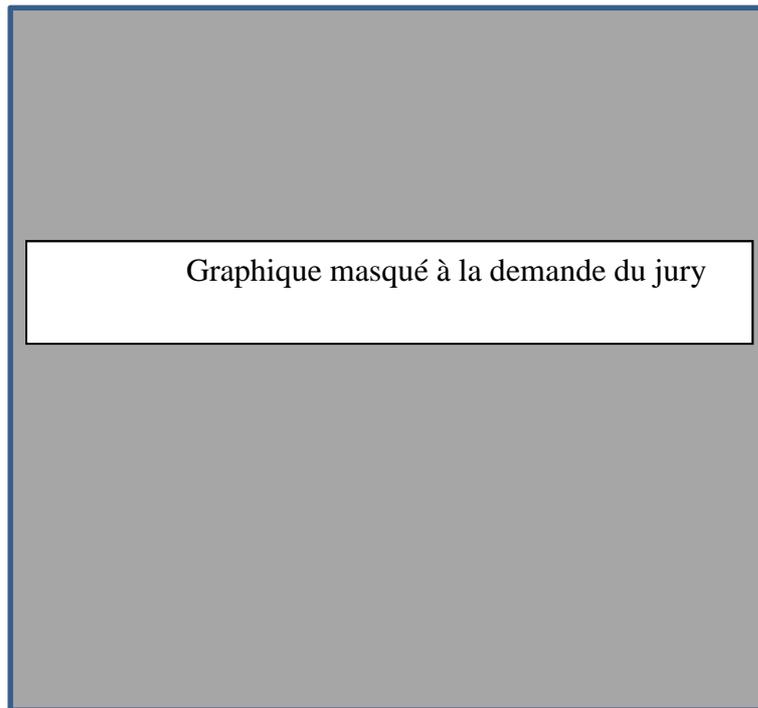
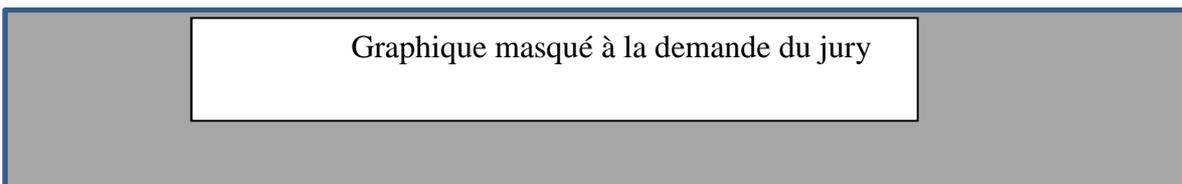


Figure 14 : Répartition géographique des magasins du baromètre B1

Le nombre d'individus concernés dans ces enquêtes est de l'ordre de 172400. Nous avons 58 points de ventes en France et 163 en Allemagne. La répartition par magasin des notes moyennes calculées à partir des résultats de satisfaction globale, est résumée dans le tableau, le graphique et la carte suivante :

Tableau 6: Distribution des notes moyennes de satisfaction par magasin



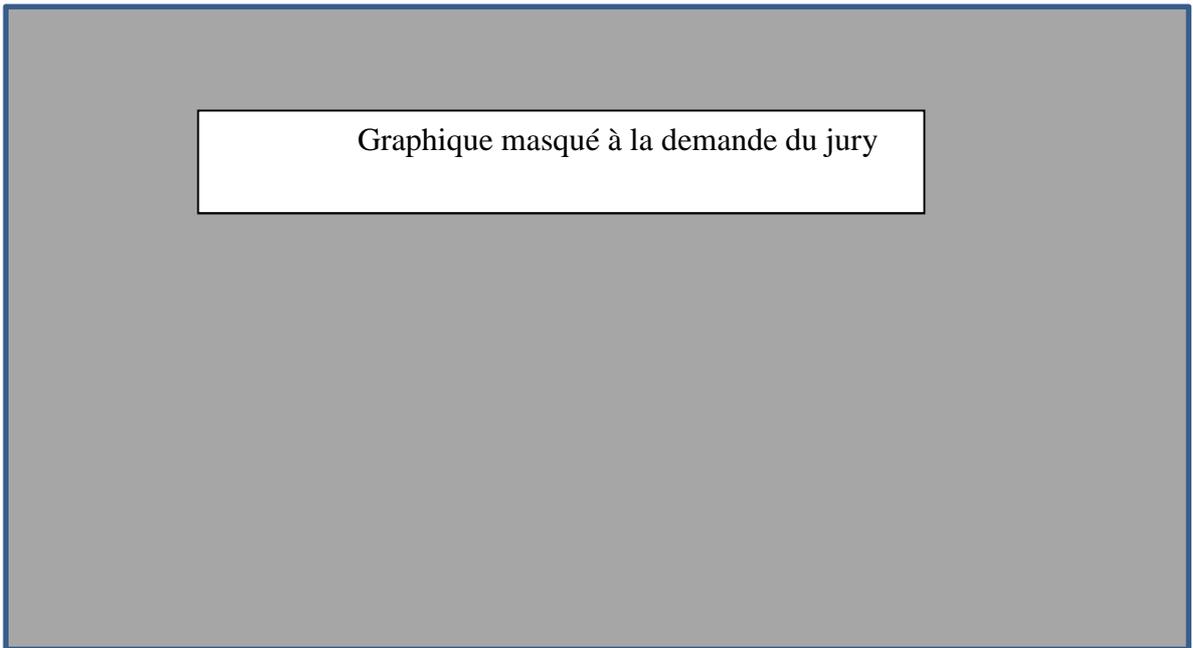


Figure 15: Distribution des notes moyennes de satisfaction par magasin

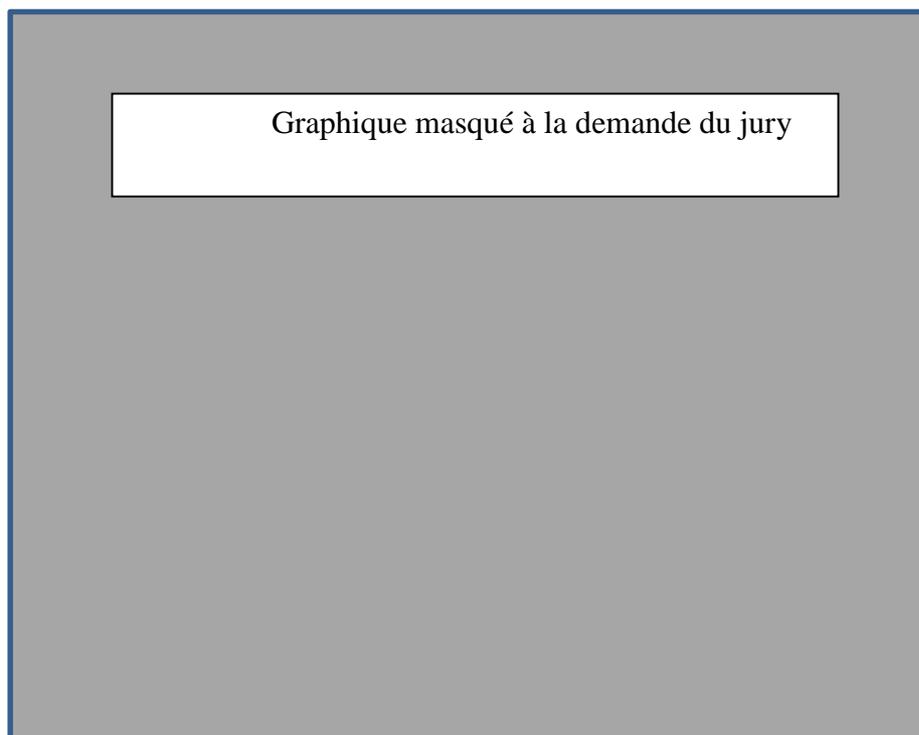


Figure 16: Répartition géographique des notes moyennes de satisfaction par magasin

Pour le besoin de notre recherche, nous avons agrégé les données au niveau du magasin. C'est sur ces niveaux que les analyses futures seront effectuées. Ce baromètre,

servira au test des deux premières propositions⁸⁰. Dans la suite, les données (notes de satisfaction individuelles) seront agrégées pour obtenir des indices aux niveaux des magasins. Ces indices une fois obtenus pour chaque magasin de l'étude, seront les nouvelles données⁸¹ qui serviront aux analyses statistiques en vue de tester les propositions.

II – Baromètre B2 : Concernant une grande enseigne de Bricolage

Contrairement au baromètre précédent, ce baromètre ne comporte pas de note de satisfaction globale. Il s'agit d'un baromètre de suivis dans le temps de critères opérationnels⁸². Dans la suite, ce baromètre servira à tester la proposition P3.

Les magasins de ces études sont situés en France métropolitaine et sont répartis sur plusieurs régions. Pour le Baromètre B2, il s'agit d'une compilation d'études sur 16 vagues trimestrielles sur 4 ans. Pour chaque vague 40000 individus sont interrogés. L'enquête est menée sur 96 magasins dans lesquels plus de 400 clients sont concernés. Les informations disponibles sont agrégées par magasin et concernent plusieurs régions. Les régions concernées sont :

Graphique masqué à la demande du jury

Les régions considérées sont large par rapport au découpage habituel (22 régions en France). Toutefois cette structure spécifique à l'enseigne est aussi conditionnée par la présence ou pas des magasins dans le territoire. Ce regroupement des régions se justifie par le parc actuel de l'enseigne. Nous remarquons aussi que la région parisienne a été distinguée en deux sous régions. Cette procédure propre à l'entreprise d'étude vient du fait que les magasins implantés dans cette région se situent d'une part et d'autre du

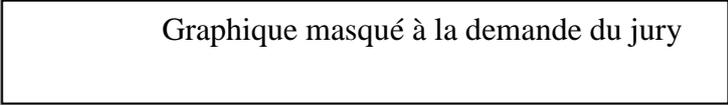
⁸⁰P1 : l'indicateur de satisfaction doit être corrélé de façon homogène aux facteurs d'actions (attributs de satisfaction) quelques soit l'unité magasin considérée. P2 : Si un indicateur détecte une différence dans une satisfaction globale/ou par attributs, cette différence doit être liée au seul effet de l'effort de Marketing

⁸¹ Il s'agit ici de « métadonnées ». ils sont le fruit d'un calcul statistique préliminaire à l'agrégation. La méthodologie qui sera présentée dans la suite, servira à justifier cette agrégation (méta-analyse) et permettra de justifier les analyses supplémentaires (l'analyse fiducio-béyésienne)

⁸² Les critères opérationnels correspondent au levier d'actions qui sont actionnables directement par les managers pour améliorer la satisfaction globale. Par exemple dans le cas d'un mauvais jugement des clients sur la propreté du magasin, l'action corrective sera prise en compte rapidement pour améliorer ce critère et de ce fait la satisfaction globale.

territoire de l’Ile de France (soit EST soit OUEST). Un seul découpage aurait caché des informations.

Ce baromètre servira à tester la proposition P3 : « sensibilité des indicateurs ». Avec la présence de trois niveaux d’agrégation magasin, région et France métropolitaine, le test de la sensibilité aura plusieurs supports. Les données récoltées se présentent comme des notes allant de 1 à 6 (la note de 6, correspond à une satisfaction élevée). L’agrégation par niveau (magasin, région, pays) est effectuée en calculant un score basé sur la distribution des notes. Le schéma suivant résume le calcul du score :

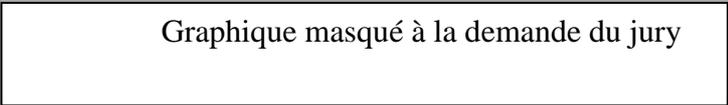


Graphique masqué à la demande du jury

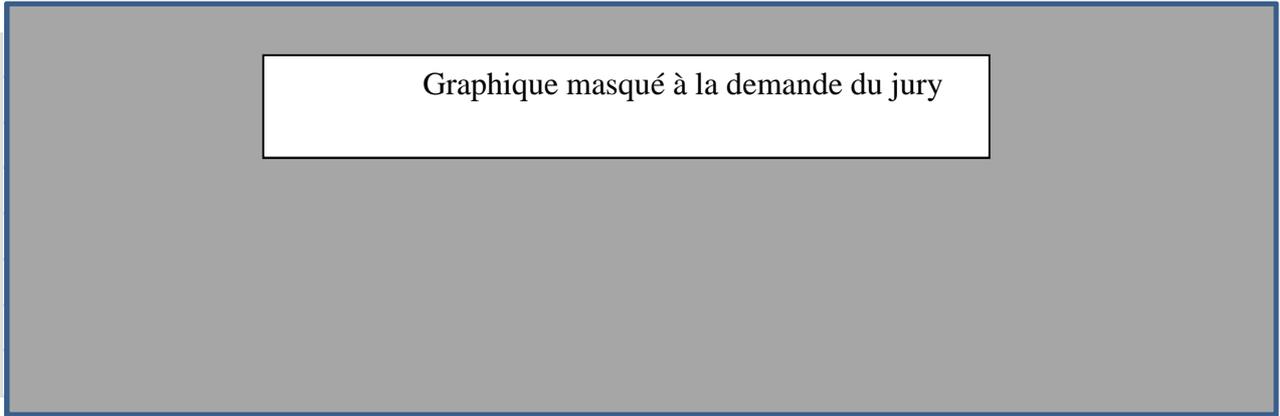
Ce score est calculé pour plusieurs critères, et à différents niveaux : magasin, région et Pays.

Les indicateurs sont calculés à partir d’un ensemble de critères opérationnels, les définitions de ces critères dans le questionnaire peuvent être résumées par :

Tableau 7: Items de satisfaction baromètre B2



Graphique masqué à la demande du jury



Nous disposons donc de 96 magasins répartis en France. Leur répartition géographique est visible sur la carte suivante :

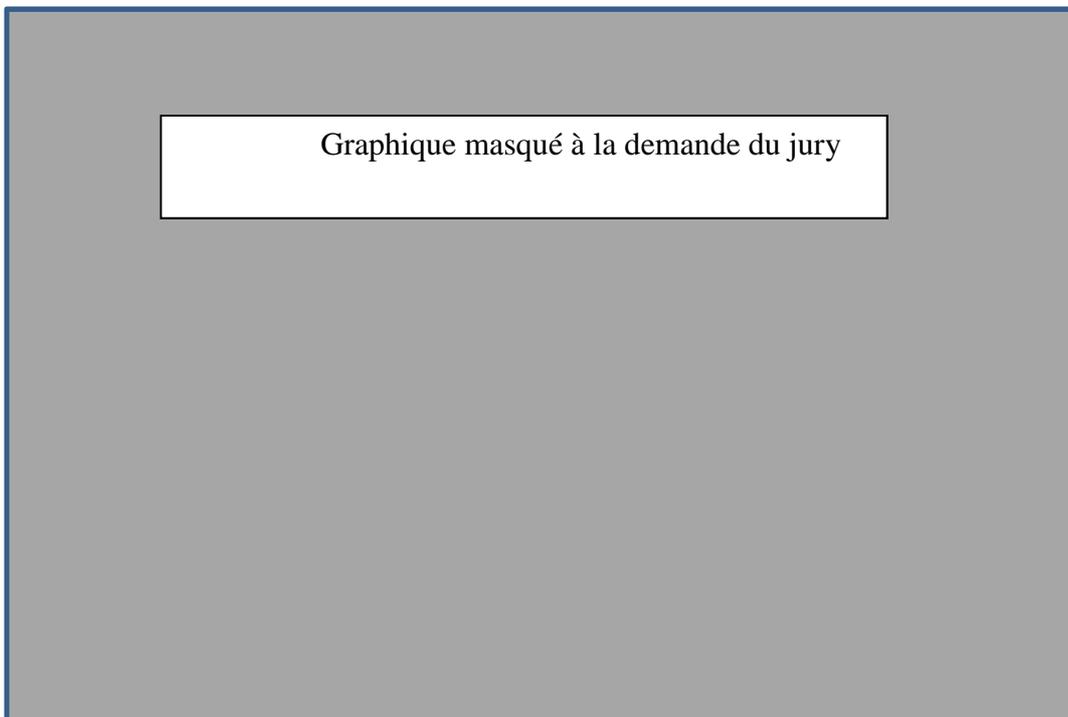


Figure 17: Répartition géographique des magasins du baromètre

Les données disponibles sont des informations agrégées, concernant des indicateurs ici c'est des scores de satisfaction, enregistrés sur 16 vagues (4 trimestres sur 4 années). Ces données sont issues d'un baromètre de suivis qui par sa nature, est destinée aux managers. Les résultats servent les décideurs à contrôler la satisfaction clients de leurs magasins dans le temps. Comparer les régions entre elles et enfin rendre compte des évolutions au niveau national de leur satisfaction clients. Une des interrogations qui émergent des analyses descriptives de ce baromètre est le manque de dynamique des

indicateurs par vague comme le montre le graphique figure 18. Cet instrument est donc un outil adapté pour juger de la sensibilité des indicateurs. Même si a priori, il semble présenter une stabilité, le fait de tester la sensibilité permettra de révéler si elle est contingente⁸³ ou systématique.

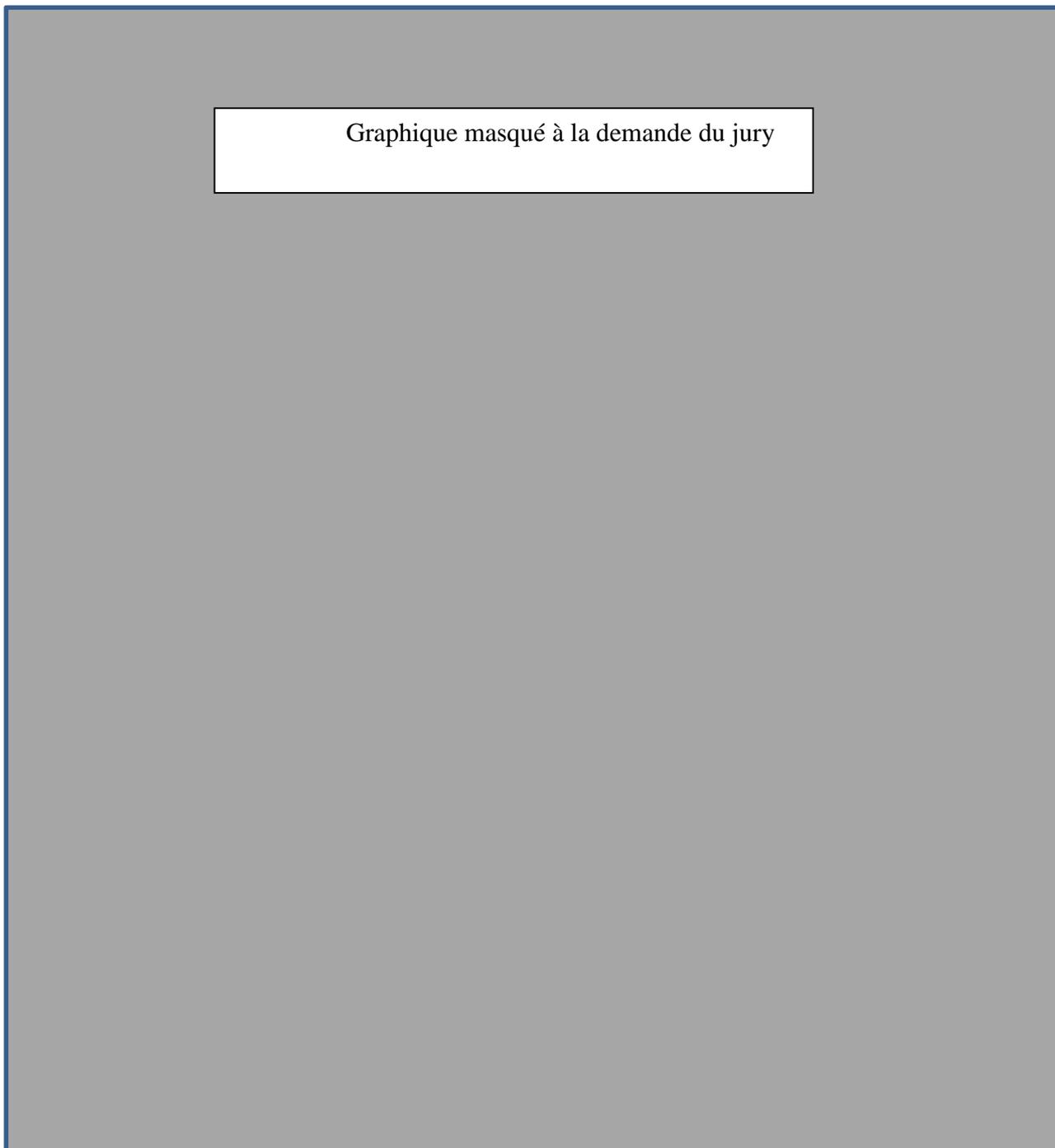
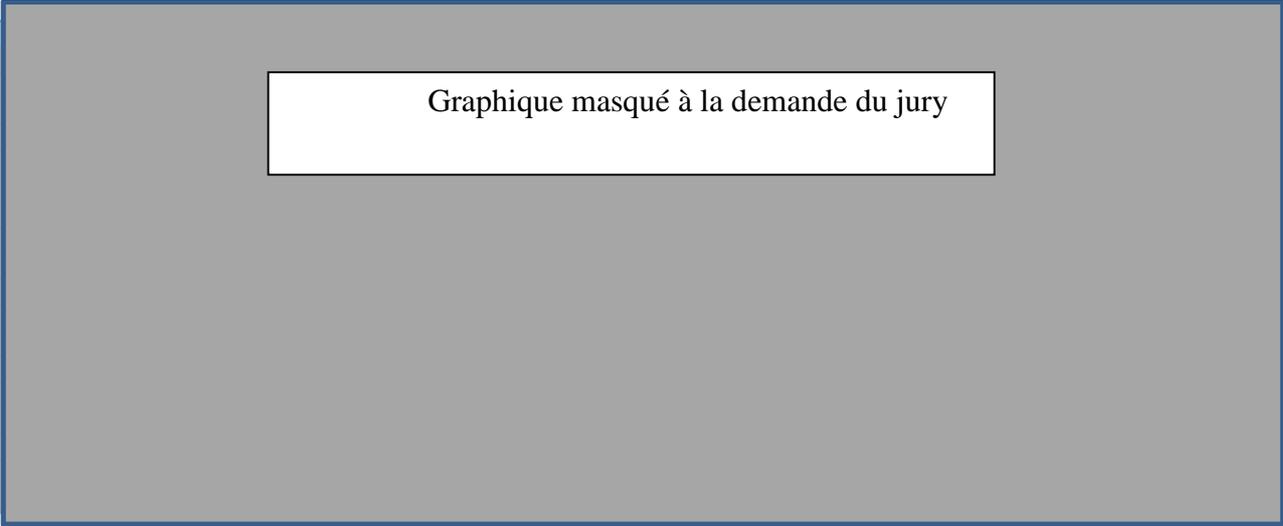


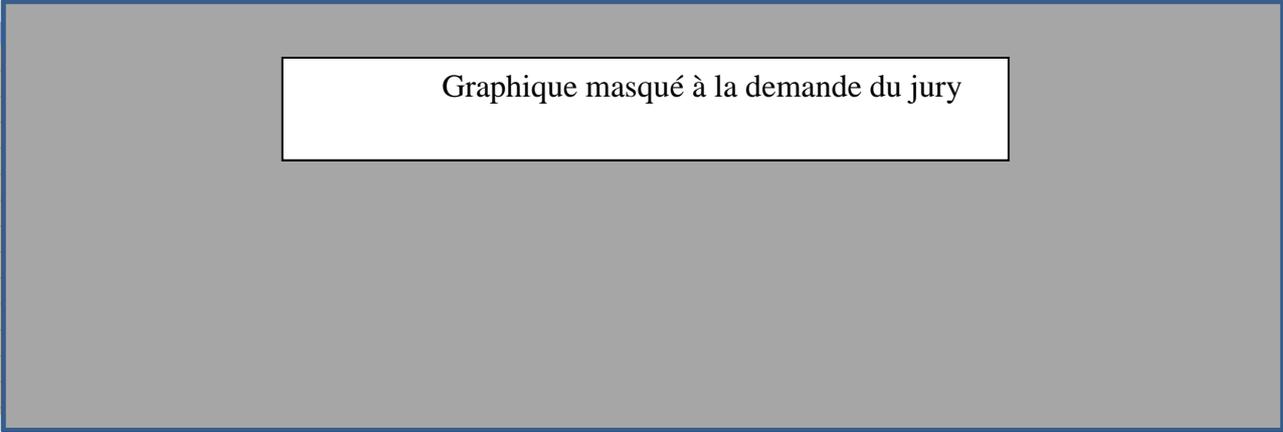
Figure 18: Dynamiques des items du baromètre B2

⁸³ Par contingente nous pensons au niveau de l'analyse. Est-ce que la stabilité est apparente quelques soit le niveau de l'analyse : magasin, région ou Pays, ou cette stabilité n'a pas la même ampleur suivant ces niveaux.



Graphique masqué à la demande du jury

Figure 19 : Extrait du Baromètre B1



Graphique masqué à la demande du jury

Figure 20: Extrait du baromètre B2

Ainsi les données que nous utiliserons pour cette recherche sont issues des enquêtes administrées par des entreprises à leurs clients. Il reflète l'état d'un business réel mettant en relation une entreprise et ses clients. L'avantage majeur de ces données reste celui du nombre de questionnaires disponibles. Ce fait nous a permis d'être confiants dans les analyses statistiques utilisées.

Toutefois, comme cela a été défini précédemment, les baromètres que nous traiterons sont structurés différemment. Le premier que nous avons noté B1 avait pour objectif principal d'établir un diagnostic de ce qui est la satisfaction globale par magasin. Suivant un ensemble de critères explicatifs, il permettra d'une part d'agréger les données via des indices (situés au niveau du magasin) et d'autre part de fournir des informations sur les facteurs de contingence nécessaires pour tester les propositions.

Pour le baromètre que nous avons noté B2, il servira à tester la sensibilité des indicateurs dans le temps, le tout en permettant de considérer divers niveaux d'analyse : niveau national, région ou encore par magasin. Cette structure différente par nature, qui sert à enregistrer et suivre un ensemble de levier de satisfaction dans le temps, est totalement adaptée à notre propos, tester la sensibilité dans le temps des indicateurs.

En guise de conclusion de cette section, nous souhaitons rappeler que comme la démarche de cette recherche ne concerne pas l'individu mais plutôt les indicateurs de satisfaction, ces outils s'adaptent bien à notre problématique, même s'il existe une faiblesse psychométrique⁸⁴. Comme le nombre de données dans ces baromètres est important les erreurs aléatoires seront fortement diminuées, ce qui est important dans notre démarche car on s'intéresse à un agrégat⁸⁵.

Après avoir présenté les données, nous proposons dans la suite d'expliquer la démarche statistique utilisée. Cette procédure est distinguée en deux processus qui sont la méta-analyse et l'analyse fiducio-béyésienne. Le rapprochement de ces deux méthodologies, nous a permis de construire un outil capable de tester les propositions de test des indicateurs. Nous commencerons notre spécification par définir le périmètre de chaque technique, l'origine, l'usage et processus de mise en œuvre, puis nous présenterons la démarche finale de test statistique.

Section 2 – **La méta-analyse une méthode de gain en puissance statistique**

La méta-analyse est une méthode de synthèse qui se veut être la plus objective possible, et ne reflétant que les faits. Elle présente un avantage majeur, celui de quantifier l'effet de la variable d'intérêt. Il s'agit de méthodes qui s'intéressent directement aux résultats des études pour inférer les conclusions. Ce type de procédure consiste à conjuguer un ensemble d'informations relatives à plusieurs sources⁸⁶, pour démontrer l'existence de l'effet, tout en gagnant de la puissance statistique. La méta-analyse revient à effectuer des analyses sur des informations agrégées, des résultats d'études, pour

⁸⁴ Suivant la considération psychométrique de la mesure. Ce point sera développé dans la dernière partie de cette thèse.

⁸⁵ Lorsqu'on s'intéresse à une résultante agrégée, la fiabilité se traduit par une bonne estimation à faible variance. Ainsi le nombre de données réduit les variances d'estimation et augmente la confiance dans les résultats.

⁸⁶ On pourra penser à plusieurs études traitant le même sujet dont l'intérêt est de les combiner pour statuer sur la pertinence du problème traité.

conclure sur l'existence ou non de l'effet de la variable qui est étudiée indépendamment dans chaque étude. A titre d'exemple, nous pourrions penser que pour commercialiser un médicament, différentes études sont effectuées (sur plusieurs groupes de sujets). Pour conclure si le médicament est efficace, il faudra alors conjuguer l'information de ce panel d'études pour pouvoir conclure. Dans ce cas, la méta-analyse constitue ici une approche d'analyse.

Le terme de méta-analyse est apparu dans les années 70. Cette technique qui reste très utilisée en médecine, s'est tout d'abord développée dans le domaine de la science sociale et de l'éducation. Les premiers travaux menés avaient pour souci d'augmenter la puissance statistique lorsque le volume de données était réduit. Le problème d'échantillon était surmonté en combinant plusieurs panels qui traitent le même thème. Sous des conditions d'homogénéité⁸⁷, les résultats obtenus, étaient plus puissants.

Le principe d'homogénéité, une des hypothèses de la méta-analyse, se traduit par le fait que l'ensemble des études (échantillons) peuvent être considérées comme similaires et donc leur regroupement sera licite. D'autre part, la question de la puissance statistique est surtout liée aux tests effectués afin de statuer sur l'effet de la variable d'intérêt. En général, un test statistique classique, met en jeu une hypothèse dite « nulle », absence de l'effet, contre son contraire appelée hypothèse « alternative ». Sous cette forme nous considérons que l'hypothèse « nulle » absence d'effet, est vraie et nous confrontons les données observées à cette hypothèse. Si les données sont incompatibles avec l'hypothèse « nulle » on rejette cette dernière, et si les données sont compatibles on ne rejette pas l'hypothèse nulle. Pour arriver à ce test on quantifie la probabilité de rejeter à tort l'hypothèse « nulle » qui correspond à un risque de tirer une conclusion incorrecte, supposer que les données soient incompatibles avec l'hypothèse nulle, quand celle-ci est vraie. Si cette probabilité appelée (*Pvalue*) est plus faible qu'un risque fixé α , on rejette, l'hypothèse « nulle » pour « l'alternative ». Sinon on ne rejette pas l'hypothèse « nulle ».

Cependant il existe un autre risque β (deuxième espèce) qui accompagne le premier α . Ce risque est celui d'accepter à tort l'hypothèse « nulle ». Il correspond au cas où on ne conclut pas à un effet quand il existe vraiment, supposer que les données soient incompatibles avec l'hypothèse « alternative » quand celle-ci est vraie. Ce risque peu considéré dans la pratique, est un indicateur de la puissance du test. En effet cette

⁸⁷ L'homogénéité consiste à dire que dans chaque étude le même effet est présent. La variation qui peut exister entre chaque résultat de chaque étude prise indépendamment, est juste due au hasard.

puissance se définit comme la probabilité de refuser l'hypothèse « nulle » à raison, c'est-à-dire la probabilité que le test détecte une différence (présence de l'effet) quand cette dernière est présente. C'est le pouvoir discriminant du test. Cette puissance est utile lorsque le résultat du test d'hypothèse est non significatif (l'hypothèse nulle non rejetée) est que l'on veut donner une confiance aux résultats. Par exemple si l'on est en présence d'un test de puissance valant 95%, et que ce dernier conclut que l'effet est non significatif (l'hypothèse nulle non rejetée), on pourra dire, avec une confiance de 95%, que l'effet observé n'est pas significatif. Cependant quand la puissance est de 20%, il faudra se garder de pareille conclusion.

La puissance du test permet de dire si les conclusions sont fiables. Donc eu égard à cela, le fait d'assembler plusieurs sources permet de faire un bon choix entre les deux hypothèses, c'est pourquoi gagner en puissance statistique est synonyme de « fiabilités ».

En général les méta-analyses sont appliquées pour des études contenant des résultats de comparaison entre la présence versus l'absence de l'effet d'une variable d'intérêt. L'expérimentation est menée en comparant deux groupes de population sur lesquels l'effet est testé. La conjugaison d'études dans ce cas est effectuée de telle sorte que seul l'effet est pris en compte. Cela est possible car la méta-analyse travaille directement sur les résultats et non sur les individus avec des indicateurs indépendants du nombre des données⁸⁸. C'est sur cette voie que notre méthodologie s'est axée.

Le schéma de la méta-analyse permet de supposer que l'information apportée est décomposée en une partie commune à toutes les études et une partie spécifique à chacune d'elles. En d'autre terme, nous supposons que l'effet de la variable d'intérêt (facteur) est présent dans tous les résultats et que les variations constatées sont causées par l'aléatoire. A titre d'exemple, les médecins ont souvent recours à ce concept pour conclure sur l'effet d'un médicament, car dans la pratique médicale, ils sont contraints de procéder à plusieurs tests avant de conclure. Par conséquent, ils se basent sur la synthèse de l'information via les méta-analyses pour affirmer l'existence de l'effet traitement, en comparant deux groupes de patients : les traités et les contrôlés. Puis à l'issue de cette comparaison une quantification de la taille de l'effet est obtenue et une conclusion est

⁸⁸ Cette notion renvoie aux statistiques descriptives qui seront détaillées plus tard dans ce texte. Une statistique est indépendante des données, si elle apporte la même information quelque soit le volume de données disponible pour l'évaluer.

tirée. En résumé la méta-analyse est employée comme un outil de synthèse sur un ensemble d'essais traitant un même thème, afin de mieux contrôler le problème.

I – La méta-analyse dans le marketing

. Pour le marketing, les problématiques étudiées sont diverses, néanmoins les structures de ces questionnements sont des paradigmes causaux qui cherchent à définir l'existence des liens entre les construits. A titre d'exemple, l'utilisation des méta-analyses dans le marketing a souvent été employée dans la publicité, le comportement des consommateurs et les études des performances des marchés. A cette occasion, plusieurs études ont été conjuguées puis analysées pour prouver si les corrélations constatées sont bien consistantes. Parmi les premiers articles qui utilisent les méta-analyses pour la problématique du comportement des consommateurs, on trouve l'article de Peterson and al (1986) dans lequel, les auteurs ont présenté la notion de l'importance de l'effet étudié. Dans cette recherche, ils ont rassemblé plusieurs études en travaillant sur les amplitudes des liaisons constatées. Nous trouvons aussi dès 1989, des articles comme celui de Sheppard and al,(1989), dans lequel la théorie de l'action raisonnée de Fishben est traitée. Dans cet exemple, Les méta-analyses ont été entreprises pour tester l'efficacité du paradigme.

Il existe d'autres thèmes dans lesquels la méta-analyse a été utilisée. Parmi ces derniers il y'a : la relation clients (Brown .S and Stayman.D, 1992 ; Palmatier and al, 2006) et la satisfaction (D.Szymanski and D.Henard, 2001) ou encore la publicité subliminale et le comportement des consommateurs, avec l'article de Trappey.C (1997). Aussi on trouve d'autres méta-analyses traitant les relations avec les publicités comparative et non comparative (Grewal and al 1998), ou encore celles étudiant les effets des émotions négatives et positives sur les réponses de la publicité (Brown and (1998), et celles traitant l'effet de l'humour dans les publicités avec l'article Eisend (2009). Comme les autres, ces méta-analyses ont aussi servi comme méthodes de synthèse. A signaler aussi que la méta-analyse a trouvé des applications dans les problématiques des orientations des marchés et le lien entre la perception du prix et la performance des marchés Rodriguez, (2004). En somme, la méta-analyse a surtout été utilisée dans les domaines du marketing comme un outil de synthèse visant à gagner en puissance statistique, et comme un moyen de tester la robustesse des modèles étudiés.

Les débuts des méta-analyses comme outils pour le marketing remontent à Farley and Donald R. Lehmann (1986), avec leur livre : *Meta-analysis in marketing : generalization of response models*. Dans leur développement, les deux auteurs soulignent les avantages de la méta-analyse pour la recherche appliquée en marketing. Ils font valoir que la méta-analyse est un bon moyen de recherche d'estimations fiables des effets au sein des problématiques marketing et soulignent aussi que les résultats des méta-analyses peuvent influencer les études futures en fixant des hypothèses plus justes. Puis la littérature s'est enrichie par d'autres articles comme ceux de Beltramini, R. F (1987), Farley and al (1995), Fern, et Monroe (1996) et Frank G.R (2001) qui eux ont mis l'accent sur l'aspect méthodologique et ont exposé des procédés quantitatifs pour mener des méta-analyses dans le Marketing.

Pour le Marketing, le principe dans l'utilisation de la méta-analyse reste en vigueur, seule les problématiques changent. Dans cette optique les méta-analyses seront affectées pour étudier les structures causales entre les concepts étudiés (ex : satisfaction ->fidélité) résumés par des indices représentant ces structures. Comme cela a été précisé précédemment, la méta-analyse est effective seulement en présence d'un ensemble d'études traitant le même thème. Cette exigence nous laisse penser qu'elle est la cause de son manque de notoriété dans le domaine de la recherche en marketing. Cependant depuis quelques décennies l'abondance des données des études d'opinions a laissé un champ d'application et de développement des méta-analyses. Mais il reste à adapter ses méthodes aux traitements de ce type de données.

A ce jour, différentes méta-analyses ont été produites dans le Marketing, les sujets de leurs développements sont divers mais la méthodologie reste semblable. Nous attirons l'attention sur le fait que les techniques de calculs statistiques restent identiques à celles utilisées en médecine, et leurs adaptations pour les problématiques du marketing restent très rares.

Dans notre problématique, cet esprit est présent mais l'intérêt lui est axé sur l'explication de l'effet via d'autres sources de données agrégées représentant d'autres sources d'information. Le tout en cherchant à expliquer la variation sans se préoccuper de gagner en puissance statistique car elle est garantie par le nombre de données (dans notre cas plusieurs milliers d'individus). Dans notre recherche, la méta-analyse servira à justifier le fait qu'on puisse travailler sur des données à des niveaux supérieures (ex : le magasin) qui sont résumées par des statistiques descriptives.

En mettant de côté nos besoins actuels, nous tenons à signaler que les méta-analyses, dans leurs applications classiques, peuvent répondre facilement au besoin du domaine du marketing. Une des principales raisons qui nous encourage à penser que le développement des méta-analyses est inévitable est surtout liée à la multiplication des sources de données. Pour un sujet, on peut trouver des centaines d'articles, ou encore des milliers d'enquêtes issues du secteur professionnel. Alors devant ce volume d'informations, les méthodes de synthèse sont inévitables. De plus nous pensons que sa progression peut être envisagée pour solutionner les problèmes rencontrés durant les études. L'objet de l'utilisation des méta-analyses peut être distingué en deux cas :

A – Présence d'un facteur ou variable modératrice du phénomène

Devant un lien de causalité entre deux construits supposé non linéaire, il est commun de s'intéresser aux facteurs ou aux variables qui peuvent moduler la relation. Dans ce cas, nous chercherons leur effet en supposant qu'il est redondant dans chaque étude. Pour affirmer que cet effet est systématique, il devient nécessaire de faire une synthèse de toutes les données pour statuer et ainsi expliquer la variation des résultats

B – Présence de résultats discordants

Ce fait est fréquent dans le secteur. Il représente le cas où un ensemble d'études affirment un constat, mais d'autres avancent le contraire. Devant ce problème de divergence de résultats, les conclusions restent difficiles à tirer. Dans ce cas de figure, la méta-analyse peut servir à donner une vision synthétique, en combinant toutes les études afin d'expliquer au mieux les constats et cela en gagnant de la puissance statistique ou en identifiant des dimensions cachées (facteurs contingents). Alors de nouveau, la méta-analyse devient un bon remède et une piste à envisager pour cerner les causes des perturbations. En général, les méta-analyses sont utilisées pour mieux identifier les variances⁸⁹ afin de garantir la puissance statistique nécessaire pour mieux généraliser les constats.

II – Méthodologie d'une méta-analyse classique

Afin de mener une méta-analyse, il est nécessaire de passer par plusieurs étapes qui peuvent se distinguer comme suit :

⁸⁹ Cette décomposition permet de dégager toutes sources de variance indépendantes du phénomène étudié, nous pensons notamment à la variance d'échantillonnage.

- Comme toutes études statistiques, il faut bien définir ces objectifs (le phénomène à étudier, les critères de jugement choisis).
- Recherche de toutes les études qui concernent la problématique.
- S'il y a exclusion d'études, en justifier la cause.
- Estimer l'effet de la variable d'intérêt en choisissant le critère de jugement adapté, et la méthodologie statistique adéquate.
- Faire une lecture critique en confrontant d'autres études, pour s'assurer de la reproductibilité des résultats.

Ce processus permet d'assurer une convenable estimation et une bonne reproductibilité des résultats. De plus il assure une démarche qui justifie les résultats finaux.

Les méta-analyses présentent l'avantage de déboucher sur la notion de « quantité » de l'effet. Pour ce faire elles utilisent des outils statistiques qui permettent de donner des mesures ou des indicateurs de cette taille. Dans la pratique, les critères de jugement de la taille de l'effet à partir desquels les indicateurs sont calculés, peuvent être soit binaires soit continus.

Dans le cas binaire, les indicateurs peuvent être les **Odds ratio** ou le **risque relatif**⁹⁰. Ces derniers permettent d'estimer la taille de l'effet. Tandis que pour les critères continus, on a le coefficient *d* de **Cohen**, le **g de Hedges** présentés par Larry V.Hedges (1981) et Cohen, Jacob (1992), ou encore le coefficient de corrélation. Tous ces indices permettent de mesurer la force de l'effet observé, chaque indice est utilisé suivant la situation de l'étude.

Pour arriver au résultat final, la méta-analyse procède par trois paliers qui sont :

- a. l'estimation de l'effet commun : il s'agit de l'estimation de l'effet de la variable d'intérêt en combinant l'ensemble de l'information disponible. Cette estimation est effectuée avec l'hypothèse de l'existence d'un effet commun.
- b. test d'association : il s'agit d'un test statistique qui sert à vérifier si l'hypothèse de l'existence d'un effet commun à toutes les études est bien valable.

⁹⁰Le risque relatif est une statistique descriptive utilisée pour mesurer le risque de survenue d'un événement dans un groupe par rapport à un autre. L'Odds ratio, pour rapport de chance, permet de juger de degré de dépendance entre des variables aléatoires binaires. D'où son utilisation dans le cas binaire : présence versus absence.

- c. test d'hétérogénéité : c'est un test statistique qui permet d'affirmer ou pas, que toutes les études sont similaires; en d'autre terme ce test affirme que le regroupement des études est possible.

Pour plus de clarté, nous allons présenter la démarche à suivre pour combiner l'information de plusieurs études. Cette procédure est similaire quelques soit le critère de jugement utilisé (binaire ou continu).

Pour estimer l'effet commun, les critères de jugement, indicateur de la force de l'effet, sont calculés pour chaque étude : Des estimations de la taille de l'effet sont obtenues. Puis à l'issue de ces estimations, une approximation de la taille de l'effet commun est obtenue. Il existe deux façons de combiner l'information, une à travers un modèle à effets fixes et l'autre à partir d'un modèle à effets aléatoires.

Le modèle fixe suppose que la valeur de l'effet commun est pareille pour toutes les études. A partir de cette hypothèse, l'estimation finale de l'effet commun est obtenue en calculant une moyenne pondérée par l'inverse de la variance de chaque étude. Cela revient à donner un poids plus important aux résultats d'études à faibles variances et un poids moindre aux autres.

Le modèle aléatoire est utilisé quand l'hypothèse d'homogénéité n'est pas vérifiée (test d'hétérogénéité non significatif). C'est à dire lorsque l'effet peut être considéré comme différent dans au moins une étude⁹¹. Cette technique prend en compte l'hétérogénéité mais ne l'explique pas.

Le modèle aléatoire se distingue du fixe dans la façon d'estimer la taille de l'effet. En effet, Il décompose la variance pour chaque étude en deux parties, une spécifique et une autre résiduelle. Après estimation de la taille de l'effet de chaque étude, la même moyenne pondérée, est effectuée pour obtenir l'effet commun final.

A travers ces approches nous constatons que faire une méta-analyse revient à combiner de l'information en moyennant les scores de chaque étude. Cette procédure est justifiée par l'hypothèse de l'existence d'un effet commun et par le fait que l'écart constaté dans chaque étude, est dû à une erreur aléatoire. Dans ce cas, procéder en moyenne, permet d'avoir une estimation consistante de la vraie taille de l'effet.

⁹¹ C'est-à-dire qu'il existe au moins une étude ou l'effet se manifeste très différemment, par son intensité ou encore dans son sens.

III – Limite de la méta-analyse

Comme toutes méthodes empiriques, les méta-analyses présentent des limites. Les déficiences constatées peuvent être résumées ainsi :

- Pour procéder à une méta-analyse il est primordial de se constituer un ensemble d'études antérieures. Le premier obstacle posé sera évidemment de bien sélectionner les études sans introduire un biais dans la méta-analyse. Un bon choix, serait d'inclure toutes les études possibles. Comme cela est difficile à réaliser, il faudra être conscient qu'un biais est toujours présent. Ce biais pourrait être introduit par les études elles-mêmes, ou tout simplement en négligeant certaines. Il est donc essentiel de prendre les conclusions de la méta-analyse avec précaution. Il est rare que de tels biais soient pris en compte puis corrigés. Mais en effectuant une analyse de sensibilité on pourra statuer sur la validité de la méta-analyse. Cette analyse consistera à inclure ou à enlever d'autres études à la méta-analyse, puis de regarder la stabilité de ces résultats. Cette démarche pourra être fastidieuse et très coûteuse en temps.
- Une autre déficience de cette procédure est le cas où les résultats sont non concluants et que seule la méta-analyse conclut. Une des explications de ce souci est soit la présence de biais dans les études, soit un manque de puissance statistique dans chacune d'elles. Pour statuer il est nécessaire de confirmer ce constat par d'autres études ayant une puissance statistique suffisante. Néanmoins, il faut que cela soit possible.
- Une autre faiblesse de la méthode est liée au paradigme fréquentiste de la méta-analyse. Dans cette optique on se base sur le théorème central limite qui assure la validité asymptotique des résultats obtenus, en vue du biais de non exhaustivité des études retenues. Il serait logique de raisonner sachant les observations et ainsi ne refléter que les faits.
- Les méta-analyses sont incapables de regrouper les études contenant des variables hétérogènes. Dans ce cas on se retrouve avec des informations manquantes. Le choix serait d'exclure l'étude, mais pourquoi ne pas inclure ces études en prenant en compte les données manquantes. Les modèles hiérarchiques bayésien sont de bonnes alternatives à ce problème.

La méta-analyse regroupe dans son fonctionnement plusieurs techniques statistiques et permet ainsi d'être souple et générale. L'intérêt porté à cette méthode dans le domaine médical est grand. Ce caractère est lié aux types des problématiques dans ce domaine et aussi à l'abondance des études traitant le même thème, qui font que la méta-analyse est devenue une méthode courante. Le marketing quant à lui, peine à s'approprier cette technique et surtout de l'adapter à ces besoins. Même si, depuis des années, différents auteurs l'ont présentée et développée, elle reste un procédé mal connu. Les explications de cette carence sont surtout liées aux processus de la méthode qui ne sont pas applicables directement aux sujets du marketing, et au fait du manque des ressources similaires nécessaires pour l'utilisation de la méta-analyse.

Toutefois, nous pensons que ces deux problématiques peuvent être contournées en incluant d'autres sources que celles de la sphère académique. Par exemple en élargissant aux secteurs professionnels, cette alternative nous permettra de contourner le problème de manque de ressources et le souci du manque d'études similaires à conjuguer. Les travaux effectués dans le domaine professionnel contiennent la même information d'une vague à une autre, par conséquent, la méta-analyse sera plus facilement applicable. Ce fait sera exploité dans cette thèse ou les méta-analyses seront appliquées exclusivement pour les données issues de baromètres professionnels. Cependant cet élargissement n'est pas une solution à tous les problèmes, car ce changement doit être accompagné par une adaptation des techniques de la méta-analyse aux questions du marketing. C'est pourquoi, nous envisageons une adaptation de la méta-analyse en incorporant l'analyse de comparaison dans le paradigme bayésien qui permet de répondre à notre problématique de l'influence des facteurs de contingence en gardant l'esprit de la méta-analyse.

Les méta-analyses comme méthode de synthèse joueront un rôle important dans le marketing, car ce domaine se renouvelle très lentement quand il s'agit de techniques quantitatives. D'autant plus que l'abondance de l'information accentue le besoin de s'adapter, afin de contrôler ces problématiques. Ainsi suivant cette logique nous présenterons divers éléments nécessaires pour répondre aux problématiques de traitement de l'information dans le marketing et ainsi, peut-être un jour, faire de la méta-analyse une composante du marketing. Comme notre question de recherche vise à évaluer les indicateurs de satisfaction qui sont par essence des quantités relatives à des unités situées à des niveaux supra individus (ex: magasin), la méta-analyse sert à justifier la démarche d'agrégation des données via des indices indépendants des données. Ces indices qui seront

vus comme des résultats sur lesquels des analyses supplémentaires vont être effectuées pour comprendre leurs variations entre unités. Là où la méta-analyse classique sert à les conjuguer entre eux pour estimer la taille de l'effet commun, nous nous les étudierons comme des quantités statistiques pour expliquer leurs variations par des sources extérieures, dans notre cas les facteurs de contingences.

Section 3 – **Existence, importance et sens de l'influence d'un l'effet**

On dit qu'un facteur possède un effet lorsque celui-ci agit sur un phénomène de telle sorte que conditionnellement au facteur en question le phénomène varie. En statistique, cela se traduit par une différence à l'intérieur des modalités du facteur, les valeurs par modalités ne sont pas toutes identiques, par rapport à l'absence de ces derniers. D'un autre côté, l'absence d'effet se traduit quant à elle, par une constance dans les valeurs conditionnellement au facteur. Ce qui revient à dire que les facteurs n'influencent pas le phénomène étudié. Souvent la notation « facteur » est associé à des variables qui créent des groupes d'observations. Ainsi l'effet du facteur dans ce cas se traduit par une différence entre les moyennes de chaque groupe (modalités) et la moyenne générale calculée à partir des observations. Si cette différence existe le facteur est « significatif » sinon il sera considéré comme « non significatif ».

Les tests statistiques classiques sont construits pour détecter si une différence entre modalités du facteur existe, puis de statuer sur la significativité du facteur. En somme un facteur significatif est un facteur influent au sens large. Il peut être important, les différences entre les valeurs de ses modalités sont grandes ou négligeable, les différences entre ses modalités ne sont pas manifestes.

IV – Importance de l'effet

L'importance d'un effet est une question pratique. En effet, il est fréquent que le praticien cherche plutôt à inférer l'importance d'un effet plus que sa significativité (existence). La réponse à cette interrogation peut s'avérer utile lorsque l'on est en présence de différents facteurs, qu'on voudrait hiérarchiser selon leur importance. De plus, tout statisticien averti sait pertinemment qu'il n'existe pas de facteur (d'un point de vu statistique) sans effet dans le cas où les données sont conséquentes. Ce qui est différenciant, c'est la force (l'importance) de son influence. Ainsi caractériser

l'importance de l'effet associée à un facteur, est plus pertinente que sa significativité (effet non nul).

L'importance de l'effet pourra être quantifiée à partir de la sémantique du domaine. Une différence d'un point par rapport à l'absence de l'effet (la valeur zéro) est importante ou encore une différence de dix points le serait aussi ou tout autant. Tout dépendant du domaine étudié. L'importance sera donc conditionnée par rapport aux phénomènes étudiés et de l'exigence de l'expérimentateur. Cependant cette approche reste subjective et peut varier d'une personne à une autre. Sa généralisation est donc difficile.

Toutefois, Il existe une autre façon de caractériser l'importance de l'effet. Cette fois en se basant sur des critères généraux calculés seulement à partir des observations. Il s'agit de calculer des statistiques « des effets calibrés » qui sont internes aux données⁹². Ainsi à l'aide de cette approche de calibrage de l'effet nous pourrions comparer les résultats à l'aide des seuils préétablis.

En suivant cette optique, statuer sur l'importance d'un effet consiste à dire si ce dernier est notable, moyen ou encore faible. Pour répondre à cette question souvent difficile dans la pratique, des valeurs repères qui déterminent la frontière entre notable, moyen, ou faible, ont été présentées par (Cohen 1977), puis par (Corroyer, Rouanet 1994). Ces valeurs repères toutes accompagnées par un calcul de l'importance de l'effet permettront de spécifier à chaque fois les frontières entre un effet notable moyen ou faible. Corroyer et Rouanet soulignent la démarche descriptive visant à extraire l'importance de l'effet dès le recueil des données, à partir d'un calibrage de l'effet en question en le rapportant à un autre effet, ou à une référence interne aux données (ex : variance résiduelle). Comme l'importance d'un effet doit être la même quelques soit le nombre d'observations servant à le déterminer, les statistiques descriptives deviennent des indicateurs privilégiés pour calculer l'importance de l'effet lors du calibrage de ce dernier.

V –Statistique descriptive

Une statistique est un indicateur servant à résumer une information contenue dans des données, comme par exemple une variance, somme des carrés etc. Cependant une statistique prend le statut d'une statistique descriptive des lors qu'elle ne dépend pas **du**

⁹² Le **D de Cohen** en fait partie.

nombre d'observations (nombre de sujets). Il s'agit d'une quantité relative à la structure des données et non au nombre. Une statistique descriptive apportera donc une information neutre, indépendante du nombre des données utilisées, mais qui reste représentative du phénomène étudié. L'utilisation de ces quantités permettra de procéder à ce qu'on appelle l'étape descriptive, qui est différente de celle de l'induction qui introduit quant à elle, des éléments dépendants du nombre de sujets pour exprimer le potentiel inductif des données.

L'étape descriptive permet de répondre à la question : Que disent les données ? L'étape inductive permet de répondre à la question de généralisation des résultats observés?

Les statistiques descriptives que nous considérons au cours de ce travail sont toutes construites à partir de la variance brute ou l'écart type brut, définies selon (Bernard, 1994). En s'inspirant des recommandations de cet auteur et en utilisant ces définitions des statistiques descriptives nous construirons, à chaque fois, un indicateur de l'importance de l'effet basé sur les statistiques descriptives ce qui permettra de s'affranchir de l'influence des données et de pouvoir comparer des indicateurs construits sur des échantillons de différentes tailles.

VI – Indicateurs de l'importance des effets et valeurs repères

Nous définirons dans cette partie divers indicateurs de l'importance des effets suivant les cas ci-après: comparaison de deux moyennes (Anova sur un facteur à deux modalités), comparaison de plusieurs moyennes (Anova sur un facteur à plusieurs modalités) ou encore la liaison entre deux variables numériques (la corrélation statistique). Les indicateurs de l'importance des effets qui seront étudiés dans cette partie sont tous définis dans l'article de Corroyer et Rouanet (1994). Une recherche dans laquelle l'importance est caractérisée par des indicateurs psychométriques « des indicateurs calibrés » dont les valeurs repères pour statuer sur leur importance, sont spécifiées. Cette étape consiste donc à décrire (la phase descriptive) l'effet d'un facteur par son importance sans passer à l'étape inductive (test statistique).

A – Importance pour une comparaison de deux moyennes

Ce cas de figure est effectif lorsque nous sommes en présence de deux groupes d'observations, données conditionnées par un facteur à deux modalités : on voudrait savoir si la différence constatée est importante ou pas. Pour ce faire on considère la

différence entre $m1$ (moyenne de la modalité 1) et $m2$ (moyenne de la modalité 2). Cet écart apparaît d'autant plus important si la dispersion est faible. Ce pourquoi la différence ($m1-m2$) est rapportée à l'écart type **intra-groupes** (statistique descriptive), ce qui permettra d'exprimer la différence en terme d'écart type : 1 fois l'écart type, 2 deux fois l'écart type, etc. Cette différence est donc « calibrée » en réduisant la différence ($m1-m2$), par l'écart type. Nous noterons **Ec** l'indicateur associé qui est défini par la formule :

$$EC = \frac{(m1 - m2)}{\sigma}$$

* $m1$ et $m2$ sont les moyennes par groupes, et σ l'écart type intra-groupes.

Les valeurs repères pour qualifier si un effet est notable ou faible par le calcul de L'Ec sont les suivantes :

Effet faible	Effet moyen	Effet important
0,2	0,5	0,8

B – Importance pour une comparaison de plusieurs moyennes

Dans une analyse de variance les conclusions sont prises suivant le test dit **F de Fisher**. Ce test est basé sur le rapport entre le carré moyen lié **au facteur** (quantité expliquée par le facteur « source systématique ») et le carré moyen intra groupe (quantité résiduelle « sources adjointes »). Plus cette quantité est grande, plus le facteur est significatif (effet non nul). La quantification de l'importance de l'effet quant à elle passe par le même principe, mais au lieu de considérer les carrés moyens, elle prendra le rapport des variances.

Ainsi pour la suite donc, nous définirons l'importance d'un effet comme un effet calibre (\hat{c}) en rapportant la variance inter (source systématique liée aux facteurs « eff ») sur la variance intra-groupes (source adjointe « *adj*»). Ainsi pour chaque effet, nous calculerons l'effet calibré, en rapportant sa variance inter (variance des moyennes des modalités), à la variance intra-groupe du modèle considéré (variance résiduelle). Cette source adjointe sera choisie à partir d'un résultat d'une analyse de la variance sur le facteur et sera souvent la part résiduelle non expliquée (variance intra-groupes).

➤ **Lien avec le F observe de Fisher-Senodor.**

L'effet calibre δ se déduit directement du **F** de Fischer à travers la formule suivante :

$$\delta = \frac{df_{eff}}{df_{adj}} \times F$$

Où df signifie le degré de liberté de sources considérées.

Ainsi nous obtenons la valeur de l'effet calibre directement du résultat du tableau de la variance en appliquant cette formule. De plus, nous pouvons prendre le paramètre η le rapport des corrélations classiques de Pearson, définit par :

$$\eta = \frac{\delta}{1 + \delta}$$

Cette quantité varie entre 0 et 1 et se traduit par la proportion de variance prise en compte par le facteur, ce qui peut être utile lors de l'interprétation. Pour ces deux indicateurs nous choisirons comme valeurs repères celles définies par Corroyer et Rouanet (1994) :

Pour δ

Effet faible	Effet moyen	Effet important
0,01	0,06	0,16

Pour η

Effet faible	Effet moyen	Effet important
0,01	0,06	0,14

Dans la suite nous avons considéré l'indice η , le rapport des corrélations classiques de Pearson, pour juger de l'importance de l'effet.

C – Importance liaison entre deux variables numériques

Dans le cas d'une variable numérique, le coefficient de corrélation linéaire est un bon représentant de l'importance de l'effet ; mais les deux indices δ et η sont aussi valables. Comme pour l'analyse de variance, il est possible d'obtenir une table résumant les variances expliquées (liée à la variable) et la variance résiduelle. Donc les deux indices peuvent être obtenus aisément à partir de test de Fisher associé à une régression linéaire. Pour un souci d'harmonisation des analyses, nous n'utiliserons l'indice η que pour qualifier l'importance de l'effet que cela soit pour un facteur ou une variable continue.

L'importance de l'effet est une notion incontournable pour le traitement statistique des effets. D'autant plus quand le sujet de la recherche vise à établir l'influence d'un facteur ou d'une variable. Un effet important peut être non significatif, ce cas de figure est souvent présent quand les données sont rares et quand les tests possèdent de faibles puissances statistiques. Ainsi le seul critère qui peut être intéressant est celui de l'importance, car majoritairement les recherches visent à établir une importance et non la non significativité. Ce pourquoi le calcul de l'importance de l'effet doit être effectué dès l'étape descriptive comme le préconise Corroyer et Rouanet (1994). Reste maintenant à dépasser le cadre descriptif et à établir cette fois-ci un test statistique sur l'importance de l'effet.

La non-significativité de l'effet, reflète une absence de conclusion. Dans ce cas, l'importance de l'effet, dire que ce dernier est notable, ne peut pas être prolongée inductivement car les données disponibles ne le permettent pas. Toutefois le calcul de la taille de l'effet (point de vu descriptif) reste toujours envisageable.

Par conséquent dans la suite, les procédures inférentielles sur l'importance de l'effet ne seront valables que dans le cas où le test est significatif à un seuil convenable (5%). Lorsque le test est non significatif seule la portée descriptive sera donc considérée et aucune procédure inductive⁹³ ne sera envisagée.

Section 4 – L'analyse de comparaison bayésienne

La notion de comparaison est une procédure naturelle dans un environnement où le besoin de contraste entre deux considérations est présent. Dans le quotidien de

⁹³ Par inductive nous volons signifier le fait de la généralisation statistique : conclusion valable au niveau de la population

l'expérimentateur cette question est étroitement liée à ses objectifs. Le but de chaque étude, dans un cadre statistique, est d'établir les différences (ou égalités) entre les multitudes de concepts étudiés. Nous sommes habitués à utiliser les statistiques pour statuer sur ce genre de problématique. Le calcul statistique fait suite à des questions posées *a priori* qui peuvent varier du simple fait de vouloir connaître si deux groupes d'individus (en termes de moyennes ou de fréquences) sont différents, jusqu'à vouloir savoir quel est le degré de différence entre plusieurs groupes. La technique statistique adaptée à ce type de comparaison est bien sûr l'analyse de variance de Fisher. Cette technique est bien connue des utilisateurs, et permet à tout praticien de statuer sur l'effet d'un facteur (différence entre plusieurs groupes d'individus) via le test F qui donne soit un résultat « significatif » (l'effet du facteur est considéré non nul) soit un résultat « non significatif » synonyme de l'ignorance. Cependant il est rare que la valeur de ce test comble les attentes d'un praticien quand sa question vise à savoir si l'effet de son facteur est important, car la vision classique des tests statistiques, reste largement insuffisante pour répondre à la question de **l'importance**.

C'est dans cette vision de refonte et d'amélioration des procédures d'inférences statistiques que l'analyse de comparaison bayésienne a été développée par Rouanet & Lépine (1975) puis par Lecoutre (1983,1984). Elle se présente sous forme de *formalisation algébrique* qui a permis de structurer les données sous forme de *plans d'analyses* dits spécifiques. Ces plans servent à planifier les données pour procéder à une *inférence spécifique* généralisant les résultats de la comparaison. L'analyse de comparaison bayésienne procède ainsi à une comparaison des données suivant un plan caractérisant l'effet étudié⁹⁴. Ce plan permet une inférence relative à l'effet considéré. Une comparaison est une question qui sera posée aux données (ex : quel est la différence entre deux moyennes ?) ; et l'analyse de fiducio-bayésienne est la procédure de test statistique pour juger de la validité de cette question.

En somme l'analyse de comparaison bayésienne grâce à ses techniques, permet de traiter toutes sortes de questions pertinentes relatives à l'analyse de la variance, notamment en identifiant l'importance des effets. L'analyse fiducio-bayésienne permet de procéder à l'inférence sur cette importance des effets en la généralisant au niveau de la population étudiée.

⁹⁴ On pourra imaginer que lorsqu'on voudra comparer deux groupes (jeune et sénior) seuls ces deux groupes seront considérés, d'où la plan de l'analyse spécifique.

I – Inférence fiducio-bayésienne pour l'étape inductive sur l'importance de l'effet

L'inférence fiducio-bayésienne, est une technique qui prolonge l'inférence fiduciaire développée dans les années trente par Fisher, sous le principe du « pivot ». Ce principe permet d'exprimer l'apport des données par un raisonnement intuitif qui vise à combler l'insuffisance de l'inférence classique. Cette méthode consiste à inverser le raisonnement de test classique. Au lieu de fixer une hypothèse particulière dont on cherche la compatibilité avec les données, on partira directement des données vers un ensemble d'hypothèse raisonnables, auquel on associe une certaine confiance (d'où l'adjectif fiduciaire).

L'inférence fiducio-bayésienne permet dans la pratique, de répondre à certaines hypothèses qui s'amènent des données en donnant une confiance (une probabilité associée à l'effet observé). Elle se distingue de l'approche classique par le fait qu'elle peut, par son raisonnement, tester l'importance de l'effet notamment en accordant une confiance⁹⁵. Ses techniques se situent comme un prolongement des méthodes d'inférence usuelles et permettent ainsi une réinterprétation de ces dernières, seuil de signification et région de confiance. Pour un exposé détaillé de cette démarche nous invitons les lecteurs à se référer à l'article de Rouanet, Lecoutre, Bernard, 1986). L'extension bayésienne du principe fiduciaire, présentée par Lecoutre (1983), apporte la justification théorique pour aborder les analyses de comparaisons (généralisation de l'analyse de la variance), connue sous le nom d'analyse de comparaisons bayésiennes.

D'autre part, Il est clair que plus le volume des données est grand plus on a de chance de conclure sur un effet non nul (significatif), même si la différence est très petite. Ce fait vient du test de **F de Fisher**. Par exemple ce test, est conçu pour détecter n'importe quelle différence, pourvu qu'on ait assez de données. Ce qui veut dire que si nous récoltons un grand volume de données nous aurons tendance à trancher pour un effet non nul, sans autant savoir s'il est important (une différence notable). De plus, il est assez rare qu'un facteur ait un effet non nul en réalité, mais il peut être notable ou non, d'où la nécessité de prendre en compte l'aspect de l'importance de l'effet et de ne pas se contenter des tests d'hypothèse. Voilà ce pourquoi l'analyse fiducio-bayésienne a été développée.

⁹⁵ La confiance ici se traduit par une loi de probabilité représentant la taille de l'effet. Cette loi permet d'attribuer un niveau de confiance sur la taille de l'effet observé. Pa exemple, nous pouvons chiffrer la probabilité que la taille de l'effet en question soit supérieur un seuil donné.

Dans leur raisonnement, les auteurs ont mis en place une panoplie d'outils issue de l'inférence bayésienne non informatifs afin d'arriver à définir des probabilités *a posteriori* sur les paramètres (ici **EC**, **η**) en leur attribuant des lois sur lesquelles l'inférence sera conduite. Par exemple on pourra chercher la probabilité que le paramètre **EC** soit supérieur à 0.5, ce qui nous renseignera sur la généralisation ou pas de la notabilité de l'effet dans le cas positif $EC > 0$. Bien sûr, cette méthode inductive complétera l'analyse descriptive faite au préalable via l'analyse de comparaison. Les lois obtenues au final sont construites de telle sorte qu'elles soient les plus objectives possibles (a priori non informatif de Jeffreys (1961)) en dégageant seulement l'information contenue dans les données. Pour reprendre une phrase des auteurs : « l'inférence fiducio-bayésienne est bayésienne par sa technique et fiduciaire par sa motivation ». Cependant grâce aux mêmes lois fiducio-bayésiennes, il est possible d'inclure une information déjà connue avant l'observation, d'autre expérience du même thème, et de conjuguer avec les informations disponibles pour une loi finale fiducio-bayésienne⁹⁶.

En somme, l'inférence fiducio-bayésienne permet de réinterpréter les tests classiques tout en ajoutant des éléments mathématiques pour mieux contrôler l'étape inductive, et répondre à la question de l'importance de l'effet. Voir Rouanet et Lecoutre (1981)

➤ Distribution fiducio-bayésienne pour l'importance de l'effet EC, η parent

Sous les mêmes hypothèses qui régissent l'analyse de variance, nous obtenons des lois dites fiducio-bayésienne correspondant aux effets parents de chaque paramètre **EC**, **η**. (Lecoutre 1983)

-la loi de **EC parent**, suit une loi de L' avec les paramètres suivants :

$$EC \propto L'_q \left(EC_{obs}, \left(\frac{EC_{obs}}{t_{obs}} \right)^2 \right)$$

Où EC_{obs} l'effet observé, t_{obs} est la valeur observée de la statistique de Student pour un test de différence de moyennes et q le degré de liberté de la variance résiduelle.

De plus, on peut considérer quand les données sont conséquentes, que l'indicateur **EC** de l'importance de l'effet, suit une loi normale :

⁹⁶ Comme nous le verrons dans la suite, la capacité d'inclure des informations supplémentaires s'adapte bien à la méta-analyse classique : conjugaison de plusieurs sources pour établir les faits.

$$\mathbf{EC} \propto N_q \left(\mathbf{EC}_{obs}, \left(\frac{\mathbf{EC}_{obs}}{t_{obs}} \right)^2 \right)$$

-la loi de δ **Parent** suit une loi de L^2 avec les paramètres suivants :

$$\delta \propto \frac{1}{q} \times L_{m,q}^2 (m \times F_{obs})$$

Où F_{obs} la valeur du test de Fisher, q est le degré de liberté de la source adjointe (variance résiduelle (n-m)) et m est le degré de liberté de la source systématique liée aux facteurs (variance expliquée par le facteur).

De même quand les données (n) sont abondantes (q est très grand), l'indicateur δ suit une loi de **Chi2 non centrée** avec les paramètres suivants :

$$\delta \propto \frac{1}{q} \times \chi_m^2 (m \times F_{obs})$$

-Pour l'inférence sur le coefficient des corrélations **η parent**, nous appliquerons, après estimation des valeurs de δ , la formule qui définit **η** :

$$\eta = \frac{\delta}{1 + \delta}$$

Nous remarquons que la procédure d'inférence est directement attachée à la procédure descriptive. Après avoir calculé les indicateurs de l'importance des effets, on applique la loi propre à chacun pour généraliser les conclusions descriptives. Il est clair que ces lois de probabilité vont nous permettre de quantifier l'importance en donnant des probabilités de dépassement d'un seuil de notabilité : valeurs repères pour juger de l'importance (ex : $\mathbf{Ec} > 0,5$). De plus dans le cas où les données sont abondantes (n très grand) la manipulation reste aisée, étant donné que toutes les lois obtenues sont connues. Cependant nous utiliserons ces lois pour effectuer de l'inférence lorsque le test associé est significatif car c'est dans ce cas qu'on peut prétendre donner une confiance sur l'importance de l'effet calculé au préalable par l'étape descriptive.

Dans le cas de l'analyse de variance, un test non-significatif à un seuil de 5%, est synonyme, dans la vision fiducio-bayésienne, que l'effet théorique cherché n'est pas proche de l'effet observé, donc la distribution est diffuse du fait de la non-significativité. De plus un test non significatif (absence de conclusion sur l'effet) est transposable naturellement aux indices η et δ (les deux indices sont des opérations arithmétiques de la

statistique F) donc faire de l'induction fiducio-bayésienne sur ces deux indices n'a pas vraiment de sens, dans le cas de la non significativité du test F.

En somme nous effectuerons de l'inférence sur l'importance de l'effet observé que dans le cas où les tests sont significatifs (effet non nul). Dans le cas contraire, on se contera de l'étape descriptive qui nous renseignera sur l'importance ou pas de l'effet, mais sans donner de confiance à l'importance.

Section 5 – Procédure de test des propositions : la méta-analyse par les procédés de l'analyse de comparaison

Dans les sections précédentes nous avons présenté le principe de la méta-analyse qui se résume par une méthodologie de gain de puissance statistique s'intéressant à des données agrégées. Puis nous avons exposé quelques procédures fiducio-bayésiennes nécessaire pour effectuer de l'inférence statistique sur l'importance des effets. Le but poursuivi précédemment était de montrer l'intérêt, d'une part de dépasser les conclusions des tests de signification classique en considérant l'importance des effets, et d'autre part de montrer que pour obtenir cette importance, il suffit d'étudier les résultats de la table de variance. Dans le développement qui va suivre, notre intention est d'adapter les procédures de comparaison bayésienne avec la méta-analyse pour répondre aux objectifs de notre recherche (test de proposition) et montrer que ces mêmes procédures peuvent être utilisées pour mener une méta-analyse classique.

I – Procédure fiducio-bayésienne conjuguée à de la méta-analyse par niveau supra-individus

Comme notre objectif est de tester des propositions qui sont relatives à des unités (ici les magasins), procéder par méta-analyse était inévitable pour justifier notre démarche. La méta-analyse comme principe nous a permis d'effectuer des calculs sur des agrégats, pour produire des statistiques (indices⁹⁷). Par conséquent les indices obtenus par magasin deviendront des données sur lesquelles une analyse statistique peut être abordée. La méta-analyse par niveau magasin revient à synthétiser l'information en agrégeant les résultats par magasin, puis d'analyse, le tout en considérant le magasin comme un individu statistique.

⁹⁷ Comme nous verrons dans le chapitre résultats. Ces indices sont des quantités basées sur des statistiques descriptives, qui ont pour but de représenter un phénomène au niveau de l'unité magasin.

A chaque magasin sera donc associé un indice (calculé selon la proposition) qui devient la quantité d'intérêt qu'il faudra expliquer. Les procédures fiducio-bayésiennes interviennent en deuxième temps pour faire des analyses sur l'échantillon des magasins (Figure 21).

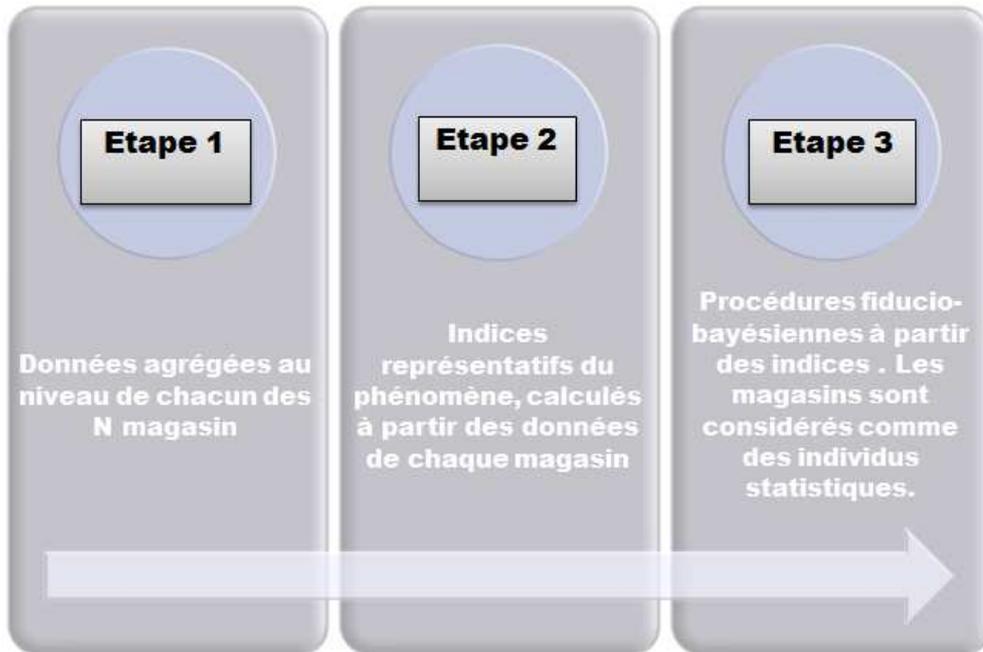


Figure 21 : conjugaison de la méta-analyse et procédures fiducio-bayésiennes

Dans notre démarche, la méta-analyse a un caractère symbolique, elle nous sert à justifier la production des agrégats (ici indice). Ces quantités « estimées » par agrégation seront précises grâce au nombre de données disponibles dans chaque magasin. La procédure d'agrégation permettant de calculer un indice pour chaque magasin est fiable.

Bien sûr ces indices choisis pour représenter les propositions auront tous des statuts de statistiques descriptives. Ils servent à traduire fidèlement les propositions sans être affectés par le nombre de données ni par les unités magasins. Cependant contrairement à la procédure classique de la méta-analyse, nous ne chercherons pas à établir si les indices qui traduisent les propositions sont importants ou non nuls, on cherchera plutôt à expliquer leurs variations. En faisant le parallèle avec les méta-analyses classiques, le magasin joue le rôle de l'étude et l'indice représente les résultats, sauf que dans notre cas

on ne conjugue pas les études pour gagner en puissance statistique, c'est garanti par le nombre de données dans chaque magasin, mais on essaiera d'expliquer la fluctuation des indices entre magasins. L'étude de la variation des indices sera rapprochée à des quantités relatives au niveau du magasin. Dans notre recherche, il s'agit des facteurs de contingence.

En résumé, les procédures fiducio-bayésiennes interviennent dans une phase d'analyse et la méta-analyse justifiée la considération des données agrégées par magasin.

II – Apport des procédures fiducio-bayésienne aux méta-analyses classiques

Une autre approche de l'utilisation des méta-analyses et les approches fiducio-bayésienne peuvent être envisagées dans les enquêtes barométriques. Il s'agit en effet, grâce aux données désagrégées (niveau individuel), de faire toutes sortes de méta-analyses, cette fois de façon classique, en établissant l'importance des effets directement par les procédures fiducio-bayésienne. Cette approche ne sera pas utilisée dans cette recherche, néanmoins son apport pourra être notable pour résoudre certains problèmes liés à l'abondance des données. En effet, dans les enquêtes barométriques, les données individuelles sont nombreuses, et si on se contentait des tests de signification, tout paraîtrait significatif et aucune information pertinente ne serait tirée. La méta-analyse comme méthode de gain en puissance statistique n'est plus probante, d'où l'intérêt de transposer la procédure de gain de puissance statistique du test statistique à l'importance des effets, en utilisant les procédures fiducio-bayésiennes.

A – Procédure statistique classique pour mener la méta-analyse

La méta-analyse est la compilation dans un nombre k d'études similaires (essais) où chaque étude fournit une indication sur l'importance de l'effet qu'on notera μ_i accompagné d'une précision π_i (inverse de la variance) propre à l'étude. La recherche de l'effet commun, et donc la vraie importance de l'effet μ , se traduit par la moyenne pondérée des paramètres μ_i par les π_i . Cette pondération est nécessaire car la variance de chaque étude n'est pas pareille (différence dans les effectifs) et conduit à donner plus d'importance aux études à faibles variances qu'aux études à fortes variances. De plus, cette estimation par moyenne pondérée estime sans biais l'effet vrai (parent). Nous retrouvons ici la conjugaison des études en suivant le principe que plus l'étude est précise, plus son poids est important dans la méta-analyse. Après ces résultats un test de

signification classique est effectué. Cependant, comme cela a été précisé avant, comme tout est significatif, le cumul le sera aussi, pourvu qu'il n'existe pas un gros déséquilibre dans les données de chaque étude⁹⁸.

B – Procédure fiducio-bayésienne pour mener la méta-analyse classique

L'analyse fiducio-bayésienne présente un avantage majeur qui est celui de gagner en précision quand les données sont conséquentes. Elle permet de conjuguer l'information via les techniques bayésiennes de telle sorte que les informations d'une étude antérieure sont conjuguées à la nouvelle, sous forme d'informations a priori. Cette propriété de gain en précision qui accompagne la croissance du nombre de données est la clé pour mener les méta-analyses. Cette conjugaison étant effectuée sur l'importance des effets directs, elle permet de mieux apprécier l'effet grâce aux distributions fiducio-bayésiennes.

De plus, cette procédure qui ressemble fort aux méta-analyses cumulatives, va permettre d'exécuter les méta-analyses via les techniques de l'analyse de comparaison bayésienne de telle sorte que la taille de l'effet sera quantifiée sur **l'ensemble des études simultanément**. Il s'agit ici d'une procédure de réactualisation des connaissances de telle sorte que l'injection des études une par une va permettre, de gagner en précision et d'obtenir un indicateur final de l'effet plus précis. Dans l'approche bayésienne l'estimation à partir de l'ensemble des études prises conjointement, est équivalente à l'introduction une par une les études sous forme de loi a priori.

Le deuxième avantage est que les tests d'association, existence de l'effet commun, et d'hétérogénéité, tous les effets dans chaque étude sont similaires, ne sont pas à considérer, car la procédure fiducio-bayésienne prend cela en considération. S'il existe une hétérogénéité, elle sera prise en compte et si l'effet commun n'existe pas la distribution finale sera diffuse (non précise) ce qui engendre des résultats non concluants sur l'importance de l'effet. Donc un effet non existant.

Ainsi on aura identifié au moins deux façons d'effectuer des analyses qui peuvent servir dans les études barométriques.

Dans notre démarche (tests des propositions) nous avons adapté la méta-analyse à nos besoins pour justifier notre démarche d'agrégation de données par niveau de magasins puis d'analyser les indices (résultats de l'agrégation) via des sources externes

⁹⁸ Voir l'article de Rouanet et al (2002) : Régression et analyse géométrique des données : réflexions et suggestions

(facteurs de contingence). Le but ici étant d'étudier les propositions traduites par ces indices faces à l'existence des facteurs de contingence.

La deuxième procédure quant à elle sera très utile lorsque les données sont conséquentes et surtout si le but poursuivi est juste d'établir les faits, à savoir faire une méta-analyse classique.

C – Application de la procédure aux données

De manière pratique nous utilisons trois indices. Chaque indice traduit donc une proposition. Ces indices sont évalués pour chaque magasin. Ce qui permet d'obtenir un échantillon statistique, sa taille est égale au nombre de magasin. La distribution des indices ainsi obtenue sera le support pour effectuer les tests.

Pour la proposition P1 : corrélation entre indicateur de satisfaction/ efforts marketing. L'indice privilégié est « le coefficient de corrélation multiple ». Ce dernier, calculé pour chaque magasin, aura pour but de juger de la liaison « indicateur de satisfaction/ efforts marketing » pour chaque magasin. Ce qui permet de traduire statistiquement la proposition.

Pour la proposition P2 : pouvoir discriminant de l'indicateur lorsque nous cherchons à évaluer la performance entre magasins. L'indice privilégié est « la variance intra-magasin ». Cet indice descriptif permet d'expliquer la forme de l'indicateur (obtenue par moyenne de note de satisfaction ou par taux de satisfaction). Comme pour P1, un échantillon d'indices est obtenu, sa taille est égale au nombre de magasins. Cet indice traduit la proposition P2 au niveau de chaque magasin. Plus la variance intra magasin est petite, plus l'indicateur est bien discriminant (les résultats de comparaison par moyenne ou taux sont très proches), plus la variance intra magasin est grande plus l'indicateur est biaisé (les résultats de comparaison par moyenne ou taux sont différents).

Pour la proposition P3 : sensibilité des indicateurs dans le temps, l'indice retenu est la « distorsion relative ». Cet indice permet de juger de l'adéquation entre une distribution des notes de satisfaction obtenues et une loi uniforme. Pour chaque vague (soit au global, soit par magasin) la distribution des notes de satisfaction est comparée à la loi uniforme. Par exemple pour les 16 vagues, nous obtenons 16 indices « distorsion relative ». Ces indices résument la proposition P3. Ce qui nous permet de juger de la sensibilité en les comparant entre eux.

Une fois l'échantillon des indices obtenu, un rapprochement est effectué entre la distribution des indices et la présence des facteurs de contingence. L'importance des effets (η) est calculée pour les facteurs de contingence considérés et une inférence fiducio-bayésienne est procédée pour donner une confiance aux résultats obtenus (étapes d'inférence statistique).

A l'issue de ces étapes, en travaillant directement sur les indices (agrégés au niveau magasin) nous statuons sur la vraisemblance des propositions. L'idée est de rejeter les propositions si les actions des facteurs de contingence sont importantes.

Chapitre 4 : Résultats

Nous avons exposé précédemment trois propositions relatives à l'usage d'un indicateur de satisfaction dans sa fonction de contrôle. Ces propositions représentent les qualités indispensables que doit avoir un indicateur de performance lorsqu'il est employé pour effectuer des comparaisons transversales et longitudinales.

Le but principal des procédures de tests qui vont suivre, est de pouvoir vérifier si les qualités d'un indicateur exprimées à travers les trois propositions, sont constatées dans la réalité de la mesure barométrique de la satisfaction. Pour ce faire, nous exploiterons des données issues de deux baromètres relatifs à deux enseignes de grande distribution représentant deux secteurs : Prêt-à-porter et Bricolage. Chaque baromètre constitue donc une étude de cas sur laquelle le test va être procédé. Le premier baromètre, que l'on notera B1, relatif au secteur du Prêt-à-porter, est composé de plusieurs items. Une partie d'entre eux concerne les attributs de la satisfaction et une autre est relative à la satisfaction générale⁹⁹. Le baromètre B1 se rapporte à plusieurs magasins répartis dans plusieurs régions et sur deux pays : France et Allemagne. B1 sera l'outil privilégié pour le test des deux premières propositions. Ces dernières concernent les fonctions de l'indicateur de satisfaction dans la procédure de comparaison transversale (entre magasins). Le deuxième baromètre B2, relatif au secteur du bricolage, concerne l'évolution dans le temps des attributs de la satisfaction par unité magasin. Dans ce baromètre, plusieurs régions de France sont renseignées avec 16 périodes d'enquêtes correspondantes à 4 années et 4 trimestres. Son caractère temporel a servi au test de la troisième proposition liée à la comparaison longitudinale.

Les procédures de test sont d'ordre statistique, c'est-à-dire qu'elles permettent de transformer les énoncés des propositions par des calculs statistiques (indices) qui représentent le contenu des propositions. Une fois ces indices obtenus, des analyses supplémentaires sont effectuées sur leurs valeurs pour remonter à la signification émise par les propositions. En se basant sur les résultats obtenus, les propositions seront testées.

Pour P1, l'objectif du test est de savoir si les indicateurs reflètent bien l'effort marketing. Il s'agit ici d'étudier les liaisons entre l'indicateur de satisfaction et les éléments qui

⁹⁹ Rappelons que dans une enquête de satisfaction clients, il y a toujours une question (item) qui concerne la satisfaction générale et d'autres questions qui sont relatives à tous les attributs de service/ou du produit qui est jugé. Cette structure est celle qui prédomine lors des études de satisfaction clients.

représentent l'effort marketing (levier de satisfaction : attributs de la satisfaction). Le lien sera étudié conditionnellement à chaque magasin. L'indice privilégié ici est le « Coefficient de corrélation multiple R ». Il permet de juger de l'étendu de la liaison : indicateur de satisfaction et Effort Marketing. Une fois cet indice calculé par magasin, nous étudierons, sur la base de l'échantillon des magasins, sa variation à la lumière des facteurs de contingence¹⁰⁰.

Pour P2, l'indicateur de satisfaction sera étudié directement dans sa fonction de comparaison transversale (entre magasins). Nous examinerons l'objectivité de l'indicateur lors de la comparaison en termes de satisfaction clients. Pour qu'il soit objectif, l'indicateur de satisfaction doit refléter seul les effets de l'action marketing. Si ce dernier détecte une différence, elle ne peut être liée qu'au fait du management.

Pour arriver à nos résultats, nous nous sommes intéressés à examiner l'indicateur selon les deux types de calcul: moyenne et taux¹⁰¹. L'idée recherchée pour cet examen est de pouvoir comparer ces deux types d'indicateurs lors d'un classement des unités magasins. En d'autre terme, nous avons voulu savoir si le type de calcul apporte les mêmes informations¹⁰² lorsque on s'intéresse à classer les magasins suivant la satisfaction de leurs clients. De même que pour P1, les facteurs de contingence seront émis pour expliquer les fluctuations dans le classement entre magasins à travers un indice commun aux deux types de calcul d'indicateurs (moyenne et taux) qui est la variance intra-magasin.

Pour la proposition P3, la sensibilité temporelle des indicateurs est étudiée à travers un indice qui résume la distribution des notes de satisfaction enregistrée dans chaque vague d'enquête. Il servira donc à la comparaison temporelle (entre vague d'enquête). Les comparaisons seront effectuées sur un niveau agrégé (sans distinguer des

¹⁰⁰ Les facteurs pris en compte ici sont : (1) l'âge, le sexe et la fréquentation des répondants, facteurs liés aux caractéristiques des individus. (2) Des facteurs géographiques région et pays, facteurs liée à l'environnement des individus.

¹⁰¹ Pour le taux, il est souvent obtenu en calculant le pourcentage des satisfaits. Pour une échelle de 1 à 10. Le taux de satisfaits est le nombre de clients ayant donné une note supérieure à 6, divisé par le nombre de clients interrogés.

¹⁰² Lorsqu'une différence, dans le classement est identifiée suivant l'un ou l'autre type de calcul, la procédure de test cherchera à connaître si les informations apportées sont les seules faits de management ou éventuellement d'autres sources de variation. Comme on notera dans les développements suivants, une différence entre les classements suivant les taux de satisfaction ou les moyennes, est due à la variance intra-magasin (variance des avis de satisfaction dans chaque magasin). Cet indice deviendra la variable d'intérêt sur laquelle la procédure de test de la proposition P2 sera effectuée.

magasins) et au niveau de chacun des magasins. Les résultats obtenus nous indiqueront le potentiel de sensibilité temporelle de l'indicateur de satisfaction. Ce qui permettra de répondre au test de la proposition P3. Le support utilisé pour ce test concerne un attribut qui est: la qualité du produit fourni. Cet attribut est indépendant du contexte magasin, c'est pourquoi il a été choisi.

Section 1 – Procédure et résultats du test de la proposition P1

Le test des propositions P1 et P2 a pour support un baromètre de satisfaction (noté B1), qui est issu d'une enquête de satisfaction menée sur deux pays : France et Allemagne. Le cumul, en termes d'individus, dépasse les 172400 observations. Nous travaillerons sur la base de 221 magasins répartis sur deux pays : France et Allemagne. Pour rappel leur répartition géographique est visible dans la carte suivante :

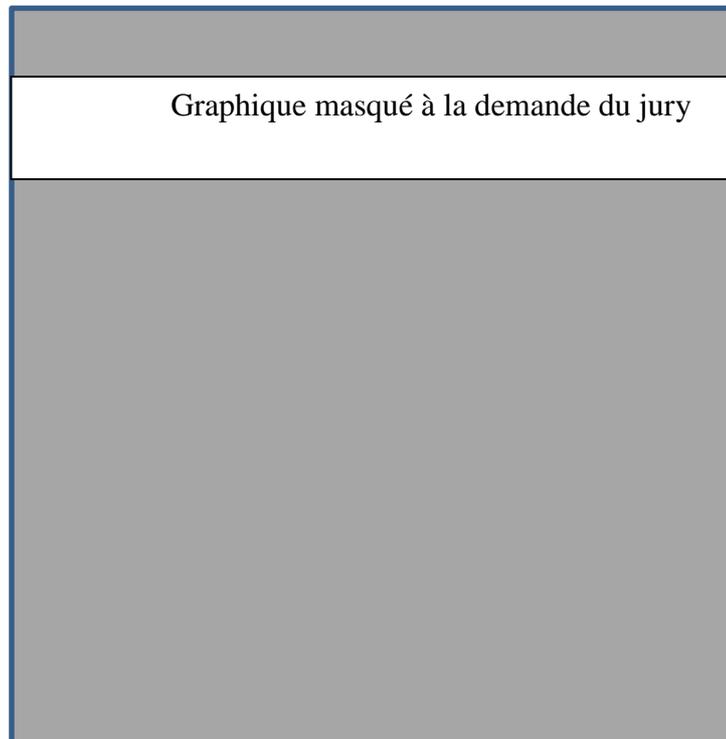


Figure 22 : Répartition géographique des magasins du baromètre B1 en fonction de la satisfaction moyenne

La question de satisfaction globale est l'item sur lequel l'indicateur est calculé. Elle est accompagnée par les vingt attributs supposés expliquer l'avis de la satisfaction clients (effort marketing). Dans l'absolu, ces attributs sont les éléments qui expliquent la variation de la satisfaction globale envers le magasin. Ils sont utilisés comme des leviers

sur lesquels l'action marketing est procédée. Par exemple, l'amélioration de l'offre doit conduire à une augmentation du niveau de satisfaction au globale ou encore une meilleure performance de l'accueil clients conduira à une augmentation de la satisfaction clients etc.

Pour cette proposition P1, nous voulons savoir si les résultats d'un indicateur de satisfaction sont le fruit du seul fait de l'effort marketing. Cette propriété est primordiale lors des procédures de comparaison sur la base des indicateurs. Si cette condition est respectée, les résultats seront interprétables en fonction. Le fait qu'un indicateur S1 est supérieur à un autre indicateur S2, doit être synonyme de différence dans l'action marketing et rien d'autre. Dans le cas des enquêtes de satisfaction, l'ensemble des actions marketing est représenté par des questions secondaires dites « attributs de la satisfaction ». Ces attributs servent à expliquer la satisfaction cumulée des clients qui est-elle même représentée par une question dite « satisfaction générale ou globale ». Ainsi pour savoir si l'indicateur représente bien l'effort Marketing, il faudra étudier la liaison entre la satisfaction générale et ses attributs. L'examen de cette relation nous renseignera sur l'information fournit par l'indicateur de la satisfaction.

La procédure vise à examiner si les attributs de la satisfaction expliquent la satisfaction globale des clients de façon semblable quelques soit l'unité magasin. Comme les attributs sont représentés par des questions censées moduler la satisfaction générale, la variation du niveau de satisfaction au global doit être accompagnée par des fluctuations dans les niveaux de satisfaction des attributs, dans le même sens, avec la même intensité et quelque soit le magasin. Les attributs de satisfaction sont des quantités qui construisent la satisfaction générale et sont par construction indépendantes du magasin étudié. Une personne satisfaite sur les attributs du service ou du produit aura une bonne satisfaction au global. Ce lien attribut/satisfaction générale, doit être uniforme quelques soient les conditions dans lesquelles il est présenté. La vraisemblance de ce fait permettra d'effectuer des comparaisons fiables entre les magasins. Si cette propriété fait défaut, les comparaisons seront biaisées, ainsi que la prise de décision.

Ainsi, l'idée poursuivie dans ce test est d'examiner le pouvoir explicatif des attributs de satisfaction sur la satisfaction générale suivant un échantillon de magasins. Le pouvoir explicatif ou la liaison doit être la même par construction, quelle que soit l'unité magasin. Sinon cela voudra dire que dans certains magasins, les attributs construisent bien l'avis global, et dans d'autres un peu moins. Ce fait devient dangereux dans le cas où

l'action future (amélioration des attributs) n'aura pas le même impact par magasin. Ce qui signifie aussi que le management par attributs n'est pas suffisant, et d'autre part que la satisfaction globale est incontrôlable.

Dans l'absolu, ce pouvoir explicatif (satisfaction générale et attributs) ne doit dépendre que du fait de l'action marketing. Les attributs de satisfaction sont les représentants de l'effort Marketing et constituent les leviers d'action. Une baisse ou une augmentation du niveau de la satisfaction globale doit être liée à une baisse ou une augmentation des niveaux de satisfaction des attributs. Ainsi la liaison entre satisfaction générale et attributs est importante car, lorsque elle fait défaut, l'indicateur de satisfaction (calculé à partir la satisfaction générale) perd tout son sens.

Le test de la proposition P1 s'est orienté sur l'étude de cette liaison. Il vise à éprouver à travers des études de cas le pouvoir explicatif des attributs, en utilisant un panel de magasins. Le but ici est d'examiner les propriétés de l'indicateur de satisfaction dans son rôle de contrôle. C'est-à-dire savoir si les variations de l'indicateur de satisfaction (satisfaction générale) par magasin sont exclusivement dues aux fluctuations dans l'effort marketing (des attributs) ou si ces sources de variation ont des causes externes.

I – Indice pour juger de pouvoir explicatif d'un ensemble de variables

Les attributs sont des questions, variables, qui deviennent des quantités statistiques lorsque les avis clients sont récoltés. La satisfaction générale est aussi une question (variable) qui est-elle même assimilée à une grandeur statistique. Pour quantifier le pouvoir explicatif des attributs envers la satisfaction générale, l'indice privilégié sera le coefficient de corrélation dans sa version multidimensionnelle. Cet indice est dit le coefficient de corrélation multiple¹⁰³, permet de quantifier la corrélation des valeurs de la satisfaction globale avec les valeurs estimées à partir des attributs de satisfaction. Ainsi il permet de donner une idée du pouvoir explicatif des attributs de satisfaction envers la satisfaction globale. Il est compris entre [0,1], plus il proche de 1 plus la liaison est forte, et inversement plus il proche de 0 plus la liaison sera considérée comme faible. Il est défini par l'équation suivante :

$$R_{y(x)} = \sqrt{\frac{SCE_{yx}}{SCE_y}}$$

103 En anglais : multiple corrélation coefficient.

Avec SCE_{yx} la part (de y) expliquée par les attributs x, et SCE_y la variation totale de y.

Cet indice mesure d'une façon générale la qualité de la prévision linéaire de la variable y à partir des attributs x. Ainsi les valeurs de $R_{y(x)}$ nous donnerons le niveau de corrélation (pouvoir explicatif) qui peut exister entre la satisfaction générale (ici noté y) et les attributs de satisfaction (ici noté x). Il sera interprétable comme une corrélation simple, car il s'agit de sa généralisation à plus de deux variables.

II – Organisation des données pour le test de P1

La proposition P1, est relative à la liaison entre les attributs de satisfaction et la satisfaction globale. En statistique, étudier une liaison se traduit aisément par un indice de corrélation. Cet indice permet de quantifier la variation de deux groupes de variables et donne une idée de l'importance de cette variation. Pour notre cas (la proposition P1) nous voulons savoir si cette liaison est sensible aux magasins, en d'autres termes, nous voulons connaître si la corrélation (attributs et satisfaction globale) varie suivant les magasins.

Pour chaque magasin considéré (221 magasins) nous avons calculé Le $R_{y(x)}$ correspondant à la liaison entre attributs de satisfaction et la satisfaction globale. Les magasins sont situés en France métropolitaine et en Allemagne. La dispersion des coefficients de corrélation multiple entre les 221 magasins, est résumée dans le tableau et les graphiques suivants :

Tableau 8: Dispersion des valeurs du coefficient de corrélation multiple

Min.	1st Quartile	Médiane	Moyenne	3rd Quartile	Max	Écart-type	Coefficient de variation
0.442	0.6032	0.6486	0.6487	0.695	0.823	0.07	10,70%

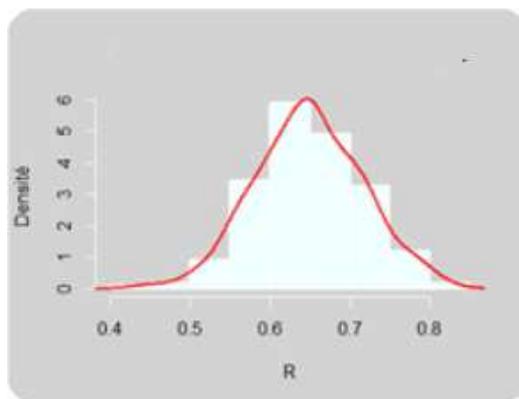


Figure 23: Variation du coefficient de corrélation multiple, sur l'ensemble des 221 magasins.

En premier lieu, nous remarquons que le coefficient de corrélation multiple $R_{y(x)}$, n'est jamais égal à 1. Les attributs n'expliquent pas 100% de la variation de satisfaction générale. Ce fait est naturel dans une considération statistique car il existe toujours une erreur aléatoire issue de la considération d'un échantillon de sondage et du fait que les attributs ne sont pas tous représentés, car seule une partie est prise en compte¹⁰⁴. Cependant la variation du $R_{y(x)}$ paraît comme une erreur systématique plus qu'une erreur aléatoire (figure 23).

Le dernier tableau (Tableau 8) nous montre aussi que le pouvoir explicatif des attributs est sensible au magasin. Avec un écart type de 0,07 et un coefficient de variation de 10,7%. Un coefficient de variation de 10,7% signifie, dans le cadre de la loi normale, que 95% des valeurs du $R_{y(x)}$ fluctuent de plus ou moins 20% de la valeur moyenne. Ce qui à notre sens, révèle une variabilité non négligeable. De ce fait le pouvoir explicatif paraît fluctué selon le magasin. Reste à savoir si cette différence est fortuite (erreur aléatoire), ou si elle est causée par la présence des facteurs contingents (erreur systématique), chose qu'on s'attèle à vérifier par les analyses qui vont suivre.

De plus ces résultats nous indiquent aussi que le pouvoir explicatif des attributs de satisfaction par magasin varie. Ceci signifie que les avis de satisfaction générale des clients exprimés dans chaque magasin ne concordent pas de la même façon entre les avis donnés sur les attributs de satisfaction : *La satisfaction globale n'est pas attachée aux mêmes éléments suivant le magasin.* Ces remarques nous orientent vers le fait que la

¹⁰⁴ Dans notre cas, nous avons 20 attributs supposés moduler la satisfaction. Toutefois il est clair que ces vingt attributs ne représentent pas à eux seuls tout l'effort marketing, d'autre non mesurée sont aussi à considérer

satisfaction globale exprimée, n'est pas le résultat des seuls attributs mais aussi des autres éléments attachés à la population sondée. Cette situation paraît totalement contingente et dépendante de l'unité magasin. Comme le jugement vient de clients et que ces clients ne sont pas les mêmes d'un magasin d'un autre, il nous a semblé pertinent de rapprocher la fluctuation des liaisons (satisfaction générale et attributs) aux caractéristiques de la population et de son environnement afin d'expliquer cette variation dans le pouvoir explicatif. L'étude de l'effet de la population sur la liaison entre attributs et satisfaction générale nous renseignera sur la consistance de l'indicateur de satisfaction à fournir des informations propres relatives au seul fait de l'effort Marketing.

D'un autre côté, un calcul supplémentaire résumant cette fois les attributs en quatre facteurs¹⁰⁵ :

- Offre
- Merchandizing
- Commodité du magasin
- Personnels du magasin

Nous a permis d'effectuer des régressions sur ces quatre facteurs à l'intérieur de chaque magasin suivant l'équation :

$$\mathbf{Satisfaction}_{ij} = a_{0j} + a_{1j}\mathbf{Offre}_{ij} + a_{2j}\mathbf{Merch}_{ij} + a_{3j}\mathbf{Commodite}_{ij} + a_{4j}\mathbf{Personnels}_{ij} + e_{ij}$$

Où (i individus et j magasins)

Les résultats de ces régressions conditionnelles par magasin, nous ont permis de remarquer que les coefficients (l'importance) de chaque facteur, varient sensiblement suivant le magasin (figure 24). Ces résultats conforte les variations du $R_{y(x)}$ constatées auparavant, et légitime l'utilisation de l'indice¹⁰⁶ (corrélation multiple) comme quantité pour juger des relations attributs et satisfaction générale.

¹⁰⁵ Il s'agit d'une analyse factorielle sur l'ensemble des attributs qui a permis de dégager quatre facteurs

¹⁰⁶ Le coefficient de corrélation multiple est la racine carré du R^2 issue de la régression linéaire multiple.

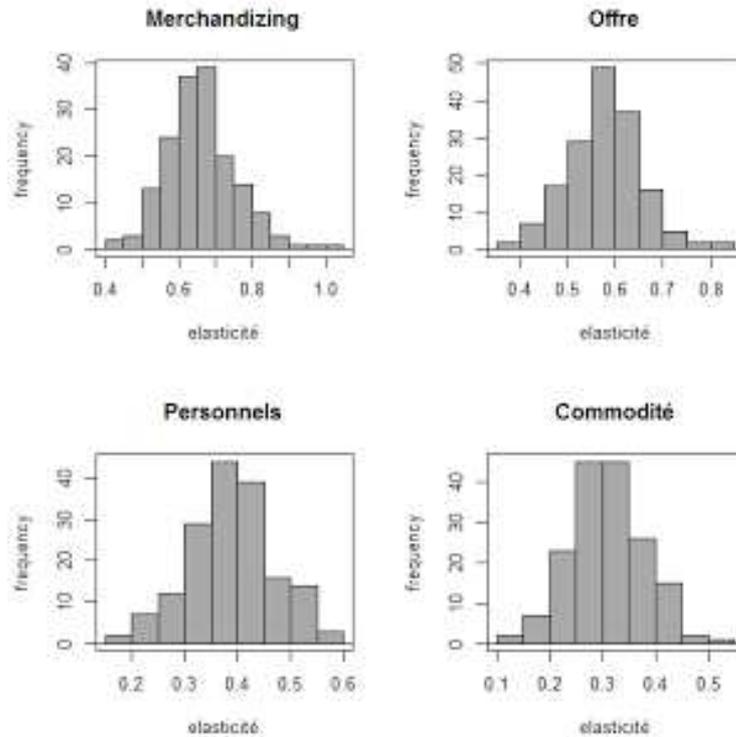


Figure 24 : Variations des coefficients des quatre facteurs

La variation des coefficients (figure 24) nous indique que chaque groupe d'attributs ne contribue pas de la même manière dans l'explication de la satisfaction globale. Ce qui veut dire aussi que la satisfaction globale n'est pas construite de la même manière suivant le magasin considéré. Ainsi la procédure de test qui suivra aura pour objectif d'expliciter la variation dans la liaison attribut/satisfaction globale, le tout en considérant l'indice de la corrélation multiple, comme variable d'intérêt pour effectuer le test de la proposition P1.

En décrivant les magasins selon leurs positions géographiques et selon les valeurs de $R_{y(x)}$ de telle sorte que les différentes valeurs du coefficient de corrélation multiple soient identifiées pour chaque magasin (figure 25). Nous remarquons que la variation du $R_{y(x)}$ peut présenter une dépendance géographique. Cette dépendance peut se traduire par le type de clients fréquentant le magasin (leurs caractéristiques), par la région ou le pays où le magasin se trouve.

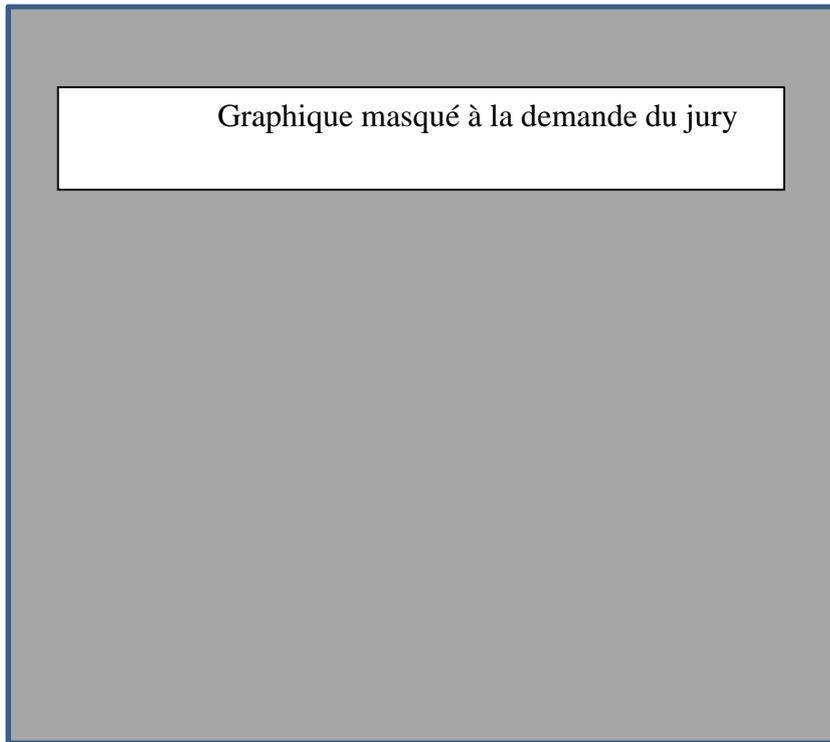


Figure 25: Répartition géographique des valeurs des coefficients de corrélation multiple suivant les magasins

Pour effectuer l'analyse supplémentaire qui servira à tester la proposition P1, nous nous sommes intéressés à d'autres informations relatives aux individus sondés (l'âge, le sexe, la fréquentation du magasin) et nous avons recueilli d'autres informations relatives à leurs environnement qui sont: la région et le pays. Ces informations supplémentaires résumant les facteurs de contingence vont servir à vérifier si la fluctuation du $R_{y(x)}$ constatée est sensible à l'action de ces derniers. Le but de cette démarche est de détecter si les variations peuvent être attachées à des erreurs systématiques, liées aux facteurs de contingence, ou si elles sont uniquement les résultantes des erreurs aléatoires.

Les données sur lesquelles le test va être basé sont les données agrégées au niveau magasin, car la quantité à expliquer ici est le coefficient de corrélation multiple $R_{y(x)}$, qui est par nature relatif à un magasin. Ainsi les données sur lesquelles l'analyse explicative va s'agencer doivent concerner le même niveau (ici le magasin). Les variables représentant les facteurs de contingence, prises en compte sont:

- La fréquentation: Elle sera représentée par un score de 10 à 60, plus la valeur est proche de 60 plus la fréquentation est élevée.
- Le sexe des répondants: Elle sera représentée par le pourcentage des femmes interrogées dans le magasin.

- L'âge des répondants: Elle sera représentée par un score de 10 à 60, plus la valeur est proche de 60 plus la population sondée est âgée.
- La région: Il s'agit d'une partie des régions en France et en Allemagne, c'est une variable catégorielle à 50 modalités.
- Le pays: variable à deux modalités correspondantes aux deux pays: la France et l'Allemagne

III – Effets des facteurs de contingence sur la variation du coefficient de corrélation multiple.

Suivant les quelques facteurs identifiés, nous étudierons la variation du $R_{y(x)}$ (pouvoir explicatif) à la lumière de ces facteurs. Cette analyse nous permettra d'apprécier les causes de variations en les rapprochant aux facteurs de contingence. Ceci nous permettra de rejeter l'hypothèse d'une fluctuation au hasard. Bien sûr, étant donné le déséquilibre dans le nombre des répondants par magasin, l'analyse via les comparaisons bayésiennes sera effectuée en tenant compte de l'effectif par magasin. Le nombre d'individus par magasin sera considéré comme un poids lors de l'analyse.

L'étape d'analyse va concerner le coefficient de corrélation multiple, c'est cet indice qui sera traité en fonction des facteurs de contingence. L'esprit de la méta-analyse reste entier, en effet, comme l'objet est de tester une proposition relative à des comparaisons transversales (entre magasins), l'agrégation des données a permis de considérer le coefficient $R_{y(x)}$ comme la variable d'intérêt car elle est attachée à un magasin et traduit directement la formulation de la proposition P1, à savoir la liaison entre indicateur de satisfaction et actions marketing.

Les résultats des analyses via les comparaisons bayésiennes, dans ce cas des **Anova**, nous donnent les conclusions suivantes:

- Effets de l'âge des répondants sur le coefficient de corrélation multiple entre magasins

Tableau 9: Effets de l'âge des répondants sur le coefficient de corrélation multiple entre magasins

Corrélation multiple	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Facteur Âge	1	0.00001643	0.00001643	1.9740	0.1614
Residuals	219	0.00182276	8.32311E-06		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
0.89%	--	--	--

Les résultats de l'analyse de la variance nous montrent que l'importance de l'effet de l'âge des répondants est faible (0,89%) sur la variation du coefficient de corrélation multiple. Ceci signifie que l'âge des répondants explique moins de 1% de la variance du $R_{y(x)}$. Autrement dit, la fluctuation entre magasins des liaisons « satisfaction (générale) et action marketing (attributs de la satisfaction) », ne peut pas être attachée à l'effet de l'âge. De plus cet effet est non significatif ($p_value=0,16$), donc absence de généralisation au niveau inductif. Cependant d'un point descriptif cela peut être considéré comme négligeable.

- Effets du sexe des répondants sur le coefficient de corrélation multiple entre magasins

Tableau 10 : Effets du sexe des répondants sur le coefficient de corrélation multiple entre magasins

Corrélation multiple	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Facteur SEXE	1	0.000016086	0.000016086	1.9323	0.1659
Residuals	219	0.0018231	8.32466E-06		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
0.87%	--	--	--

Les résultats de l'analyse nous montrent une faible importance de l'effet du sexe des répondants sur la variation du coefficient de corrélation multiple. Cet effet est non significatif ($p_value=0,16$), et peut être considéré comme négligeable suivant les résultats descriptifs.

- Effets de la fréquentation du magasin sur le coefficient de corrélation multiples entre magasins

Tableau 11 : Effets de la fréquentation du magasin sur le coefficient de corrélation multiples entre magasins

Corrélation multiple	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Fréquentation	1	0.00003597	0.00003597	4.3685	0.03776
Residuals	219	0.00180322	8.23388E-06		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
1,95%	27,00%	6,00%	0,00%

Les résultats nous indiquent qu'il existe un effet significatif ($p_value=0,03$) de la fréquentation sur la variation du coefficient de corrélation multiple $R_{y(x)}$. Cet effet n'est pas important d'un point descriptif (1,95%), mais on ne peut avancer le fait qu'il soit négligeable, car on a 72% de chance qu'il soit supérieur au seuil de 1%. L'effet paraît comme faible, mais non négligeable. On conclut donc à une influence modérée du score de la fréquentation sur la variation des coefficients $R_{y(x)}$. Ce qui signifie aussi que la variation du pouvoir explicatif des attributs de la satisfaction par magasins est modulée par l'ancienneté des clients (leurs fréquentations). Ces résultats traduisent aussi que les profils des personnes interrogées diffèrent, en termes de fréquentation, entre magasins. Cette différence affecte, de façon modérée, la liaison : « Satisfaction globale et attributs ». L'action de cet effet ne sera pas considérée comme importante; en revanche elle sera posée comme existante.

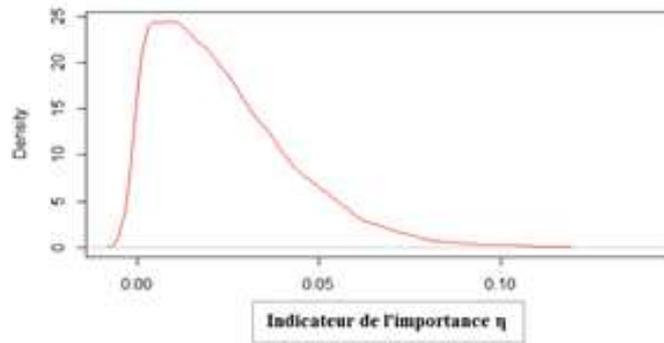


Figure 26 : Distribution des valeurs de l'importance de l'effet de la fréquentation

- Effets de la région et le pays sur la variation du coefficient de corrélation multiple entre magasins

La région et le pays sont des facteurs emboîtés, cela veut dire que la région est forcément nichée dans un pays, ainsi une analyse directe par les procédés des analyses de variance peut fausser le résultat car l'effet de l'un est confondu avec l'autre. De plus, comme nous possédons seulement un échantillon par région (toutes les régions ne sont pas représentées), il faudra raisonner sachant ce plan hiérarchique. Pour remédier à ce problème, nous utiliserons des analyses d'**Anova** dite « hiérarchique »¹⁰⁷ à partir du plan **Magasin<Régions<Pays>>**¹⁰⁸. La lecture du plan permet de voir que l'effet du pays ne peut être que relatif à la variation $R_{y(x)}$ dans la région, et l'effet région ne peut être que relatif à la variation du $R_{y(x)}$ entre magasin. Il existe ici deux niveaux dans lesquels les coefficients de corrélation multiple $R_{y(x)}$ varient: le niveau magasin et le niveau région. Deux niveaux qu'il faudra prendre en compte lors de l'analyse de telle sorte que:

- a. L'effet du pays sera analysé en fonction de la région en considérant la variance expliquée par la région comme résiduelles (source d'adjointe) du pays
- b. L'effet de la région sera analysé en considérant cette fois la variance résiduelle classique (ici inter magasins).

¹⁰⁷ Le mot hiérarchique ici renvoie au traitement statistique des plans incluant des variables qui sont subordonnées l'une par rapport à l'autre. Il s'agit de découper la variance résiduelle du modèle selon plusieurs variables. Il y aura donc plusieurs variances résiduelles dont chacune doit correspondre à une source de variation.

¹⁰⁸ Le plan signifie que le magasin est niché dans une région qui elle-même est nichée dans un pays.

- c. Puis, en suivant ces deux considérations que l'importance de l'effet sera calculée pour chaque facteur (pays et région) selon les conclusions du test de signification.

Les résultats obtenus sont:

➤ **Pour l'effet du Pays**

Tableau 12 Effets du pays sur la variation du coefficient de corrélation multiple entre magasins

Corrélation multiple	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Erreur Région					
Pays	1	0.1422	0.1422	43.1176	3.1E-08
Residuals	49	0.1616	0.0033		
Erreur within					
Residuals	170	0.6932	0.004077647		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
46.82%	0.00%	100.00%	100.00%

La table de variance précédente est découpée selon les deux sources de variance : régions (Error : Région) et selon la variance inter magasins (Error : Within). L'effet du pays est calculé par rapport à la région. L'analyse donne un résultat significatif ($p_value=0,00$). Cela signifie que le pays a un effet sur la variation des coefficients de corrélation $R_{y(x)}$ (ici par région). De plus, comme l'affirme les résultats du tableau, l'effet est important (46,82%) soit presque 47% de la variance des $R_{y(x)}$ entre régions expliquée par le pays, le tout avec une confiance de 100% que cette importance soit supérieure à 10%. Le pays est donc un facteur qui explique en partie la variation des coefficients de corrélation multiple avec une importance notable.

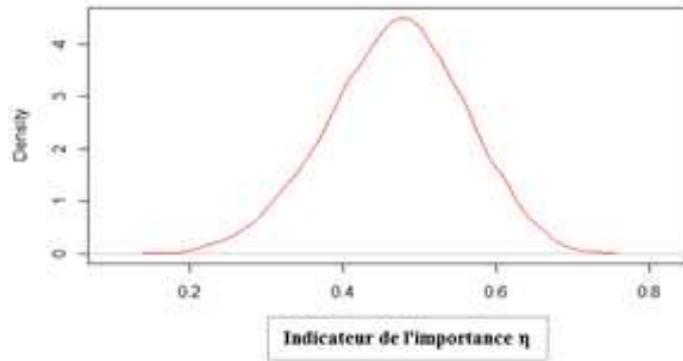


Figure 27: Distribution des valeurs de l'importance de l'effet du Pays

➤ Pour l'effet de la région

Tableau 13 Effets de la région sur la variation du coefficient de corrélation multiple entre magasins

Corrélation multiple	Degré de liberté	de somme des carrés	des somme des carrés Moyen	F	P value
Pays	1	0.1422	0.1422	34.8731	1.86E-08
Pays: Région	49	0.1616	0.0033	0.8088	0.806
Residuals	170	0.6932	0.004077647		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
18,90%	--	--	--

Ici, nous nous intéressons à étudier l'effet région par rapport à la variance des coefficients $R_{y(x)}$ inter magasins (Residuals). Comme la région est nichée dans le pays, l'effet région devient l'effet de l'interaction de la région à l'intérieur du pays (ici représenté par pays : région). Cette quantité correspond tout simplement à **Région<Pays>**. D'après les résultats, l'effet de la région est non significatif (p_value=0,80), ce qui nous empêche de procéder à l'inférence sur son importance. Néanmoins d'un point de vue descriptif, la région possède un effet tout aussi important que le pays (18,9%). Malheureusement on ne peut conclure sur sa notabilité inductive (inférentielle).

Cependant cette absence de significativité peut être vue comme un résultat lié au nombre important des régions (50) par rapport aux nombre de magasin disponibles 221.

En effet comme le fait d'être significatif suppose le non rejet de l'hypothèse nulle, donc l'absence de conclusion, le fait que cet effet soit important sans être significatif n'est pas paradoxal. De plus comme le test de F est sensible au nombre de données, le peu de données disponibles dans notre cas (221 magasins devant 50 régions) handicape le test F pour voir la différence. Cependant, étant donné que le test de signification est non concluant, nous ne pouvons pas affirmer l'importance de l'effet région avec une certaine confiance. Toutefois, comme un test non significatif ne signifie pas forcément une absence d'effet, mais plutôt une incapacité de conclure, l'importance dans la vision descriptive devient nécessaire pour cerner si l'effet est présent.

Par ces résultats, les effets de la région et du pays seront tous deux considérés comme importants dans l'explication des variations du coefficient de corrélation multiple $R_{y(x)}$, avec une certitude sur l'effet pays. Cela conduit à affirmer que les différences du pouvoir explicatif des attributs sur la satisfaction sont sensibles aux facteurs géographiques.

IV – En résumé

Les facteurs contingents sont situationnels, ils concernent la population sondée. L'ampleur de leurs effets sera attachée à la seule étude de cas actuelle, et ne pourra être généralisée de façon globale. D'ailleurs, c'est pour cette raison que le sens de l'influence des effets n'a pas été précisé. Cependant ces effets auront à notre sens la même tendance à faire fluctuer le pouvoir explicatif des attributs dans l'explication de la satisfaction globale.

Nous avons détecté à travers ces analyses, que les facteurs comme la région et le pays ont des effets notables sur la variation des coefficients de corrélation multiple calculés par magasin. C'est-à-dire que la représentativité de l'indicateur de satisfaction est sensible aux positions géographiques du magasin. Ces conclusions cachent une autre information qui est celle de la provenance des clients. Comme pour deux magasins les clients ne sont pas les mêmes, et proviennent de endroits différents, la représentativité de l'indicateur (indicateur représentant bien l'effort marketing) est plus attaché aux profils de personnes, et donc à la zone géographique d'où ils viennent. Aussi nous avons remarqué que la fréquentation des clients peut également moduler cette variation et affecter la représentativité de l'indicateur de satisfaction. Cette représentativité est attachée à la capacité de l'indicateur à bien représenter l'effort marketing.

Ainsi ces résultats constituent les contre-exemples qui prouvent que les variations des pouvoirs explicatifs des attributs par magasin, ne sont pas dues uniquement à l'effort marketing, ni à des erreurs aléatoires, mais très probablement à des facteurs tiers qui sont les facteurs de contingence. Ceci met donc à mal la proposition P1, et met en évidence la possibilité que l'indicateur de satisfaction globale calculé dépend aussi de certains éléments extérieurs qui dépassent le champ de l'action marketing. Par cela, l'évaluation statistique de l'indicateur de satisfaction ne paraît pas totalement objectif, et la comparaison transversale en termes de performance subit ce constat.

En conclusion, l'évaluation des indicateurs de satisfaction subit des erreurs qui peuvent être attachées à des facteurs de contingence qui conditionnent leurs résultats. Ces erreurs influencent donc l'indicateur dans sa fonction de pilotage et causent un problème d'objectivité: l'indicateur de satisfaction reflète donc des variations des actions marketing et de certaines caractéristiques individuelles et environnementales. Par cela, les comparaisons transversales (entre magasins directement) sont sujettes à erreurs.

Section 2 – **Procédure et résultats du test de la proposition P2**

Le test de cette proposition se base sur le même support que la proposition P1. Cependant ici il n'est plus question d'étudier la liaison indicateur et attributs, mais plutôt l'analyse de la structure statistique de l'indicateur à travers son évaluation par la moyenne ou le taux. Pour cet examen, nous nous focalisons sur la seule question de satisfaction globale mesurée par un item. L'évaluation de l'indicateur peut avoir plusieurs supports (ex : ensemble des attributs), toutefois, comme la satisfaction globale constitue l'élément privilégié dans les procédures de contrôle, nous nous sommes retréints à cette dernière.

Afin d'expérimenter (tester) la proposition P2, nous examinerons l'indicateur de satisfaction dans sa fonction de contraste directe entre unités magasins. Quel que soit le type de l'indicateur utilisé, moyenne ou taux de satisfaction, il se doit d'apporter une information homogène (même information quelques soit sa forme). L'indicateur de satisfaction doit être fidèle. Cette propriété est essentielle pour que le pilotage de la performance soit pertinent La notion de fidélité renvoie ici à la capacité de l'indicateur à mesurer fidèlement le phénomène : ici, c'est la satisfaction générale.

Dans cette procédure, notre analyse vise à vérifier si les différences entre niveaux de satisfaction signalées par l'indicateur peuvent être attribuées au seul effet de l'effort

marketing¹⁰⁹. Par exemple: si le niveau de satisfaction dans un magasin M1 est supérieur à celui du magasin M2, cela signifie que l'effet de l'effort marketing (management) est meilleur dans M1 que dans M2 et que les ressources doivent être allouées au magasin M2 pour améliorer sa performance. Un indicateur doit être capable de traduire cette propriété et ne doit en aucun cas détecter des différences autres que celles représentées par ce schéma.

Dans le processus de management de ressources, la satisfaction est un critère sur lequel les managers se basent pour comparer les performances de leurs magasins. Dans cet exercice de contrôle, on est souvent amené à comparer des notes moyennes ou des taux de satisfaits dans chaque unité (magasin). Ce procédé sert avant tout à établir un classement pour que les ressources soient allouées en fonction de la performance et du besoin d'amélioration.

L'idée poursuivie dans ce test est d'étudier les valeurs données par un indicateur comme la moyenne ou le taux, lors d'une comparaison entre magasins. La démarche vise à vérifier si les deux types d'indicateurs traduisent bien, et surtout de la même manière les effets de l'action Marketing.

Dans la pratique, le calcul des indicateurs par magasin sert à classer les magasins selon le niveau de satisfaction obtenu. L'hypothèse implicite derrière cette pratique est de dire qu'un magasin ayant un niveau de satisfaction clients supérieur à un autre, est considéré comme meilleur en termes de management. Cette hypothèse qui est à la base de la prise de décision, a intérêt à être vérifiée. Sinon, dans le cas contraire, un biais décisionnel sera introduit et un manque d'équité entre magasins sera produit. Ici, il s'agit de la légitimité de l'indicateur et surtout de son calcul. Ainsi comme pour le test de la proposition P1, la problématique concerne la capacité de l'indicateur à assurer son rôle de comparaison.

I – Indice pour juger de la représentativité de l'indicateur lors de la comparaison

Dans ce cas, l'objet poursuivi sert à comprendre les différences des niveaux de satisfaction via les indicateurs: moyenne et taux de satisfaction. Les deux indicateurs ont le même support; c'est à dire la question de satisfaction globale. Pour voir si ces deux types d'indicateurs produisent les mêmes résultats, il nous faut un autre indice qui soit

¹⁰⁹ L'effet de l'effort marketing, représente les actions émanant du management du magasin. Si deux magasins enregistrent deux niveaux de satisfaction distincts, la différence est due au management.

descriptif (indépendant du nombre des données), qui soit relatif à la question de satisfaction. L'indice privilégié dans notre cas est la variance des notes de satisfaction. Ce choix vient du fait que lors d'une comparaison entre magasins par la moyenne, on suppose une égalité des variances intra-magasin. Ainsi la variance comme indice de contrôle devient un paramètre à privilégier car si les variances intra-magasin sont différentes, la procédure de comparaison subit un faussage statistique¹¹⁰. La variance est ici considérée comme un outil pour juger de la fiabilité des comparaisons par moyennes.

Pour le taux de satisfaction, comme son nom l'indique, il correspond à une fréquence qui ne tient pas compte de toute la distribution des valeurs. Un taux est calculé majoritairement comme le pourcentage des notes supérieures à un seuil (ex:6/ pour une échelle à 10 points). De ce calcul, il résulte le fait que si tous les répondants donnent une note de 10, le taux sera équivalent à la situation où tous les répondants donnent une note de 8 ou de 9. Ainsi le taux sera identique dans les deux situations, mais en réalité les deux situations ne sont pas les mêmes.

En conséquent, les deux formes d'indicateurs sont tributaires de la distribution des notes, le fait de tenir compte de la variance nous permettra de considérer un paramètre de la distribution qui est essentiel pour comprendre les valeurs données par la moyenne et les taux.

La variance des notes sera calculée pour chaque magasin (variance-intra), et elle sera l'indice privilégié pour tester l'hypothèse P2

II – Organisation des données pour le test de P2

Pour chaque magasin considéré (221 magasins), nous avons fait un classement suivant les différents niveaux de satisfaction obtenus par la moyenne et les taux. Puis nous avons calculé la correspondance par le coefficient de corrélation des rangs¹¹¹. Le but dans une première étape est de voir si le classement est identique. Un classement identique signifiera que les deux formes d'indicateurs (moyenne ou taux de satisfaits) apportent la même information; ici c'est la situation attendue. Seulement, dans le cas contraire, si les deux indicateurs fournissent des classements différents, la situation devient problématique suivant l'hypothèse de l'objectivité de l'indicateur.

¹¹⁰ ROUANET Henry, LEPINE Dominique : Problème de méthodologie statistique (1). Introduction à l'étude de la robustesse des méthodes usuelles d'inférence statistique. *Mathématiques et Sciences humaines* tome 46, (1974), p21-33.

¹¹¹ Coefficient de corrélation non paramétrique de Kendall, qui traduit la concordance entre deux classements.

Chaque calcul de moyenne et taux par magasin sera accompagné par un calcul de la variance et les mêmes facteurs considérés dans la proposition P1. Les magasins sont situés en France métropolitaine et en Allemagne.

Lors du classement des magasins par moyenne et par taux, le taux de concordance des deux classements est de **76%**. Cela signifie aussi qu'il existe 24% des paires qui sont discordantes: 24% des comparaisons deux à deux ont subi des classements différents selon la moyenne et le taux de satisfaction. Comme exemple, nous avons pris le cas de deux magasins classés différemment selon les deux types d'indicateurs. Les résultats sont représentés dans le tableau suivant

Tableau 14 Classement par taux ou moyenne exemple A

Magasin	Moyenne	Taux	variance	effectif
M1	8.11	93%	1.40	1 260
M2	8.13	89%	1.97	929

Nous remarquons que le premier magasin à une note moyenne légèrement inférieur au deuxième, mais il présente un taux de satisfaction nettement supérieur. Ce paradoxe confirme la présence d'une influence autre que celle de l'effort marketing qui influe sur l'indicateur calculé sous ces deux formes : moyenne ou taux. Cependant en regardant les variances de deux magasins, le premier possède une variance moins élevée que le deuxième. Comme les deux moyennes sont proches, la variance joue un rôle de contraste qui s'est manifesté par un taux de satisfaction qui donne une information contradictoire par rapport à la moyenne. De plus, en regardant les effectifs par magasin nous pourrions penser que la variance est petite car l'effectif est grand. Néanmoins la situation inverse peut aussi être vérifiée comme le confirme le tableau suivant:

Tableau 15 Classement par taux ou moyenne exemple B

Magasin	Moyenne	Taux	variance	effectif
M1	8.29	91%	1.68	951
M2	8.28	92%	1.85	1152

Suivant ces constats, il paraît que la comparaison est « faussée » par le fait que les deux variances sont différentes. Il est clair que cette erreur cause un problème dans les interprétations des résultats par moyennes ou par rapport aux taux des satisfaits. Ce cas engendre des problèmes de décision. Ainsi les indicateurs selon le type de calcul : par

moyenne ou par taux, sont sensibles à l'effet de la variance. Reste à étudier si cette différence des variances est juste accessoire (erreurs d'échantillonnage) ou si elle a des sources extérieures.

Pour les tests de la proposition P2, la variance intra magasin, est devenu un paramètre d'intérêt car l'explication des erreurs de classement par moyenne ou par taux vient de ses effets. C'est pourquoi, comme le test de l'hypothèse P1, nous effectuerons l'analyse supplémentaire en relation avec d'autres informations relatives aux individus sondés (l'âge, le sexe, la fréquentation du magasin) et à leur environnement (la région ou le pays). Le but est de vérifier si la fluctuation de la variance intra magasin est expliquée par la présence de ces facteurs de contingence.

La distribution des variances récoltées pour chacun des 221 magasins est résumée par la table et le graphique suivants:

Min.	1st Qu.	Médiane	Moyenne	3rd Qu	Max	Ecart-type	Coefficient de variation
0.76	2.07	2.43	2.45	2.84	4.29	0.60	24.44%

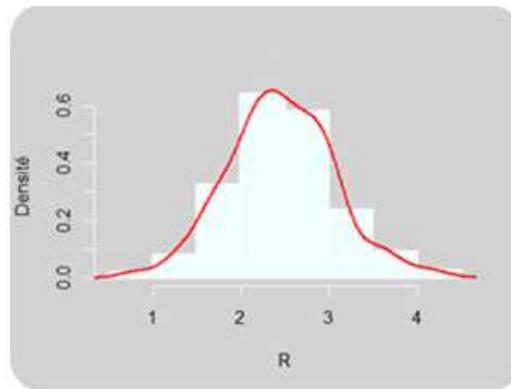


Figure 28 : Dispersion des valeurs des variances intra-magasin sur l'ensemble des 221 magasins

Le dernier tableau nous montre que les variances varient par magasin (figure). Avec un écart-type de 0,6. Le coefficient de variation des variances intra-magasins est de l'ordre de 24%. Ce qui à notre sens justifie l'idée que les variances dans chaque magasin sont différentes les unes des autres. Ce résultat montre aussi que la distribution des notes de

satisfaction paraît fluctuer selon le magasin. Cette variation peut engendrer des inexactitudes dans les informations recueillies lors du calcul par moyenne ou par taux. Reste à savoir si cette différence est fortuite (erreur aléatoire), ou si elle est causée par la présence des facteurs contingents (erreur systématique).

A titre illustratif, nous avons représenté géographiquement les variances intra-magasin (figure 29) et comme pour la proposition P1, une dépendance géographique est apparente. Ceci encourage l'explication de la fluctuation des variances intra magasin par les facteurs de contingence.

Dans la suite, les variances intra magasin (variance calculée à partir des individus restreints à leurs magasins) sera la variable d'intérêt (la variable à expliquer). Cette variable d'intérêt représentera la quantité qui testera la proposition P2. Elle sera étudiée à la lumière des facteurs de contingence (identique à ceux utilisés pour la proposition P1), afin de statuer sur la nature de ses fluctuations (variation fortuite ou de sources contingente)

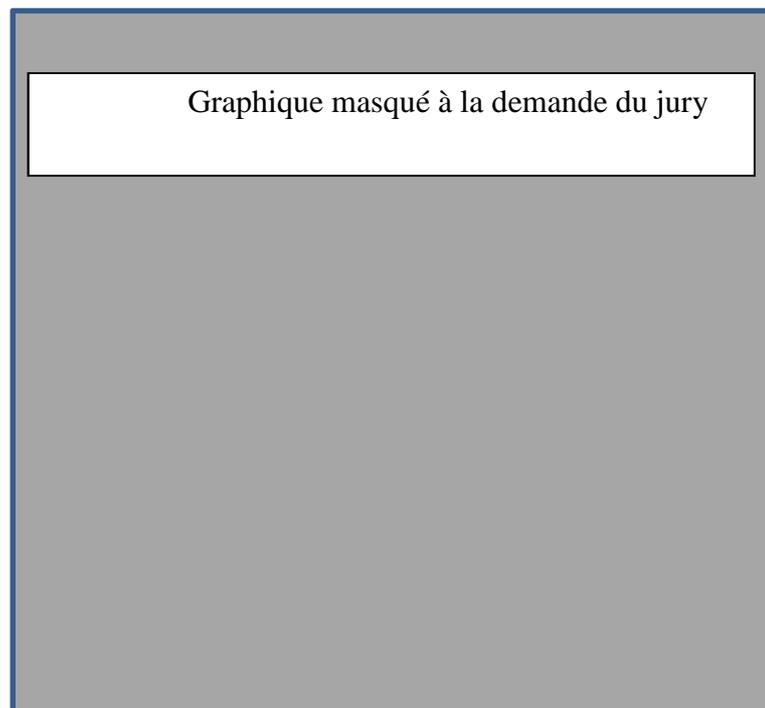


Figure 29 : Répartition géographique des variances selon les magasins

III – Effets des facteurs de contingence sur la fluctuation de la variance intra magasin.

Toujours en suivant l'analyse avec un esprit méta-analyse, la variance intra-magasin sera expliquée à partir des données agrégées par magasin. Dans le même principe, les données vont être étudiées à la lumière des facteurs de contingence agrégés au niveau magasin. Le tout en tenant compte des effectifs dans l'analyse de comparaisons bayésienne. Le but est de savoir si la fluctuation de la variance peut être attachée aux facteurs de contingence, ce qui permettra de statuer sur la proposition P2. Nous voulons savoir si les différences détectées par l'indicateur sont exclusivement dues aux efforts marketing dans chaque magasin ou si elles sont produites par d'autres sources. Les résultats des analyses via les comparaisons bayésiennes, le cas de l'**Anova**, nous donnent les conclusions suivantes :

- Effets de l'âge des répondants sur la variance

Tableau 16 : Effets de l'âge des répondants sur la variance

Variance intra magasin	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Facteur âge	1	0.015187	0.015187	23.2548	1.656E-06
Residuals	219	0.143022	0.000653068		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
9,60%	0	85 %	0.45

Les résultats de l'analyse nous montrent que l'âge des répondants a un effet significatif ($p_value=0,00$) sur la fluctuation de la variance par magasin. Ce qui signifie que la structure de la population selon les âges induit une différence dans la variance, donc la structure des distributions des notes de la satisfaction relative à chaque magasin. Cet effet est moyennement notable (9,6%) avec une confiance de 85%. Ainsi l'âge des répondants peut être considéré comme ayant un effet non négligeable sur la fluctuation des variances intra-magasins. Le déséquilibre dans l'âge des répondants par magasin, influe sur l'étendue des notes de satisfaction exprimées. Autrement dit, cet effet joue un rôle modulateur de la forme des distributions des notes de satisfaction. Ce qui peut induire une

différence dans l'information apportée par le calcul de la moyenne et le taux de satisfaction.

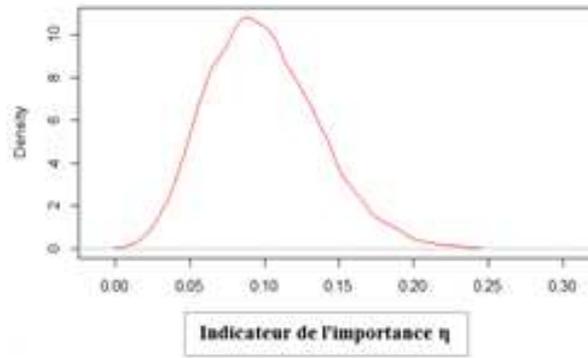


Figure 30 : Distribution des valeurs de l'importance de l'effet de l'âge sur la variance intra magasin

➤ Effets du sexe des répondants sur la variance.

Tableau 17: Effets du sexe des répondants sur la variance

Variance intra magasin	Degré de liberté	de somme des carrés	somme des carrés Moyen	F	P value
Facteur SEXE	1	0.00001532	0.00001532	0.0212	0.8843
Residuals	219	0.158193	0.000722342		

Importance de l'effet	Probabilité <1%	Probabilité >6%	Probabilité >10%
0.009%	--	--	--

Contrairement à l'âge, les résultats de l'analyse nous montrent que le sexe des répondants n'a pas d'effet significatif (p_value de 0,88). De ce fait, ce dernier peut être considéré comme négligeable étant donné son importance très faible (0,009%)

➤ Effets de la fréquentation des répondants sur la variance

Tableau 18: Effets de la fréquentation des répondants sur la variance

Variance intra magasin	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Fréquentation	1	0.0072303	0.0072303	10.488	0.001388
Residuals	219	0.15098	0.000689406		

Importance de l'effet	Probabilité <1%	Probabilité >6%	Probabilité >10%
4.57%	4.20%	0.3	4.70%

Les résultats nous indiquent qu'il existe un effet significatif ($p_value=0,001$) de la fréquentation sur la fluctuation de la variance intra-magasin. Cet effet n'est pas important (4,5%), cependant on ne peut pas dire qu'il est négligeable, car nous avons 95,8% de chance qu'il soit supérieur au seuil de 1%. Ainsi l'effet de la fréquentation sera considéré comme faible mais existant. Ce constat traduit le fait que plus les clients sont hétérogènes en termes de fréquentation (anciens clients et nouveaux clients) plus la variance varie. Donc comme constaté précédemment pour l'âge, la variable fréquentation joue un rôle sur la fluctuation des variances, de ce fait les résultats des indicateurs par moyennes et par taux deviennent sensibles aux types des clients en termes de fréquentation.

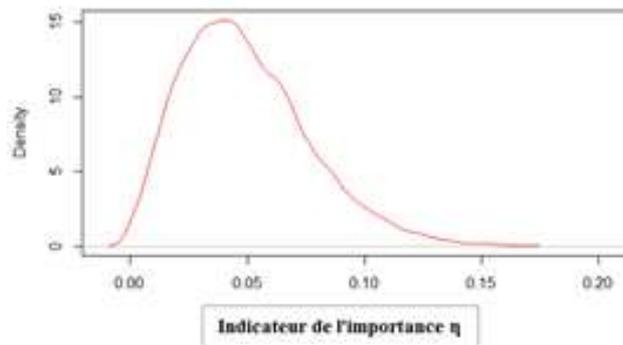


Figure 31 : Distribution des valeurs de l'importance de l'effet de la fréquentation sur la variance intra magasin

- Effets de la région et du pays des répondants sur la variance

Comme pour les analyses précédentes concernant P1 sur le pays et la région, l'analyse sera effectuée en tenant compte de la structure hiérarchique.

➤ Pour l'effet Pays

Tableau 19: Effets du pays des répondants sur la variance

Variance intra magasin	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Erreur Région					
Pays	1	13.68	13.68	40.4538	6.49E-08
Residuals	49	16.57	0.3382		
Erreur within					
Residuals	170	50.14	0.295		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
45.20%	0.00%	100.00%	100.00%

L'effet du pays par rapport à la région est significatif ($p_value=0,00$), cela induit qu'il a un effet sur la fluctuation des variances intra-magasins. De plus comme l'affirme les résultats du tableau ci-dessus: l'effet est important avec une confiance de 100%. Le facteur pays semble donc jouer fortement sur les différences entre les variances intra-magasin. Ainsi selon le pays (ici la France ou l'Allemagne), les distributions des notes de satisfaction sont plus ou moins étalées, ce qui implique, comme pour les facteurs précédents (âge et la fréquentation), que les classements peuvent être différents selon l'indicateur choisi (moyenne ou taux).

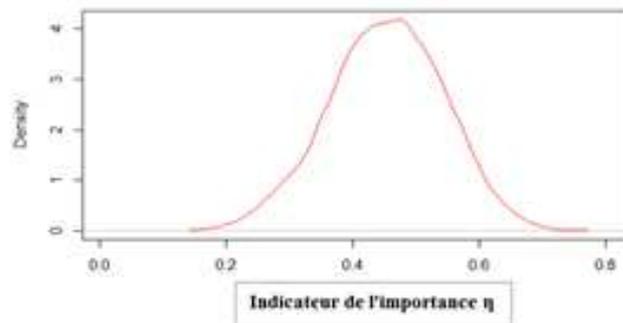


Figure 32 : Distribution des valeurs de l'importance de l'effet du Pays sur la variance intra magasin

✦ Pour l'effet la région

Tableau 20: Effets de la région et du pays des répondants sur la variance

Variance intra magasin	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Pays	1	13.68	13.68	46.3821	1.6E-10
Pays: Région	49	16.57	0.3382	1.1465	0.26
Residuals	170	50.14	0.295		

Importance de l'effet η	Probabilité <1%	Probabilité >6%	Probabilité >10%
24.84%	-	-	-

L'effet de la région est non significatif ($p_value=0,26$), ce qui empêche de généraliser l'importance au niveau inductif. Néanmoins, l'effet de la région est important d'un point descriptif. 24,8% de la variance est expliquée par ce dernier. Cependant, comme pour les résultats sur le coefficient de corrélation multiple, le nombre de régions considéré par rapport au nombre de magasins, fait que les résultats soient non significatifs et nous ne pouvons pas accorder une confiance à cette importance.

IV – En résumé

Dans les deux cas d'étude, les facteurs géographiques comme la région et le pays, ont des effets notables sur la différence des variances intra-magasins. Les variables « fréquentation des clients et l'âge des répondants » peuvent elles aussi moduler cette variation. Ceci induit que cet ensemble de facteurs explique les différences dans les classements suivant l'indicateur considéré comme résultat de moyenne des notes ou comme résultats d'un taux de satisfaction. Ainsi une des causes qui font que le type de calcul : moyennes ou taux donne des classifications différentes, sont les facteurs liées aux profils des individus. Il suffit que les profils des individus sondés soient très différents entre magasins, pour que l'indicateur soit incapable de distinguer réellement les différences dues au management à celles liées à des causes extérieures.

Ces résultats constituent le contre-exemple qui prouve que les disparités des variances ne sont pas dues à des erreurs aléatoires (fluctuation de l'échantillon de sondage). Ceci met donc à mal la proposition P2, car les variances modulent le calcul de l'indicateur, moyenne ou taux. Donc un contraste dans les variances causées par les

facteurs contingents implique l'incapacité de l'indicateur à remplir son rôle de contrôle : différences des niveaux de satisfaction entre magasins, par moyenne ou taux, non essentiellement attachées au seul fait du management. Par conséquent l'évaluation statistique de l'indicateur de satisfaction ne paraît pas totalement objectif, et la comparaison transversale subit ce constat lors de l'utilisation de la moyenne ou du taux de satisfaits.

En conclusion, la comparaison directe par moyenne ou taux subit l'effet des facteurs de contingences, dont certains sont attachés aux caractéristiques individuelles et à leur environnement, ce qui cause des problèmes de légitimité à l'indicateur de satisfaction dans sa fonction de contrôle. En somme, l'indicateur de satisfaction est sensible à la forme du calcul utilisé (moyenne ou taux). Une sensibilité qui engendre des biais de décision.

Section 3 – **Procédure et résultats du test de la proposition P3**

L'indicateur de satisfaction est par essence une quantité non stable car il doit refléter le concept de satisfaction qui par définition évolue dans le temps. Pour qu'un indicateur puisse remplir sa fonction il doit donc évoluer dans le temps. Dans l'absolu, ses variations reflètent les actions marketing qui sont modifiées au cours des périodes. Dans la pratique, ces actions sont en mutation permanente visant à accompagner le contrôle de la performance. L'indicateur doit donc à tout prix refléter ce changement, une stabilité inhibera toute prise de décision. Si une moyenne ou un taux de satisfaction est stable, cela veut dire que les actions marketing précédentes (à t-1) sont non concluantes, ce qui peut fausser l'appréciation future des nouvelles actions.

La procédure de test de la proposition P3, vise à étudier la sensibilité des indicateurs dans le temps. L'indicateur de satisfaction aura comme support une question de satisfaction d'un attribut, car c'est sur ces quantités que le suivi temporel de la performance est basé. Les managers veulent souvent savoir si une amélioration a été apportée sur un attribut (ex : l'accueil, la propreté) par rapport aux vagues précédentes. Ce suivi peut être vu au global (ensemble de leur parc) ou par unité magasin. Notre but est d'apprécier l'évolution des indicateurs sur ces deux niveaux : global (sans distinguer les magasins) ou par magasin afin de répondre aux préoccupations formulées dans la proposition P3.

L'idée poursuivie dans ce test est d'examiner l'évolution des indicateurs en détachant les deux niveaux (ensemble du parc / par magasin). L'intérêt est de tester la sensibilité qui doit être semblable dans les deux plans, selon l'hypothèse que l'indicateur doit refléter les avis clients peu importe le niveau de calcul. Le test vise à étudier la vraisemblance de la proposition P3. Il sera effectué à partir d'un échantillon de magasins suivant un attribut de la satisfaction. L'idée est de vérifier si l'indicateur est sensible sur les deux niveaux (global et par magasin) afin de s'assurer de la bonne démarche du pilotage de la performance.

I – Indice pour juger de la sensibilité dans le temps des indicateurs.

L'objet de l'analyse va être accès sur le test de l'existence de l'invariance dans le temps des indicateurs. Pour prouver la stabilité/ou instabilité des indicateurs, nous nous sommes muni d'un baromètre concernant une grande enseigne de Bricolage. L'exploitation de ce baromètre pour prouver l'invariance dans la mesure d'une période à une autre, a été effectuée à l'aide d'une technique statistique dite de « calcul de distorsion ». Il s'agit d'un outil statistique descriptif visant à révéler si les concentrations entre deux distributions de fréquences sont éloignées l'une de l'autre. Cette notion de « distorsion » sera utilisée pour comparer les distributions des notes enregistrées sur un critère donné (l'attribut de satisfaction) à une distribution uniforme. Le fait de comparer les dispersions à l'aide de cette notion de distorsion peut être vue comme une comparaison sur toute la famille des indicateurs (moyenne, variance, taux etc.) d'où l'intérêt de cet indice.

Nous définirons l'indice de distorsion (Térouanne, 1995)¹¹², entre deux distributions de fréquences p , et q , par :

$$\phi(p, q) = \sum_{x,y} \phi_{x,y}(p, q)$$

Avec $\phi_{x,y}(p, q) = p(x)q(y) - q(x)p(y)$

Cet indice de distorsion $\phi(p, q)$ est compris entre 0 et 1. Si l'indice est égale à 0, nous nous trouvons dans le cas où les distributions de fréquences sont égales $p=q$, et s'il est égal à 1, les deux distributions sont totalement différentes.

¹¹² La Distorsion est un indice qui peut être vu comme une mesure de liaison (dépendance) entre variables nominales. Dans notre cas, il est utilisé comme une mesure de degré de ressemblance entre deux distributions de probabilité.

Comme la problématique de notre étude, est de prouver que les distributions d'une vague à une autre, sont identiques, nous prendrons l'indicateur $\lambda(p, q)$ défini par : $\lambda(p, q) = 1 - \phi(p, q)$. Ce qui nous permettra de traduire la ressemblance entre deux distributions avec une valeur proche de un et proche de zéro dans le cas d'une dissemblance

Cet indice sera calculé relativement à la distribution uniforme¹¹³ Il concernera chaque vague pour chaque magasin et au global (l'ensemble des magasins). Il sera noté λ_p (indice de distorsion relative) et sera utilisé pour juger de la sensibilité des indicateurs. L'idée de prendre la loi uniforme comme repère, permettra d'avoir des indices semblables par vague et comparable entre eux.

II – Organisation des données pour le test de P3

Pour chacun des 97 magasins considérés, nous avons calculé l'indice de distorsion relative λ_p , pour chacune des 16 vagues (4 trimestres et 4 années). Puis nous avons procédé de la même manière sur l'agrégat des magasins.¹¹⁴ La procédure de calcul d'indice concerne **un attribut de satisfaction** sur lequel l'avis exprimé est indépendant du choix du magasin. Pour ce faire, nous avons concentré notre analyse sur un attribut qui concerne la satisfaction envers la qualité des produits fournis par l'enseigne. Pour nous, la qualité du produit vendu dans les magasins est identique et ne dépend en aucun cas du contexte du magasin (à la différence de la propreté du magasin par exemple).

Les magasins concernés sont situés en France métropolitaine et sont répartis sur plusieurs régions (figure). Pour le Baromètre B2, il s'agit d'une compilation d'études sur 16 vagues trimestrielles sur 4 ans. Pour chaque vague 40000 individus sont interrogés. L'enquête est menée sur 97 magasins dans lesquels plus de 400 clients sont interrogés. Les informations disponibles sont agrégées par magasin, région et au niveau de la France. Nous disposons donc de 97 magasins répartis en France. Leur répartition géographique est visible dans la carte suivante :

¹¹³ Toutes les éventualités ont la même probabilité.

¹¹⁴ Nous nous entendons par agrégat des magasins, le cumul des données sans distinction des magasins.

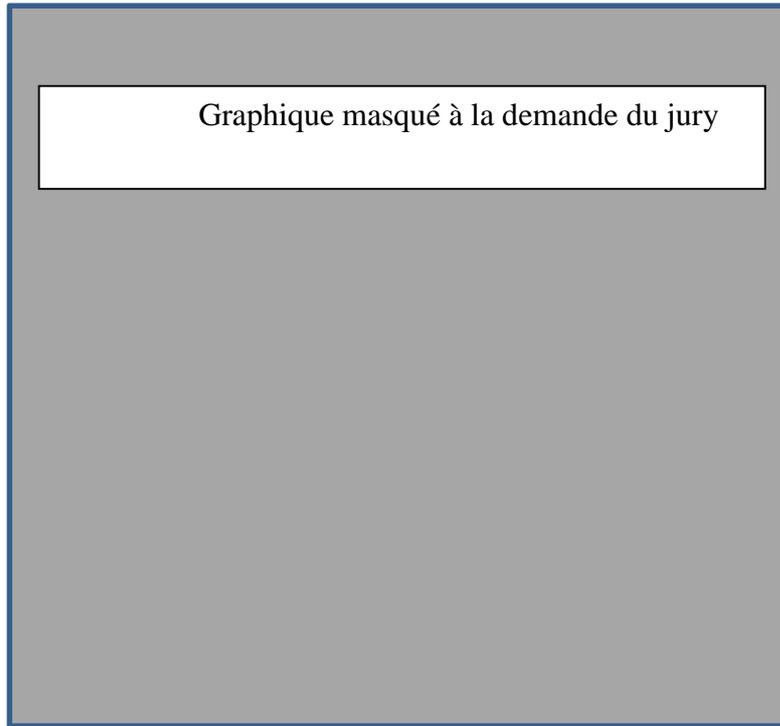


Figure 33 : Répartition géographique des magasins du baromètre B2

La dispersion de l'indice λ_p (correspondant à la distorsion entre la distribution des notes de satisfaction et la loi uniforme calculée sur chacune des 16 vagues) est résumée comme suit :

➤ Au niveau global (sans distinguer les magasins)

Min.	1st Qu.	Médiane	Moyenne	3rd Qu	Max	Coefficient de variation
0.377	0.3804	0.3816	0.382	0.3843	0.3865	0.73%

Les résultats nous indiquent que sur les 16 vagues, les indices de distorsion relative à la loi uniforme λ_p bougent peu. Un coefficient de variation de 0,73% indique une faible variation autour de la moyenne. Ce qui veut dire que d'une vague à une autre la forme des distributions des notes de satisfaction reste similaire. Ce qui induit aussi que les indicateurs de satisfaction (moyenne ou taux) sont sensiblement identiques d'une vague à une autre.

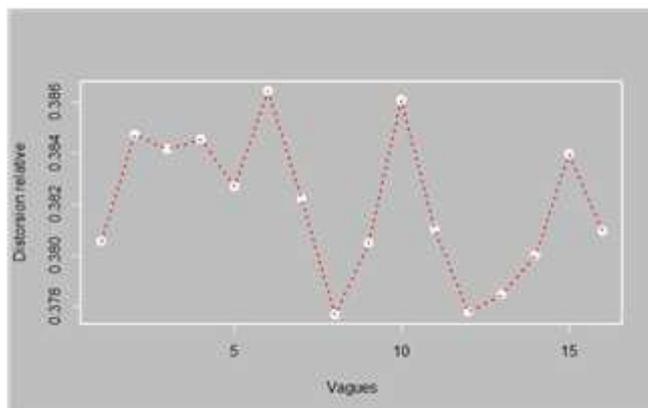


Figure 34 : Évolution des indices de distorsion relative au niveau agrégé (ensemble du parc) suivant les 16 vagues

Cependant en regardant le graphique des 16 valeurs des indices (figure 34). Nous constatons que même si l'étendue des valeurs est limitée (stabilité), il semble avoir une saisonnalité. Un test de saisonnalité basé sur une **Anova**, pour voir s'il existe des différences systématiques entre les trimestres (ce qui prouve une saisonnalité), nous donne les résultats suivants :

Tableau 21 : Un test de saisonnalité des λ_p

Indices de distorsion relative	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Trimestres	3	4.5175E-05	1.50583E-05	2.455	0.1133
Residuals	12	7.3599E-05	6.13325E-06		

Les différences entre les indices de distorsion relative λ_p , entre les trimestres n'est pas significatif (p_value 0,11). Néanmoins l'importance de cet effet (η) est de 38%. Ce qui prouve une différence importante qui peut être interprétée comme une saisonnalité par trimestre. La non significativité est surtout liée à la rareté des données seulement 16 observations.

L'analyse au niveau global (sans distinguer les magasins) a relevé une tendance de stabilité dans le temps avec une petite fluctuation qui peut être attachée à une forme de saisonnalité. Ainsi au niveau global les indicateurs de satisfaction ne semblent pas être sensibles à un effort marketing qui lui varie fortement d'une vague à une autre.

➤ Au niveau magasin

Par rapport au niveau agrégé, les variations des indices de distorsion relative λ_p sont plus sensibles sur les 16 vagues. Pour révéler les variations des indices nous avons calculé pour chaque magasin leur étendue λ_p . Cette étendue de variation a été captée par le coefficient de variation¹¹⁵. Les résultats du calcul pour chaque magasin de ces coefficients de variation sont représentés par le tableau et le graphique suivant :

Min.	1st Qu.	Médiane	Moyenne	3rd Qu	Max
2.42%	3.61%	4.28%	4.68%	5.63%	9.58%

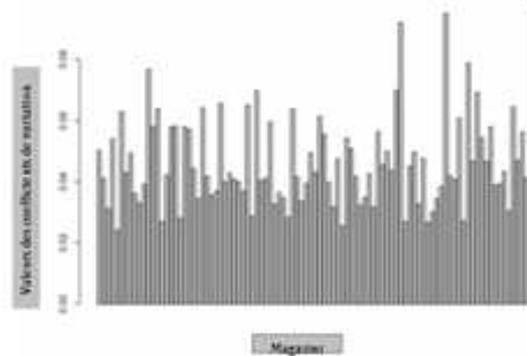


Figure 35 : Valeurs des coefficients de variation des indices de distorsion relative, sur les 16 vagues pour chaque magasin

Les résultats contrairement au niveau agrégé, montrent que la distorsion relative des notes de satisfaction dans chaque magasin, varie entre les vagues. Là où pour le niveau global le coefficient de variation (des indices de distorsion relative λ_p sur les 16 vagues) est de 0.73%, la variation par magasin oscille entre 2.42% et 9.58%. Ce fait nous indique qu'un indicateur de satisfaction est plus sensible au temps au niveau des magasins et le devient moins au niveau agrégé.

Pour le niveau agrégé (ensemble des magasins), nous avons constaté une saisonnalité liée aux trimestres, reste à savoir si ce fait est valable dans chaque magasin et aussi savoir ce qui pourrait expliquer cette saisonnalité. Dans notre théorie de facteurs de

¹¹⁵ Coefficient de variation : écart-type/moyenne : $c_v = \frac{\sigma}{\mu}$

contingence nous avons postulé qu'il existe des facteurs notamment liés à l'environnement des individus qui peuvent perturber la mesure et créer ainsi ce que l'on a appelé l'effet de « la structure hiérarchique ». Cette structure se traduit par le fait qu'un individu niché dans un magasin qui appartient à une région qui elle-même est incluse dans un pays, induit une structure qui conditionne les individus à une certaine homogénéité qui se traduit dans notre cas par une distribution des notes de satisfaction, similaire d'une vague à une autre.

Suivant ce raisonnement, l'analyse qui va suivre servira à voir si les différences d'une vague à une autre, au niveau magasin, sont significatives. L'analyse sera aussi accompagnée par une autre s'intéressant aux saisonnalités et à des facteurs de contingence relative à l'environnement comme la région. Le but poursuivi ici est de tester la sensibilité de l'indicateur par magasin et de vérifier l'existence d'une telle structure qui provoque une stabilité au niveau agrégé (ensemble des magasins). Nous finirons notre démonstration par un test relatif à cette structure hiérarchique qui crée une invariance des indicateur au niveau global, conditionnent les magasins dans des niveaux.

III – Sensibilité dans le temps et facteurs de contingence

Cet examen a pour but de qualifier l'importance des différences dans les indices de distorsion relative λ_p entre les vagues. Ce calcul sera effectué en distinguant les niveaux magasins et en présence du facteur région. L'analyse conjointe de toutes ces informations vise à associer les vrais effets à chaque source (vague, région) en séparant leurs effets conjoints. Il s'agit d'un plan de type **Magasin*Vague<Région>**. C'est pour cette raison que l'effet de la vague ne peut être analysé que sachant la région et vice versa. Chaque facteur a une source de variation différente. Pour la région, c'est la variation inter magasins (Error magasin) et pour la vague la source de variation est l'interaction magasin : vague (Error: Magasin: vague). Dans ce cas, on analyse un plan de données appareillées, les données concernent les mêmes magasins sur plusieurs vagues. Par conséquent l'étude de l'effet de la vague sera rapporté à la source : interaction magasin et vague, et l'analyse de l'effet région sera rapporté à la variance entre magasins. Les résultats de cette décomposition sont les suivants :

Tableau 22 : Effet de la région sur la sensibilité de λ_p

Indices de distorsion relative	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Erreur Magasin					
Région	8	0.1008	0.0126	5.7008	7.68E-06
Residuals	88	0.1945	0.0022		
Erreur Magasin :Vague					
Vague	15	0.0097	0.000646667	1.9165	0.0183
Région: vague	120	0.0314	0.0003	0.7755	0.96
Residuals	1320	0.4454	0.0003		

L'effet de la région sur les différences de l'indice de distorsion relative λ_p est significatif ($p_value=0,00$). L'importance de cet effet est de l'ordre de 34,13% avec une confiance de 100% qu'elle soit supérieure à 10%. Ce qui lui accorde le statut d'effet important. Pour l'effet de la vague (différence entre vague), il est significatif ($p_value= 0,01$), son importance est de 2,12%, avec une confiance de 99% qu'elle soit supérieur à 1%, ce qui lui donne le statut d'un effet non négligeable. Pour ce qui est l'effet de la région sur la vague, l'interaction reste non significative ($p_value 0,96$) mais avec une importance de 6,6%. Cela signifie qu'au niveau descriptif, la variation de l'indice λ_p selon la région diffère suivant les vagues avec une importance moyenne, sans qu'elle soit confirmée au niveau inductif.

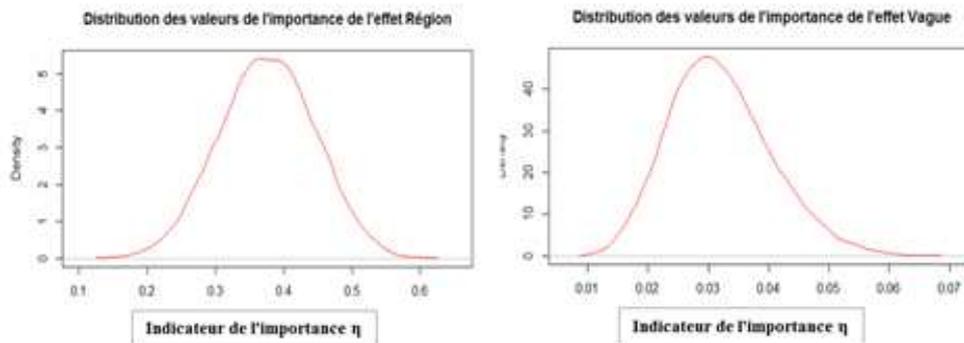


Figure 36 : Importance des effets région (à gauche) et vague (à droite)

Les deux graphiques (figure 36) indiquent que l'effet de la région est plus important que celui de la vague d'enquête. Ce qui laisse entendre que la sensibilité de l'indicateur est plus liée à une source systématique entre régions que celle liée à la vague de sondage

De ces résultats nous tirons les conclusions qu'il existe bien une différence dans les notes de satisfaction par magasin suivant les vagues, ce qui engendre une meilleure sensibilité de l'indicateur de satisfaction par magasin. L'existence de l'effet région nous indique que dans ces variations il existe un caractère systématique entre régions. De plus ces différences systématiques peuvent dépendre de la vague (effet d'interaction). En somme, il existe bien une sensibilité des indicateurs au niveau magasin, mais cette sensibilité dépend fortement de la localisation géographique du magasin (ici la région). Toutefois reste à connaître la nature de cette dépendance.

IV – Saisonnalité dans le temps et facteurs de contingence

Suivant les résultats précédents, dépendance vague/région, avec une importance de 6,6%, nous avons effectué le même calcul statistique mais en considérant cette fois les trimestres (4 modalités). Les conclusions de l'analyse sont les suivants :

Tableau 23 : Effet de la région sur la sensibilité de λ_p

Indices de distorsion relative	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Erreur Magasin					
Région	8	0.1008	0.0126	5.7008	7.68E-06
Residuals	88	0.1945	0.0022		
Erreur Magasin :Vague					
Trimestre	3	0.003	0.0010	2.9636	0.0317
Vague	12	0.0067	0.00056	1.6547	0.071
Trimestre: vague	24	0.0061	0.000254167	0.7533	0.79
Région: vague	96	0.0253	0.000264	0.7810	0.93
Residuals	1320	0.4454	0.00034		

En étudiant, les variations des trimestres, l'effet de la vague devient non significatif ($p_value=0,07$) avec une importance qui passe de 2,17 % à 1,48%. Cependant, le facteur trimestre (saisonnalité) quant à lui est significatif ($p_value 0,03$) avec une importance de 0,66%. Donc la variation de l'indice de distorsion relative, est attachée à la fois à la vague

et aux trimestres. L'effet est faible pour le trimestre mais il reste influent, l'effet de la vague perd de son importance mais reste plus important que celui du trimestre. Donc leurs effets sont présents sans pour autant que cela soit vraiment notable.

Pour l'interaction région par trimestre et par vague, ils sont tous deux non significatifs avec une différence dans l'importance concernant les vagues (5,37%) par rapport aux trimestres (1,4%). Ainsi aucune dépendance entre les régions et la vague ne peut être confirmée. Il semble aussi qu'il existe une tendance de saisonnalité trimestrielle comme pour le niveau agrégé. De plus l'effet région est important au niveau des magasins ce qui conforte l'existence d'un effet de structure hiérarchique qui agit au niveau de chaque vague. Cette structure conditionne les résultats à une certaine homogénéité au niveau agrégé (sans distinguer les magasins), notamment la stabilité d'une vague à une autre.

Une analyse supplémentaire, liant l'effet de la région sur les différences dans l'indice de distorsion relative λ_p , par trimestres¹¹⁶ nous fournit les résultats suivants :

Indices de distorsion relative/ trimestres 1	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Région	8	0.02423	0.00302875	6.358	9.48E-08
Residuals	379	0.18055	0.000476385		

Indices de distorsion relative/ trimestres 2	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Région	8	0.01988	0.002485	5.834	4.87E-07
Residuals	379	0.16144	0.000425963		

Indices de distorsion relative/ trimestres 3	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Région	8	0.02831	0.00353875	7.795	1.06E-09
Residuals	379	0.17205	0.000453958		

Indices de distorsion relative/ trimestres 4	Degré de liberté	somme des carrés	somme des carrés Moyen	F	P value
Région	8	0.03454	0.0043175	10.362	3.84E-13
Residuals	379	0.15791	0.000416649		

¹¹⁶ La considération trimestre, renvoie à la restriction aux trimestres 1, sur les 16 vagues soit 4 trimestres (97 magasins fois 4. Soit 388 données) et à la restriction aux trimestres 2 pour les 16 vague etc...

Les importances des effets de la région, respectivement pour les 4 trimestres sont de: 11,8%, 10,9%, 14,1% et 17,9%. Ils sont tous importants et ont sensiblement les mêmes grandeurs. Ces résultats appuient un effet de structure: différence systématique entre régions qui semble avoir la même ampleur dans chaque trimestre. Ce fait a été dégagé dans l'analyse précédente en apportant une absence d'effet important de l'interaction région et trimestre.

V –En résumé

Les résultats nous montrent qu'il existe une saisonnalité significative, mais relativement faible. Les effets de la région (sensiblement identique par trimestres) se sont révélés importants. La saisonnalité indique aussi que suivant chaque trimestre des profils de clients lui sont spécifiques. Nous pensons que cela est spécifique à l'activité étudiée (le Bricolage). Suivant la période, les travaux de bricolage ne sont pas les mêmes et les clients n'ont pas le même comportement suivant l'importance du produit. Pour la stabilité elle est moins marquée au niveau magasin par rapport au niveau global. De ce fait, l'indicateur de satisfaction subit des phénomènes qui le conditionnent à être plus ou moins sensibles suivant le niveau de l'analyse : par magasin ou sans distinction des magasins. Ainsi la proposition P3 est elle aussi mise à défaut dans cette étude de cas, et les indicateurs subissent les effets de structure ce qui met à mal leurs fonctions de contraste dans le temps

Section 4 – Existence d'une structure hiérarchique

Les conclusions apportées par les analyses précédentes peuvent se résumer par :

- Les indicateurs sont plus sensibles au niveau magasin qu'au niveau agrégé (sans distinguer les magasins)
- Les indicateurs présentent des différences plus marquées entre régions par rapport aux vagues d'enquêtes.

A titre illustratif le graphe suivant (figure 37), représentant l'évolution des scores de satisfaction des magasins sur 16 vagues dans quatre régions. Nous remarquons que la variation des scores de satisfaction par magasin semble différente par région. Il existe donc des tendances d'évolution des scores, conditionnées par région.

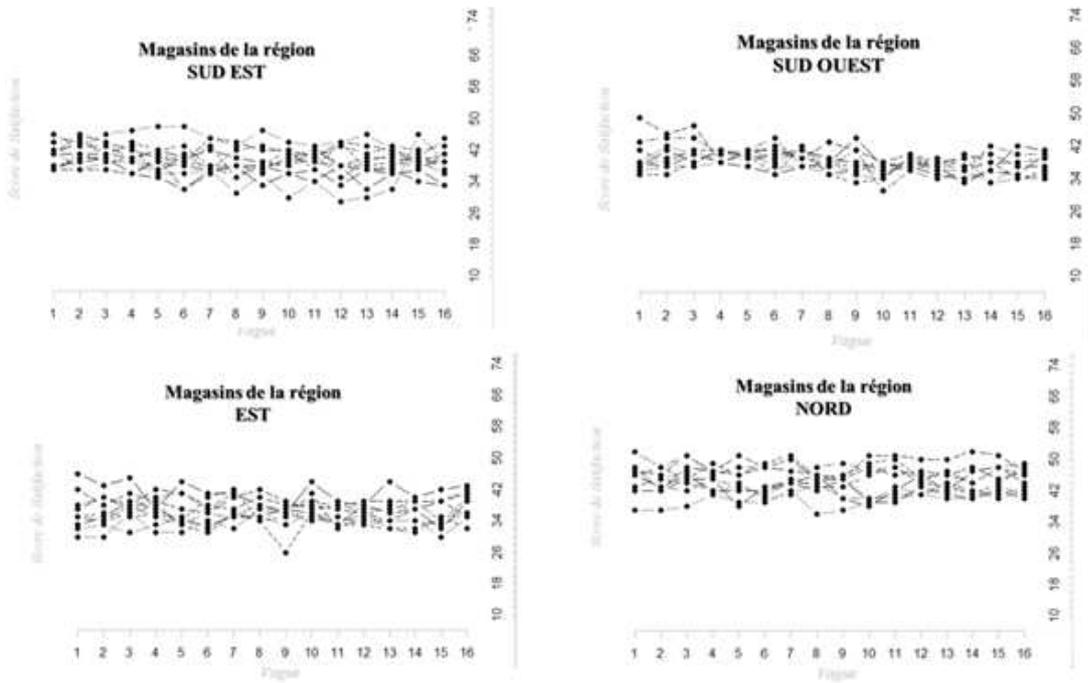


Figure 37 Évolutions des scores de satisfaction enregistrés par magasin

Un autre point à relever (Figure 37) est le parallélisme des courbes dans l'évolution des scores de satisfaction pour les magasins. Ceci signifie que certains magasins sont toujours au « top » et d'autres restent toujours dans le « Bottom ». Aussi les scores de satisfaction enregistrés, même s'ils varient dans le temps, présentent une certaine homogénéité spatiale (i.e. à l'intérieur d'une région). Une explication naturelle, pourrait être liée aux managements mêmes des magasins qui sont différents par nature. Ce qui induit une variance naturelle entre magasins. Toutefois, étant données les résultats avancés dans les analyses précédentes, la région semble aussi jouer sur les différences entre indicateurs. Ainsi l'existence d'une structure hiérarchique passe par l'existence d'une variation inter-magasin importante, qui est subordonnée à une variation inter-région. C'est ce fait-là qui va être étudié dans cette partie.

Par ces constatations, nous allons construire un modèle statistique, qui traduit cette structure. Le but recherché est d'étudier si une telle structure est vraisemblable. Le modèle sera construit étape par étape, (niveau par niveau) et aidera à évaluer l'étendue de la variation à chaque niveau : Intra-magasin : scores par vagues (niveau 1), inter-Magasin (niveau 2), inter Régions (niveau 3) (Schéma).

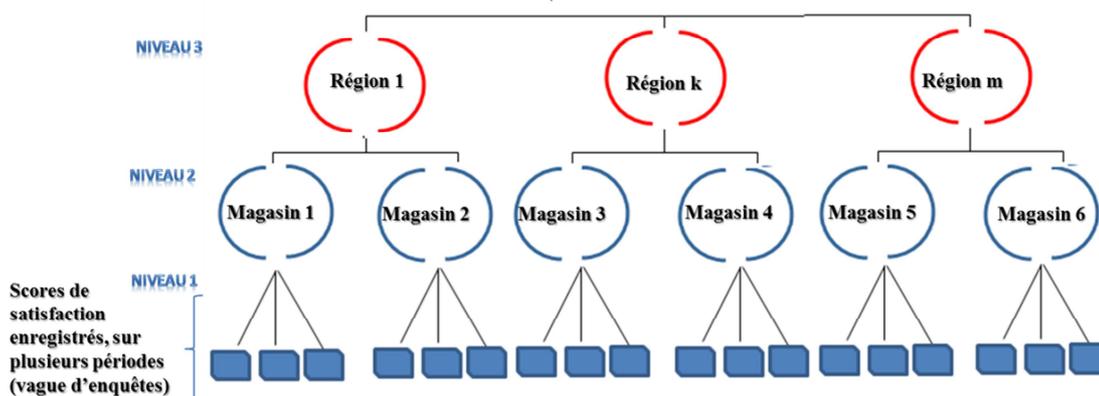


Figure 38 : Structure hiérarchique testée

Les données disponibles sont les scores de satisfaction enregistrés dans chaque magasin pour 19 périodes de sondages. Nous avons 88 magasins¹¹⁷ concernés répartis sur 9 régions de la France métropolitaine. Les données (figure 38), se structurent comme suit :

- Des scores de satisfactions pour 19 vagues, pour chaque magasin.
- Chaque magasin appartient à une région.
- Les scores de satisfaction varient entre 10 et 60. Ils sont calculés à partir des notes récoltés sur l’item « qualité des produits ».

Grouped Data: Score ~ Vague | Région/MAG

Score	Vague	MAG	Région
43	1	3	NORD
46	2	3	NORD
42	3	3	NORD
46	0	3	NORD
42	4	3	NORD
42	5	3	NORD
44	6	3	NORD
42	7	3	NORD
45	8	3	NORD
39	9	3	NORD
39	10	3	NORD
43	11	3	NORD
41	12	3	NORD
40	13	3	NORD
42	14	3	NORD
40	15	3	NORD
36	16	3	NORD
39	17	3	NORD
44	18	3	NORD

Figure 39:Extrait des données d’analyse

¹¹⁷ Il s’agit de mêmes données utilisées pour la proposition P3, toutefois 19 vagues ont été considérées au lieu de 16. Ainsi contenu de la présence des données manquantes, nous avons considéré 88 magasins au lieu de 97.

La structure hiérarchique permettra de dégager la variation des scores à l'intérieur des magasins suivant les vagues (niveau 1), entre les magasins (niveau 2) et entre régions (niveau 3). Pour que chaque niveau existe, il faudra que la variance qu'il produit soit importante.

I – Prise en compte du niveau Magasin

Il s'agit ici de distinguer la variance de ce niveau. Comme les scores concernent plusieurs périodes on s'attend que ces derniers varient dans le temps. Il s'agit de ce qui est communément appelé « mesure répétées ». La considération de ces deux niveaux permettra de décomposer la variance totale suivant la période de mesure et suivant le magasin. En terme d'équation cela se traduirait par :

$$\text{Niveau 1} \quad S_{ti} = M_{0i} + e_{ti}$$

$$\text{Niveau 2} \quad M_{0i} = \gamma_{00} + u_{0i}$$

Les deux équations réunies nous donnent :

$$S_{ti} = \gamma_{00} + u_{0i} + e_{ti}$$

Avec $u_{0i} \sim N(0, \sigma_u)$, $e_{it} \sim N(0, \sigma_e)$ et S_{ti} les scores de satisfaction enregistrés dans chaque vague. L'indice t est relatif au 19 vagues et i est relatif aux 88 magasins.

Les deux quantités σ_u^2 et σ_e^2 correspondent respectivement à : la variance de niveau 1 (variation des scores entre les vagues à l'intérieur des magasins) et à la variance de niveau 2 des scores moyens entre les magasins. La première quantité σ_e^2 nous renseignera sur l'étendue de la sensibilité de scores sur les 19 variables et la deuxième σ_u^2 nous informera sur la différence dans les scores issus des magasins eux-mêmes. La quantité γ_{00} représente la moyenne générale, que l'on lui associe une variance σ_u^2 relative à chaque magasin j (niveau 2) et une variance σ_e^2 associée à chaque score de la vague t (niveau 1 intra-magasin).

L'application du modèle aux données, nous donne les résultats suivant :

Tableau 24: Modèle hiérarchique avec niveau magasin.

Paramètres	Modèle
Effets Fixes	
Constante γ_{00}	38.8343
Effets aléatoires	
Niveau 2 variance inter-magasin : σ_u^2	10.253
Niveau 1 : variance intra-magasin : σ_e^2	5.732
Qualité du modèle	
Deviance :-2 log V	7976
AIC	7982
BIC	7999

Nous remarquerons que la variance inter-magasin σ_u^2 est importante. Ce qui signifie que les scores de satisfactions différent entre vagues avec une étendue (écart-type) de plus de 3 points suivant les magasins ($\sqrt{10.243}$), quand l'étendue de variation des scores entre les vagues, est de 2 points ($\sqrt{5.732}$). Ceci nous indique que les variations des scores de satisfaction se répartissent sur au moins deux niveaux. Pour juger de l'importance de cette décomposition nous utiliserons l'indice de rapport des variances ρ « Rho »¹¹⁸. Il est défini par : $\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$ Dans notre cas, la valeur de ρ est de : 64%. Cet indice pourrait être vu comme le degré de « ressemblance » des scores de satisfaction à l'intérieur de chaque magasin. Nous noterons que cette valeur est importante et confirme l'existence d'une structure hiérarchique (ici deux niveaux). D'autre part, nous remarquons aussi que la variance intra-magasin (entre les scores de satisfaction sur les 19 vagues) est de l'ordre de 34%, avec un écart type de 2 points ($\sqrt{5.732}$). En mettant en parallèle cette valeur à celle calculée au global (sans distinguer les magasins) qui est de 0.56 points, nous remarquons que les scores de satisfaction varient fortement à l'intérieur de chaque magasin d'une vague à une autre, par rapport au niveau global. Ceci confirme que les indicateurs sont nettement plus sensibles au niveau magasin par rapport au niveau global.

¹¹⁸ Le rho est un indice qui mesure la corrélation intra-classe. Il représente le degré de « ressemblance » des individus au sein d'un même groupe.

En somme, le modèle étudié indique que plus la moitié de la variance totale des scores entre les vagues, se situe entre magasins. Afin d'explicitier cette variance, nous introduisant maintenant la région (variable explicative) afin de connaître si cette variance inter-magasin peut être expliquée par cette information. Pour cela notre modèle devient :

$$\text{Niveau 1} \quad S_{tik} = M_{0i} + \gamma_{1k}R_{tik} + e_{tik}$$

$$\text{Niveau 2} \quad M_{0i} = \gamma_{00} + u_{0i}$$

Les deux équations réunis nous donnent :

$$S_{tik} = \gamma_{00} + \gamma_{1k}R_{tik} + u_{0i} + e_{tik}$$

Avec $u_{0i} \sim N(0, \sigma_u^2)$ et $e_{tik} \sim N(0, \sigma_e^2)$. L'indice t est relatif au 19 vagues, i est relatif aux 88 magasins et k indexe les 9 régions étudiées.

L'application du modèle aux données nous donne les résultats suivants :

Tableau 25 : Modèle hiérarchique avec niveau magasin. La région variable explicative.

Paramètres	Modèle
Effets Fixes	
Constante	37.06899
NORD	7.29417
NORD OUEST	4.3718
OUEST	0.68699
PARIS EST	0.4468
PARIS OUEST	0.08412
RHONE ALPES	0.40788
SUD EST	2.07455
SUD OUEST	0.93101
Effets aléatoires	
Niveau 2 variance inter-magasin : σ_u^2	5.2699
Niveau 1 : variance intra-magasin : σ_e^2	5.7343
Qualité du modèle	
Deviance :-2 log V	7901
AIC	7923
BIC	7983

Commençons par analyser les paramètres obtenus pour chaque région. Pour la région « Nord » un paramètre de 7.29 est affiché. Il signifie que par rapport à la constante (score moyen de la région de référence¹¹⁹) il a 7.29 points en plus. Ce qui induit des différences par régions en termes de scores de satisfaction. Notons qu'ici que ces valeurs ne sont qu'à titre illustratif¹²⁰. Il faut les interpréter comme des tendances.

Passons maintenant, aux valeurs des variances inter-magasin σ_u^2 et la variance des scores intra-magasin σ_e^2 , nos paramètres d'intérêt. Nous remarquons que la variance intra reste inchangée : elle passe de 5.732 à 5.7343. Par contre la variance inter σ_u^2 passe de 10.253 à 5.2699. Ce qui fait une variation « Pseudo R² »¹²¹ de 48.6% = $\frac{(10.253-5.2699)}{10.253}$. Ainsi le facteur région explique de façons importantes des différences dans les scores moyens entre les magasins. Toutefois, ce facteur semble ne pas avoir d'effet sur la variance des scores d'une vague à une autre. Ces résultats déjà notés dans les analyses précédentes : absence d'interaction vagues et régions avec des effets de la région, confirme que la région constitue un niveau supplémentaire qui structure les données.

II – Prise en compte du niveau Région

Etant donné l'effet de la région, sur les variations des scores de satisfaction inter-magasin, considérer la région comme un niveau supplémentaire est envisageable. Dans cette perspective le modèle devient :

$$\text{Niveau 1} \quad S_{tik} = M_{0ik} + e_{tik}$$

$$\text{Niveau 2} \quad M_{0ik} = \gamma_{00k} + u_{0ik}$$

$$\text{Niveau 3} \quad \gamma_{00k} = \gamma_{000} + v_{00k}$$

Les deux équations réunis nous donnent :

$$S_{tik} = \gamma_{000} + v_{00k} + u_{0ik} + e_{tik}$$

Avec $u_{0ik} \sim N(0, \sigma_u^2)$, $v_{00k} \sim N(0, \sigma_v^2)$ et $e_{tik} \sim N(0, \sigma_e^2)$. L'indice t est relatif au 19 vagues, i est relatif aux 88 magasins et k indexe les 9 régions étudiées.

¹¹⁹ : Comme dans le cas d'une Anova, lorsqu'on traite un facteur la constante correspond à une modalité de référence ici c'est la région « EST ». La constante affiche un score de satisfaction moyen de 37.06.

¹²⁰ Il faudra se garder de tirer des conclusions liées aux régions car ce n'est pas toutes les régions qui sont représentées. Ici les données concernent une enseigne, un secteur précis.

¹²¹ Indice inspirée du R² de la régression linéaire MCO, il permet de juger de la part d'explication lorsqu'une variable explicative est introduite dans le modèle.

Dans ce modèle la quantité σ_u^2 représente la variance de score de satisfaction moyen (constante γ_{000}) de chaque magasin i au sein de chaque région k (niveau 2). La quantité σ_v^2 représente quant à elle la variance des scores moyens entre les régions (niveau 3).

L'application du modèle aux données est accompagné par une étape supplémentaire qui consiste à standardisé les scores de satisfaction pour chaque période (vague). L'intérêt ici est d'ordre pratique, par le fait que les variances estimées dans chaque niveau vont s'interpréter comme des pourcentages, étant donné que la variance totale est égale à 1. L'application nous donne donc les résultats suivants :

Tableau 26 : Modèle hiérarchique avec niveau magasin et région

Paramètres	Modèle 1	Modèle 2
Effets Fixes		
Constante	-3.361e-16	0.004887
Effets aléatoires		
Niveau 3 : variance inter-régions : σ_v^2		0.359
Niveau 2 : variance inter-magasin : σ_u^2	0.653	0.336
Niveau 1 : variance intra-magasin : σ_e^2	0.342	0.342
Qualité du modèle		
-2 log V	3275	3239
AIC	3281	3269
BIC	3297	3247

Dans ce tableau, nous avons représenté le modèle 1 (deux niveaux : Vagues et Magasin) contre le modèle 2 (trois niveaux : Vagues, Magasin et Région). Nous remarquons que la variance inter-magasin (modèle 1) est nettement diminuée, elle passe de 65% à 33%. C'est la même tendance remarquée lorsque la variable région a été intégrée comme variable explicative. Lorsque le niveau région est ajouté, nous remarquons que la variance totale des scores de satisfaction se répartie sur chacun des niveaux : 35% pour la région, 33% entre magasins et 34% au niveau le plus bas. Cette décomposition nous indique que le modèle 2 est plus en adéquation avec les données : avec une déviance (-2 log Vraisemblance) significative¹²² et des critères AIC ou BIC qui indiquent une amélioration.

¹²² La déviance est un critère de jugement des modèles, plus elle est basse plus le modèle est adapté aux données. Elle est à utiliser seulement dans le cas où les deux modèles différent seulement dans leurs parties aléatoires.

Enfin nous remarquons que la constante est nulle, ce fait est naturel car les données (scores de satisfaction) ont été standardisées.

Ainsi nous concluons que les données qui présentent une stabilité apparente au niveau global (sans distinguer les magasins) sont compatibles avec une structure hiérarchique qui conditionne ces données et qui crée une stabilité lorsque la structure n'est pas prise en compte.

Section 5 – Conclusion du chapitre

Les éléments apportés par ces résultats nous indiquent que les indicateurs de satisfaction subissent des erreurs qui sont attachées à des sources indépendantes de l'effort Marketing. Suivant les trois propositions testées, nous avons pu prouver par contre-exemple, qu'un indicateur n'est pas toujours représentatif de ce qu'il doit mesurer. Nous avons mis en lumière une situation de discordance entre le concept (ici la satisfaction) et son indicateur. Dans cette étude de cas, les sources de perturbation, qualifiées de facteurs de contingence, ne constituent qu'une partie des éléments susceptibles de causer des problèmes d'objectivité de l'indicateur.

Par cela, nous concluons que la mesure barométrique de la satisfaction doit être plus vigilante vis à vis sources liées aux caractéristiques des individus et de leur environnement. La non prise en compte de ces derniers peut affecter les qualités de l'indicateur de performance, jusqu'à fausser les prises de décision.

Pour nous, la comparaison transversale directement sur les indicateurs, est donc une pratique insuffisante pour le contrôle de la performance. Un classement entre magasin suivant les indicateurs de satisfaction est très fragile surtout lorsqu'on néglige les facteurs géographiques¹²³.

Pour les comparaisons longitudinales, la stabilité est conditionnée suivant les niveaux de calcul. Dans notre recherche, une stabilité apparente au niveau global (c'est-à-dire sans distinction des magasins), le devient moins lorsqu'on descend au niveau de chaque magasin. Cette stabilité est le fruit d'une considération purement statistique, dû à une structure hiérarchique.

Nos conclusions indiquent que dans le cas des indicateurs basés sur des sondages, on ne peut se contenter d'un calcul comme la moyenne ou le taux. Il faudrait passer au-delà de

¹²³ Par essence les magasins se trouvent dans des zones différentes géographiquement, qui peuvent aussi différer entre elles en termes des profils sociodémographiques et socioéconomiques.

ces instruments en prenant en compte d'abord la provenance des clients, de leurs valeurs (en termes de chiffres d'affaires) et surtout raisonner suivant les structures influentes (les facteurs de contingence) lorsqu'elles existent.

Chapitre 5 : L'indicateur de satisfaction face à son principal concurrent, le cas du Net Promotor Score

Dans un objectif de contrôle de la performance via les indicateurs non financiers¹²⁴, les recherches pointent la nécessité de trouver des indicateurs tournés vers les clients. La satisfaction et ses indicateurs constituent une des possibilités, pour juger des comportements des clients. Dans les enquêtes de satisfaction, la présence de certaines questions touchant de près à la satisfaction (intention de ré-achat, recommandation) a permis l'émergence d'autres indicateurs basés sur la recommandation dont le plus populaire est le Net Promotor score¹²⁵

Le Net Promotor score s'inscrit dans la lignée des indicateurs de performance. Il est utilisée au même titre que l'indicateur de satisfaction pour contrôler la performance entre unité de gestion (ex : magasin) dans le temps et dans l'espace. Il se situe en aval de l'indicateur de satisfaction car la recommandation est établie comme conséquence de la satisfaction (Anderson 1998).

A la lumière des éléments développés dans cette thèse : l'action des facteurs de contingence sur les indicateurs de satisfaction, le NPS risque lui aussi de souffrir des faiblesses que nous avons pu montrer. Le NPS a pour support « la recommandation », il vise à contrôler la stabilité du portefeuille clients et ainsi de faire état de la dynamique commerciale. Le NPS par construction doit être la conséquence de l'indicateur de la satisfaction, et doit hériter de ce fait de ces faiblesses, car comme l'indicateur de satisfaction, il est le résultat d'une mesure agrégée, et donc la présence des facteurs de contingence peuvent aussi l'affecter.

Le Net promotor score paraît donc fragile pour contrôler la performance, au même titre que l'indicateur de satisfaction. Toutefois dans la pratique managériale, il devient très populaire car les managers trouvent en lui un bon indicateur de performance, sans pour autant que cela soit vérifié par les recherches académiques.

Devant les faiblesses des indicateurs de satisfaction à juger les comparaisons inter-entreprises, le NPS s'est imposé petit à petit. Il faut dire que la facilité à appréhender ce

¹²⁴ Nouveaux indicateurs introduits suite à des critiques des indicateurs exclusivement comptables. Ce sont des indicateurs tournés vers le marché rendant compte de la satisfaction des clients.

¹²⁵ . Le Net Promotor score est une différence entre deux taux (individus recommandant fortement la firme moins individus ne recommandant pas la firme)

Promotor Score

dernier en termes de simplicité de la question, de l'échelle et du mode de calcul, a favorisé sa duplication à des fins de comparaison entre les entreprises. Un point que l'indicateur de satisfaction peine à combler.

A l'heure actuelle, Le NPS s'est établi naturellement auprès des managers et son utilisation s'est démocratisée parfois même en remplaçant l'indicateur de satisfaction. Devant ce conflit d'usage, les indicateurs de satisfaction se trouvent délaissés au profit du NPS, le tout avec des fondements théoriques très limités (Morgan et Rego 2006). Ainsi il existe aujourd'hui un débat sur l'utilisation du Net Promotor score comme substitut de l'indicateur de satisfaction. D'ailleurs son initiateur Reichheld (2003), le présente comme l'indicateur nécessaire et suffisant pour prédire et pour juger du comportement des clients.

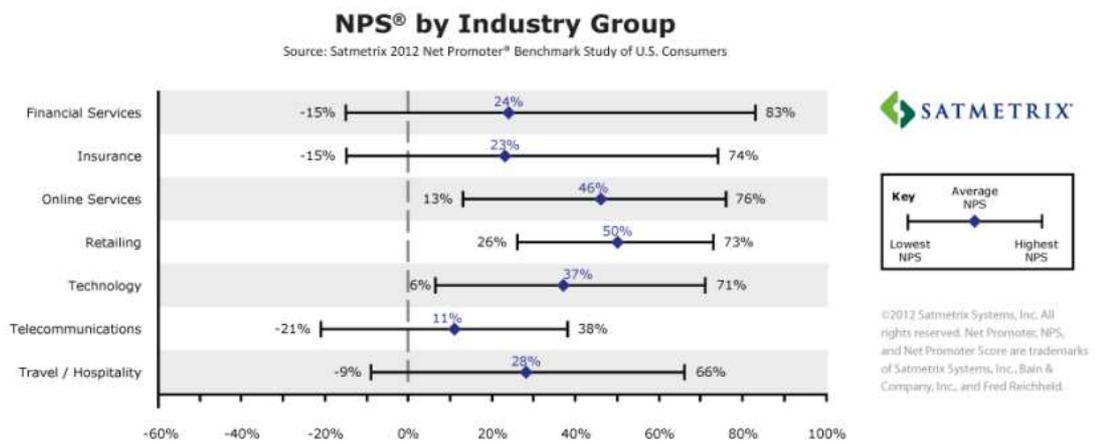


Figure 40: Benchmarks 2012 du NPS par industrie, source site web SATMETRIX

Secteur	Score minimum	Score global secteur	Score Maximum	Etendue de l'intervalle
Health Insurance	67	72	73	6
Life Insurance	78	81	81	3
Property and Casualty Insurance	77	78	81	4
Internet Retail	74	81	86	12
Internet Portals and Search Engines	74	79	82	8
Internet Social Media	61	69	78	17
Computer Software	75	77	77	2
Personal Computers	77	80	86	9
Banks	66	77	79	13
Specialty Retail Stores	76	79	83	7
Limited-Service Restaurants	73	80	83	10

Figure 41 : Benchmarks 2012 par industrie des scores ACSI, source site web ACSI

Promotor Score

A notre sens, le succès du NPS, vient de son caractère ondoyant, par rapport à des indices basés sur la satisfaction clients (le cas de l'ACSI¹²⁶ en figure 41). La figure 40 résume les différentes valeurs du NPS obtenues pour différents secteurs. Elle nous montre une fluctuation importante du NPS à l'intérieur de chaque secteur et par rapport à l'ensemble. D'un autre côté, pour l'indice ACSI, la fluctuation reste relativement faible par rapport à la précédente. L'ACSI est un indice à base 100. Il voit rarement ses valeurs descendre en dessous de 50. De plus, si la fluctuation d'un indicateur est utile au manager pour un contrôle au sein de l'entreprise, elle reste inquiétante¹²⁷ lorsqu'on s'intéresse à une comparaison entre entreprises.

C'est pourquoi il est urgent de comprendre ces faits et de déterminer s'il y a une prééminence du Net Promotor Score sur l'indicateur de satisfaction. Puis d'identifier les faiblesses et les points forts de cet indicateur dans le cadre de la mesure barométrique de la satisfaction clients.

Notre démarche dans ce chapitre, par les développements qui vont suivre, vise à comparer les informations apportées par les deux indicateurs: celui basé sur la satisfaction et celui basé sur la recommandation. Pour parvenir à nos fins, la première comparaison s'est fixée sur les deux supports des indicateurs: satisfaction vs recommandation, puis dans la seconde nous avons examiné la liaison entre les deux indicateurs dans leurs utilisations comme indices de performance. Le but de ces comparaisons est de mettre en lumière ce qui est commun et ce qui ne l'est pas.

I – La satisfaction et la recommandation

La question de la satisfaction clients est un sujet primordial. Depuis les travaux fondateurs (Oliver 1980 ; Howard 1968; 1969 ; Andreasen 1977) elle est considérée comme un vecteur de développement et de croissance. Une satisfaction élevée conduira probablement au ré-achat et à un bouche à oreille favorable (Lebarbera et Mazursky, 1983). La satisfaction est devenue un moyen de pilotage des activités commerciales et se situe au cœur de la politique centrée clients. L'intérêt porté à la satisfaction clients a permis la mise en place de sondages réguliers auprès des consommateurs. Ils sont devenus aujourd'hui des moyens incontournables pour le contrôle de la croissance. Ces enquêtes faites trimestriellement ou annuellement ont un ultime but qui est d'explicitier la

¹²⁶ American Customer Satisfaction Index

¹²⁷ Un indice basé sur des résultats de sondages, reste forcément affecté par la question de la représentativité de l'échantillon. Une grande variation de l'indice n'est pas indépendante de la fluctuation des échantillons.

Promotor Score

croissance sous ses deux dimensions: verticale et horizontale. Une composante verticale qui découle du comportement du ré-achat des consommateurs et une composante horizontale qui résulte de la conquête de nouveaux clients. Pour arriver à cet objectif, divers indicateurs ont vu le jour dont les plus connus sont: le taux de satisfaits (TS) et le Net Promotor Scor (NPS). L'utilisation de ces deux indices prédictifs, est devenue inévitable dans un quotidien professionnel préoccupé par la mesure de la future croissance.

II – La dynamique des deux dimensions de la croissance future

Dans une vision axée clients, deux composantes de la croissance se distinguent et qui sont les résultats de deux comportements des consommateurs. Toute entreprise cherche à recruter (dimension horizontale) et à fidéliser (dimension verticale) (Figure 42). Cependant dans la pratique cette dernière se voit contrainte à contrôler une dimension plus qu'une autre. Par exemple le secteur des banques ou de la téléphonie, adopte une politique centrée plus sur le recrutement des nouveaux clients que sur la fidélisation. Contrairement au secteur du luxe qui lui est plus axé sur la fidélisation de la clientèle déjà présente, car l'extension de leurs portefeuilles n'est pas toujours possible. Ce dernier cas est aussi présent dans un marché concurrentiel où la conquête de nouveaux clients est difficile.

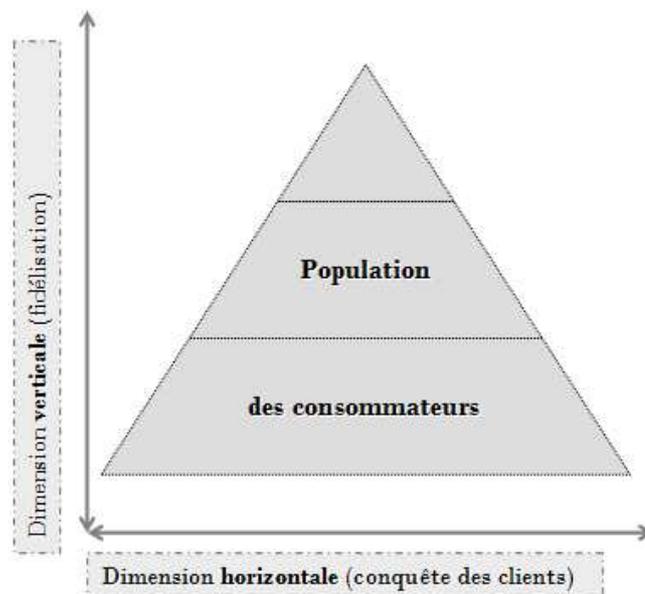


Figure 42: Les deux composantes de la croissance

Promotor Score

Ainsi à travers ces éléments on voit le caractère contingent des deux dimensions (privilégier l'une ou l'autre selon le secteur) et l'asymétrie dans le contrôle entre les deux composantes de la croissance future par les entreprises. C'est pourquoi le contrôle de la croissance a soit été axé sur la recommandation soit sur la satisfaction à travers les deux indicateurs comme le Net Promotor score (NPS) ou le taux de satisfaits (TS).

III – Le taux de satisfaits (TS) et Net Promotor Score (NPS) comme indices concurrents

Le taux de satisfaits revient à faire un calcul de pourcentage dans la population des consommateurs cibles, et le Net Promotor Score revient à calculer la différence de pourcentages entre ceux qui recommandent fortement (promoteurs) et ceux qui ne recommandent pas du tout (détracteurs). Le choix de ces deux indicateurs pour sonder la croissance est conditionné aux deux dimensions de cette dernière telle que:

- le **NPS** est utilisé pour le contrôle de l'extension du portefeuille clients en donnant des indications sur les futurs clients conquis: « *plus on vous recommande, plus le nombre des clients augmente et le profit avec* ». Il permet donc de sonder la perspective d'extension horizontale sur la population des consommateurs.
- Le **TS** quant à lui est un indice prédictif du maintien du portefeuille clients: « *plus les clients sont satisfaits plus ils seront fidèles, donc un profit assuré* ». Il permet donc de sonder la perspective d'un maintien vertical sur la population des consommateurs.

Cependant, malgré leur utilisation qui peut être dissociée en termes d'usage, les deux indices font l'objet de débats au sein de la communauté des professionnels d'études. Cette forme de concurrence entre les deux indices a été introduite en même temps que l'indice NPS par son initiateur Reichheld (2003). L'auteur préconise l'utilisation de la seule question de recommandation pour assurer la prospérité et la vraie croissance des entreprises. De ce fait, plusieurs firmes ont centré leurs actions sur la seule question de « recommandation » en utilisant le NPS comme l'ultime prédicateur de la croissance future.

Bien sûr nous sommes d'accord sur l'utilisation de ce score pour une évaluation dans le temps, car la modification sensible du taux de clients qui ne recommande plus

Promotor Score

l'entreprise, est un résultat préoccupant. Mais il paraît peu probable qu'une seule question pourrait capter un phénomène bidimensionnel et aussi complexe que la croissance future. De plus en admettant que le NPS est un indice suffisant pour sonder la croissance, que devient la question de la satisfaction qui est défendue dans la littérature marketing pour son pouvoir prédictif de la fidélité clients donc du profit et de la croissance (Fornell, et al 1996; Vanhamme, 2002). Récemment des chercheurs comme Timothy L. Keiningham et al (2007) se sont intéressés à corréliser les deux indicateurs NPS et ASCI (l'American Customer Satisfaction Index) à la croissance. Les résultats de la comparaison indiquent que les deux indicateurs sont corrélés à la croissance, au même niveau. Leurs résultats nous montrent que le NPS ne peut pas être considéré comme le seul indicateur de la croissance.

IV – Satisfaction et recommandation quelle liaison

La littérature marketing a déjà établi que la recommandation (bouche à oreille) est une des conséquences de la satisfaction (Anderson 1998). Selon cet auteur, le bouche à oreille serait surtout le fait de clients très satisfaits ou au contraire très insatisfaits. D'autres chercheurs comme Kraft et Martin (2001) lient la recommandation positive à la satisfaction élevée, à la norme sociale et aux facteurs personnels. D'autres comme Richins et Bloch (1986) postulent que la liaison entre l'insatisfaction et la recommandation négative est modulée par l'implication envers le produit et l'expertise du consommateur. Ainsi au vu de ces éléments, la recommandation et la satisfaction paraissent comme deux concepts étroitement associés. Donc en suivant le paradigme de Disconfirmation des attentes (Oliver 1980), le modèle s'étend pour accoler la satisfaction à la recommandation sous une forme de schéma mental qui se situe au niveau de l'individu et qui peut être résumé ainsi :

Promotor Score

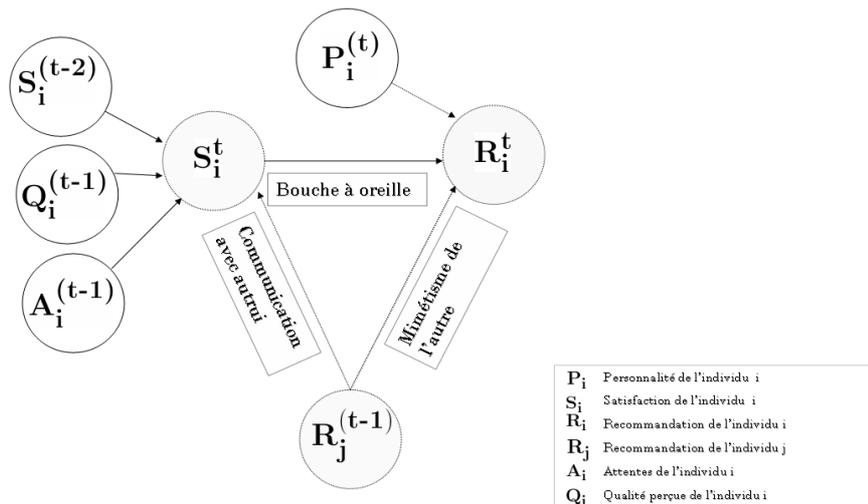


Figure 43 : Schéma de liaison entre la satisfaction et la recommandation au niveau d'un individu

La **satisfaction** de l'individu à un instant donné est liée à celle exprimée précédemment, à ses attentes *a priori*, à la qualité perçue et à la recommandation antérieure d'un autre individu captée par lui.

La **recommandation** de l'individu à un instant donné est influencée par sa satisfaction au même instant, par sa personnalité et par la recommandation d'un autre individu à un instant antérieur

Ainsi donc, établir une mesure sur la recommandation d'un ensemble d'individus est dépendante de celle basée sur la satisfaction de ces mêmes individus. Au vu de ces éléments, il serait logique de considérer que les informations fournies par chaque outil (NPS ou TS) sont liées et il semble préférable de s'intéresser à l'exploitation des deux indicateurs simultanément, au vu de la liaison entre leurs supports, au lieu de les confronter pour un choix unique.

L'objet de ce chapitre est de raisonner à partir des éléments de construction des deux indicateurs afin de répondre aux deux interrogations suivantes :

1-Etre un promoteur est- t-il suffisant ou juste nécessaire pour être satisfait?

2-Etre un détracteur est- t-il suffisant ou juste nécessaire pour être insatisfait?

Les réponses à ces deux questions nous permettront de traduire la dépendance entre les deux concepts de satisfaction et de recommandation en un sens d'implication entre les

Promotor Score

deux indicateurs. Renseigner la façon dont les deux indices se comportent entre eux, nous permettra de déterminer les informations communes et celles qui sont indépendantes.

V –Méthodologie

Pour répondre à ces interrogations nous avons exploité la méthode d'analyse implicative présentée par Bernard et Charron (1996). Cette méthode permet de donner le sens de l'implication à partir d'un tableau de contingence 2x2, en utilisant l'indice de H de Loevinger. Cet indice permet de juger du degré d'implication entre deux événements -ici entre les promoteurs et les satisfaits et entre les détracteurs et les insatisfaits.

D –H de Loevinger

Soit deux événements A et B, pour juger de leurs dépendances on peut commencer par chercher le sens de liaison entre ces deux événements. On dira que A implique B ($A \Rightarrow B$), si la réalisation de A conduit à celle de B, ce qui signifiera que le succès de l'un conduit au succès de l'autre. Une telle implication ($A \Rightarrow B$), signifie que A est suffisant pour avoir B, et que B est nécessaire pour avoir A (la non réalisation de B, conduit à la non réalisation de A)

En statistique, juger de l'implication entre deux événements passe par l'utilisation du H de Loevinger (1947, 1948), qui est définie comme un indicateur descriptif permettant de quantifier le sens de l'implication en donnant l'intensité de cette dernière. **Plus l'indice est proche de la valeur 1 plus l'implication sera considérée comme forte**, donc nous pourrons connaître le sens à privilégier grâce à la vraisemblance quantifiée par ce dernier. Pour ce faire, des pourcentages sont calculés à partir du croisement des deux variables satisfaction & recommandation catégorisées dans le tableau suivant:

Tableau 27 : Structure de croisement NPS et indicateur de satisfaction

	Promoteurs « note >=9 »	Détracteurs « note <=6 »	
Satisfaits « note >=7 »	f _{sp}	f _{sd}	f _S
Insatisfaits « note <=6 »	f _{ip}	f _{id}	f _I
	f _P	f _D	

Promotor Score

Dans notre recherche, nous voulons établir si le fait d'être un promoteur est suffisant pour être satisfait (Promoteurs \Rightarrow Satisfaits) ou être promoteur est juste nécessaire pour être satisfait (Satisfaits \Rightarrow Promoteurs). De même, on voudrait savoir aussi: si le fait d'être un détracteur est suffisant pour être insatisfait (Détracteurs \Rightarrow Insatisfaits) ou être détracteur est juste nécessaire pour être insatisfait (Insatisfaits \Rightarrow détracteurs). La réponse à ces questions est possible en exploitant le H de Loevinger comme suit :

- A- Le H de Loevinger qui permet de juger du degré de l'implication :
« Promoteurs \Rightarrow Satisfaits », se définit par: $H_{PS} = 1 - \left(\frac{f_{ip}}{f_{I} \times f_{P}} \right)$
- B- Le H de Loevinger qui permet de juger du degré de l'implication
« Satisfaits \Rightarrow Promoteurs », se définit par: $H_{SP} = 1 - \left(\frac{f_{sd}}{f_{D} \times f_{S}} \right)$
- C- De plus le H de Loevinger est égal pour les propositions suivantes :
(Promoteurs \Rightarrow Satisfaits) et (Insatisfaits \Rightarrow Détracteurs)
(Satisfaits \Rightarrow Promoteurs) et (Détracteurs \Rightarrow Insatisfaits)

Donc cet indice nous permettra de fixer le sens de l'implication en comparant l'intensité obtenue pour les implications testées. Afin de comparer les degrés d'implication, nous utiliserons les valeurs repères définies par les Bernard et Charron(1996) sur lesquelles le H pourra être examiné. Si $H < 0.40$ il y'a absence d'implication. Si $0.4 < H < 0.60$ il y'a une tendance à l'implication. Si $H > 0.60$ il y'a une quasi-implication. Nous choisirons ainsi les implications les plus fortes suivant ces valeurs pour définir le sens à privilégier.

E – Inférence sur le H de Loevinger

Une fois les valeurs de H calculées pour chaque implication il devient nécessaire de donner une **garantie** au résultat descriptif obtenu. En termes de statistiques, cela se traduit par le niveau de confiance qu'on peut accorder à une valeur de H quelconque. Dans notre étude nous évaluerons la probabilité pour que H soit supérieur à la valeur repère 0.6. Plus cette probabilité est grande plus la garantie accordée à une quasi-implication est forte.

Suivant les recommandations des deux auteurs Bernard et Charron(1996) et en utilisant l'inférence bayésienne non informative. Nous avons calculé pour chaque H le degré de

Promotor Score

confiance pour choisir les implications les plus consistantes entre: (Promoteurs \Rightarrow Satisfaits) ou (Satisfaits \Rightarrow Promoteurs).

L'inférence sur H part du principe que le tableau croisé (Table1) est une réalisation d'une loi de Dirichlet *a posteriori* (sachant les données observées) à **k =4** paramètres :

$$(fsp, fip, fsd, fid | \text{Données}) \sim Di\left(\left(fsp + \frac{1}{k}\right), \left(fip + \frac{1}{k}\right), \left(fsd + \frac{1}{k}\right), \left(fid + \frac{1}{k}\right)\right)$$

A partir de cette loi on pourra simuler la loi de H de Loevinger *a posteriori* suivant les formules suivantes :

$$\mathbb{P}(H_{PS} | \text{Données}) = 1 - \left(\frac{\beta(fip, f_0 - fip)}{\beta(fip+fid, f_0 - (fip+fid)) * \beta(fip+fsp, f_0 - (fip+fsp))} \right)$$

$$\mathbb{P}(H_{SP} | \text{Données}) = 1 - \left(\frac{\beta(fsd, f_0 - fsd)}{\beta(fsd+fid, f_0 - fsd+fid) * \beta(fsd+fsp, f_0 - fsd+fsp)} \right)$$

Avec $f_0 = \sum f$ et $\beta(a, b)$ une loi bêta avec les deux paramètres de forme a et b

Ainsi grâce à cette loi simulée nous pouvons établir la probabilité que H dépasse 0.6 ce qui nous renseignera sur le degré de crédibilité des quasi-implications testées.

VI – Résultats

L'analyse a été effectuée pour trois secteurs qui sont : l'énergie, la téléphonie mobile et l'industrie. Les données disponibles concernent six entreprises, avec un cumul des données de 5000 individus (les données en annexe). Les résultats de comparaison de chaque sens d'implication sont les suivants:

Tableau 28 Résultats d'implication NPS et Taux de satisfaction

Entreprise:	Secteur	Promoteurs \Rightarrow Satisfaits		Satisfait \Rightarrow Promoteurs	
		H de Loevinger	Confiance : Pr (H>0.6)	H de Loevinger	Confiance : Pr (H>0.6)
A	Téléphonie	0.90*	100%	0.53	0.06%
B	Téléphonie	0.75*	95%	0.44	2.3%
C	L'industrie	0.74*	90%	0.53	29%
D	L'industrie	0.98*	100%	0.2	0%
E	Energie	0.98*	100%	0.69*	88%
F	Energie	0.94*	100%	0.58	39%

*quasi -implication

Promotor Score

Ces chiffres nous montrent que l'implication **promoteurs** \Rightarrow **satisfaits** est plus forte que dans le sens **satisfaits** \Rightarrow **promoteurs**. Car les H correspondants à ce sens sont tous nettement supérieurs par rapport au sens **satisfaits** \Rightarrow **promoteurs**, le tout avec une garantie satisfaisante (Pr (H>0.6)).

Ceci nous amène à considérer que le sens « Promoteurs \Rightarrow Satisfaits » est le plus consistant. De plus grâce à la propriété du H de Loevinger l'implication « Insatisfaits \Rightarrow Détracteurs » est aussi l'implication qu'il faudra privilégier. Ce raisonnement rejoint, la logique : « si on est insatisfait on ne vous recommandera pas et si on vous recommande on est forcément satisfait de vous ».

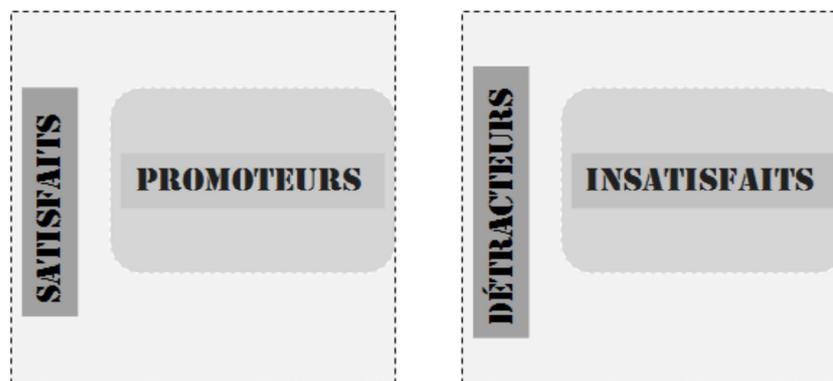


Figure 44: Les sens d'inclusion (d'implication) entre les populations des satisfaits insatisfaits et promoteurs détracteurs

En termes ensemblistes, cette dépendance peut se traduire par le fait que la population des insatisfaits est incluse dans la population des détracteurs. Et la population des promoteurs est incluse dans la population des satisfaits. (Figure 44)

VII – Interprétation de la dépendance des indicateurs

Le NPS par son calcul de différence de pourcentages : **promoteurs moins détracteurs**, nous renseigne sur la croissance future à travers la conquête de nouveaux clients, mais il juge mal les clients déjà présents. Le taux de satisfaits surveille les clients présents mais capte mal le phénomène de recrutement. Cependant les implications obtenues livrent deux informations complémentaires qui comblent les insuffisances des deux indicateurs pris séparément.

*Les **insatisfaits**, ne favorisent pas le développement **vertical et horizontal** dans la population des consommateurs

Promotor Score

*Les **promoteurs** contribuent à la fois des deux sortes de développement **vertical**, et **horizontal** dans la population des consommateurs

En termes d'indicateurs ces remarques se traduisent par: Le fait que le NPS sous-estime la croissance dans sa dimension verticale (tous les détracteurs ne sont pas néfastes à la croissance) et le taux des satisfaits surestime la croissance dans sa dimension horizontale (les satisfaits ne sont pas tous des promoteurs).

Donc les entreprises doivent réduire l'insatisfaction qui peut être néfaste à son développement et augmenter le nombre de ses promoteurs donc des satisfaits. Ainsi, il suffira juste de surveiller les insatisfaits et les promoteurs ce qui assurera la prise en compte de toute l'information disponible pour prédire la croissance de l'entreprise dans ses deux dimensions verticale et horizontale.

➤ Indicateur composite

Un nouvel indice composite serait donc efficace s'il s'intéresserait aux promoteurs et aux insatisfaits. Dans ce cas, le support de l'indice examinera les deux noyaux : **positifs et négatifs** tout en étant suffisant pour prédire la croissance sous ses deux formes : verticale et horizontale.

Une idée d'indicateur serait donc une différence entre les pourcentages des: **Promoteurs moins Insatisfaits**. Le fait de s'intéresser à cette différence revient à prendre une sorte d'intersection minimale des deux indicateurs (TS et NPS) qu'il faudra surveiller. Ainsi plus cette quantité est grande plus la croissance est positive sur ses deux dimensions.

Assurer un fort taux de promoteurs et un faible taux d'insatisfaits, est l'objectif qui doit être poursuivi. Raisonner sur ces deux quantités permettra de se munir d'un indicateur de performance qui résume la croissance future issue du comportement clients, sur ses deux composantes (recrutement et fidélisation).

VIII – Conclusion

Les dénouements de cette étude montrent que le taux de satisfaits et le NPS véhiculent une information commune dans leurs prédictions de la croissance, ce qui prouve leurs liaisons. Néanmoins ils semblent que tous les deux sont incomplets dans leurs utilisations séparées, d'où l'intérêt de les adopter simultanément pour le contrôle de la croissance sur ses deux composantes.

Promotor Score

Nous avons signalé au début de notre développement que le succès du NPS est souvent lié à sa capacité à « bouger ». Toutefois à notre sens, le degré de sa variation est lié au mode de calcul (une différence de pourcentages). En appliquant, la même procédure pour la satisfaction, des résultats mouvants sont aussi constatés. De ce fait la question qu'il faudra poser à l'avenir est de savoir si le NPS est un bon indicateur de performance en donnant une signification à sa sensibilité.

Nos conclusions ont montré, une dépendance du NPS au Taux de Satisfaction, ce qui a permis de dégager un sens d'exploitation qu'il faudra privilégier à l'avenir pour mieux capter le phénomène de la croissance future (insatisfait vs promoteur). Il faudrait aussi étendre ces résultats à d'autres secteurs pour confirmer cette hypothèse d'autant plus que nous constatons que la liaison n'est pas identique entre les secteurs. Ces résultats confirment l'hypothèse que l'intérêt porté au produit/service et l'expertise du consommateur, jouent des rôles modérateurs entre la satisfaction et la recommandation (Richins et Bloch, 1986). Donc en plus de la différenciation par secteur, il devient nécessaire de prendre en considération l'objet étudié (l'importance du produit/service aux yeux du consommateur) et le degré d'expertise du consommateur (BtoB, BtoC) pour interpréter les résultats de ces deux indices. Cette différence par secteurs rejoint aussi le fait que chaque entreprise privilégie une dimension de la croissance plus que l'autre. Néanmoins selon les résultats obtenus nous constatons que lorsque l'intérêt est porté sur la recommandation il ne faut pas négliger les insatisfaits (tous les détracteurs ne sont pas néfastes) et si l'on veut s'intéresser à la satisfaction, il ne faut pas négliger les promoteurs (tous les satisfaits ne sont pas bénéfiques).

De ce fait les entreprises doivent conjuguer les deux dimensions de la croissance car les deux composantes: verticale et horizontale vont de pair et s'influencent mutuellement.

Discussion générale

Le projet de cette thèse vise l'étude de la satisfaction dans le cadre barométrique. Notre propos ambitionne l'examen de l'indicateur de satisfaction dans ses fonctions de contrôle de performance. Il s'agit de faire un diagnostic de son utilisation comme outil de comparaison entre unités de gestion (le cas du magasin). L'aboutissement de notre recherche met en évidence l'existence contingente de certains facteurs qui perturbent l'indicateur dans son rôle de comparaison d'un point de vue transversal (entre magasin) et longitudinal (l'évolution des indicateurs dans le temps). Ces facteurs sont établis comme sources explicatives de l'effet du contexte qui régit la mesure barométrique de la satisfaction clients. Nous avons aussi signalé que les indicateurs de satisfaction (le cas des taux de satisfaction) et le Net Promotor Scor peuvent être vus comme des marqueurs complémentaires pour juger de l'activité dont l'utilisation conjointe paraît plus discriminante qu'une utilisation indépendante. Un élément supplémentaire qui émane de nos travaux réside aussi dans la procédure de test qui sert à définir l'action d'un facteur de contingence sur l'indicateur de satisfaction. Ainsi nous préconisons que toute recherche sur les facteurs de contingence doit avoir comme socle la procédure statistique présentée dans cette thèse.

Il faut signaler aussi que nous avons utilisé pour arriver à ces résultats des données barométriques issues d'enquêtes professionnelles. Cette démarche est légitime étant donné que les problématiques traitées concernent la mesure barométrique de la satisfaction et non le concept lui-même. Cette position nous a amené à élargir le champ de cette thèse à une considération qui s'attache plus au domaine de la statistique (quantitative) avec un cadre de recherche qui s'approche des outils de contrôle de gestion. Sans prétention à apporter des éléments théoriques à ce domaine, nous nous sommes attelés à inspecter l'application des indicateurs de satisfaction comme outils de gestion, d'où la bifurcation vers le domaine de contrôle de gestion. Il faut rappeler aussi que même si la satisfaction a été beaucoup étudiée dans le Marketing, elle reste néanmoins aussi importante dans une littérature de comptabilité qui a fait d'elle un élément de pilotage de la performance des entreprises.

Notre réflexion s'est détachée de la considération standard d'une recherche ciblée

sur la modélisation des concepts. Nos développements concordent avec une vision axée sur l'application (point de vue empirique). Par cela, nous nous sommes intéressés à mettre en évidence les difficultés pratiques de la mesure de satisfaction clients en identifiant les sources de leurs existences. Étant donné notre distinction claire entre le concept de la satisfaction et l'indicateur, tous nos résultats ont eu des portées associées aux indicateurs et non sur le concept. Ce cas de figure justifie l'approche statistique de nos problèmes, car par essence un indicateur de satisfaction est le résultat d'une agrégation statistique (moyenne taux ou score).

Nous discuterons dans cette partie deux positions. La première émane directement des résultats et vise la nécessité de procéder à des modélisations statistiques plus fines pour que les indicateurs de satisfaction soient plus fiables. Dans ce cadre-là, nous pensons notamment à la modalisation multiniveaux. La seconde position qui sera discutée est à caractère « épistémologique ». Elle vise à susciter l'intérêt de l'utilisation des données barométriques comme moyens de recherches afin de traiter des problématiques très importantes jusque-là inexplorées. Il s'agit ici de proposer des pistes pour élargir le champ de recherche en Marketing, en axant les nouvelles investigations sur des données barométriques.

Section 1 – La modélisation multiniveaux une approche naturelle pour les données barométriques

Les résultats de test des indicateurs de satisfaction dans leurs fonctions de contrôle de performance, ont permis de mettre en évidence l'existence de sources contingentes susceptibles d'inquiéter le processus de comparaison des performances entre points de vente (magasins). Les conséquences de ces perturbations sont le fruit de facteurs environnementaux comme le pays, la région, ou encore des facteurs liés aux caractéristiques individuelles comme l'âge, le sexe, ou encore la fréquentation clients. Par définition, ces facteurs sont considérés comme des éléments attachés à une mesure agrégée sur un ensemble d'individus. Comme l'indicateur de satisfaction est une computation d'un groupement d'avis de satisfaction, ces facteurs de contingence agissent comme des quantités qui le conditionnent en bloc en créant une nouvelle source de variances au niveau de l'individu et au niveau du magasin. Ce fait nous laisse penser que les deux niveaux (individu et magasins) doivent être considérés et distingués lors de la mesure de satisfaction, car leurs interactions semblent la cause des perturbations des

indicateurs de satisfaction. En pratique, admettre qu'un indicateur agrégé est transposable à des conclusions individuelles, sous-entend que les avis de satisfaction des clients, sur lesquels l'indicateur est calculé sont homogènes. Pourtant l'existence de facteurs de contingences est une preuve de la non validité de cette considération. C'est sur cette voie de dissociation entre les niveaux magasins et individus que les analyses futures doivent être axées.

Ces deux niveaux sont différents par nature et la non prise en compte de leurs propriétés induit des mesures de satisfaction différentes qui inhibent la comparaison en vue de contrôler la performance. Parler de niveau sous-entend une structure dans les données. Dans nos propos, elle s'est manifestée avec une disposition qui lie le client à un magasin qui est lui-même attaché à une région et un pays. Ce fait met l'indicateur de satisfaction à un niveau magasin caractérisé par des variables attachées aux points de ventes (facteurs environnementaux) et un niveau client identifié par toutes sortes de variables sociodémographiques. Il en vient, qu'une analyse doit pouvoir dissocier les caractéristiques du niveau agrégé (magasin) et du niveau individuel (le client).

L'analyse multiniveaux vise à prendre en compte ces structures en expliquant la variabilité de chaque niveau (ex : individus et magasins) par des caractéristiques attachées à la sous population (ici les facteurs d'environnement) et par des caractéristiques individuelles (ici les facteurs sociodémographiques). Par cela l'analyse multiniveaux permet de décomposer la variance de chaque niveau pour mieux contrôler l'hétérogénéité liée à la structure de l'information. Ce qui induit de raisonner suivant la structure hiérarchique qui dans ce cas est naturelle.

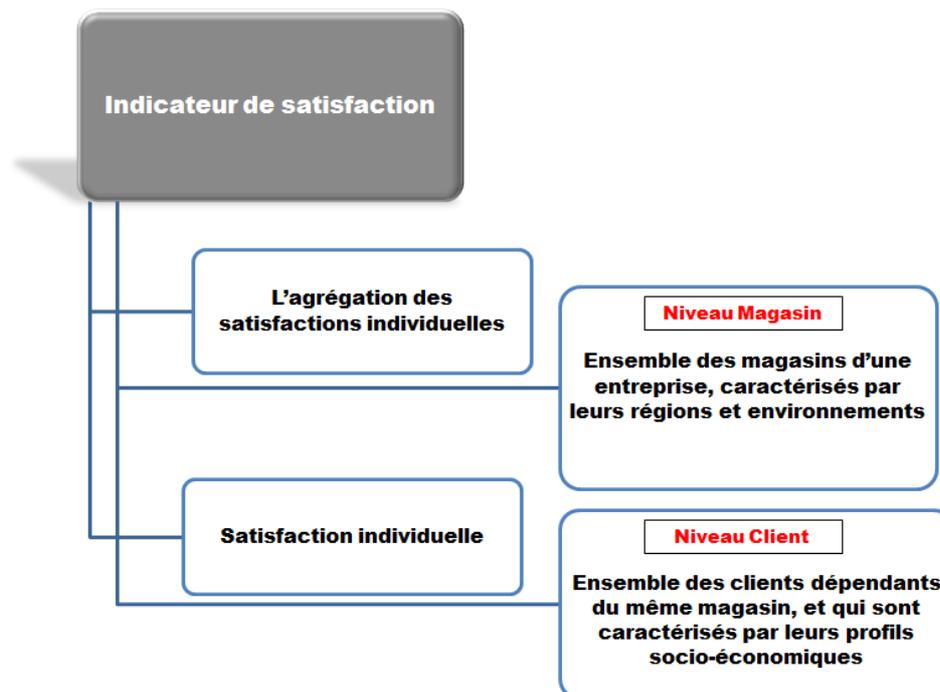


Figure 45: Dépendance de l'indicateur de satisfaction, aux niveaux d'analyse.

I – L'analyse multiniveaux en Marketing

L'analyse multiniveaux n'est pas étrangère au domaine du marketing, cependant elle reste conditionnée par une organisation spécifique des données qui induisent des approches par niveau. Dans ce dernier cas l'analyse multi-niveaux est incontournable. Le cas des données de sondage par magasin est un des exemples des structures multi-niveaux confirmé par la présence de facteur de contingence. Jean Claud. Ray et Daniel Ray (2008), dans leurs articles (*analyse multi-niveaux en Marketing*) présentent l'approche de la modélisation multi-niveaux en donnant un exemple d'application à partir des données de satisfaction et fidélité récolté par magasin. Dans leurs examens du lien entre la fidélité et la satisfaction, ils ont introduit des variables au niveau individuel comme le sexe des répondants et des variables au niveau magasin comme l'intensité concurrentielle. Ces deux groupes de variables ont servi à expliquer le lien en explicitant la variabilité de chaque niveau (clients, magasins) par les informations qui lui sont propres. Dans leur développement, ils préconisent l'utilisation des analyses multi-niveaux dans le cas où les données ont une structure naturelle qui distingue les niveaux, et lorsque le besoin de les

séparer est inévitable. Ce fait se traduit par des sources de variances différentes où chacune est attachée à une classe ; d'où l'intérêt de séparer les niveaux lors de l'analyse.

II – L'intérêt de l'analyse multiniveaux

L'utilité première de cette analyse est avant tout un avantage de modélisation statistique. C'est-à-dire une mesure plus fidèle du phénomène étudié. L'analyse multiniveaux apporte une meilleure estimation en décomposant les erreurs pour chaque niveau et ainsi expliciter les dépendances. Par exemple des clients qui fréquentent le même magasin ont de forte chance de partager les mêmes caractéristiques, ce qui induit des autocorrélations qui peuvent fausser le calcul statistique. Ainsi un modèle qui décompose les niveaux en associant à des variables qui leurs sont spécifiques permet d'expliquer les dépendances attachées au niveau individu (client) et celles situées au niveau agrégé (le magasin). Dans notre cas, cette approche doit être applicable lors de l'évaluation de l'importance des attributs de la satisfaction : modèle multi attributs¹²⁸.

La non considération de la structure multi-niveaux, peut conduire à certaines erreurs d'interprétation des résultats, comme *l'erreur atomiste* (D.Courgeau et B. Baccaini, 1997). Cette dernière est due à la négligence de la structure multi-niveaux en ignorant le contexte où les individus évoluent. Ce fait, se manifeste quand on considère que seules les caractéristiques individuelles influencent les comportements. D'autres erreurs d'interprétation peuvent aussi exister lorsque l'on cherche à déceler des comportements individuels à partir de mesures agrégées. Cette erreur est connue sous le nom d'erreur écologique. Elle a été identifiée pour la première fois dans l'article de Robinson (1950) qui conclut par : « on ne peut pas utiliser une corrélation écologique, mesurée au niveau agrégé, comme substitut d'une corrélation individuelle ». Ces deux types d'erreurs peuvent conduire à des conséquences fallacieuses, et à une mauvaise prise de décision. C'est pourquoi, l'idée de travailler simultanément à divers niveaux d'agrégation en vue d'expliquer le comportement individuel permet de raisonner à la fois suivant les caractéristiques des individus et ceux de la sous-population (ensemble de clients d'un magasin) qui va elle aussi affecter le comportement de l'individu qui en fait

¹²⁸ Durant notre recherche, il a toujours été question d'une satisfaction globale qui doit être expliquée par des attributs de service et/ou du produit (ex : satisfaction de l'accueil). Ce modèle connu sous le nom de modèle multi-attributs, est utilisé par les entreprises pour hiérarchiser (définir l'importance des attributs) pour mieux cibler et améliorer la satisfaction globale.

partie¹²⁹. Ces analyses duales : individu et contexte dans lequel il vit, permettent d'éviter l'ensemble des erreurs d'interprétation, en ayant un modèle plus représentatif de la structure des données.

III – La modélisation multiniveaux pour les données barométrique.

Les résultats de notre thèse indiquent que l'indicateur de satisfaction calculé de façon agrégée sur les magasins, subit des perturbations de la part des facteurs de contingence. Dans ces conclusions, la région ou le pays semblent être des facteurs influents. Ceci implique une variation, des valeurs données par les indicateurs, selon la région et le pays en plus de la variation des notes de satisfaction au niveau de l'individu, en conséquence l'existence de plusieurs niveaux est confirmée. De plus comme le but des indicateurs de satisfaction est de comparer les points de ventes, la disparité selon les facteurs comme la région ou le pays deviennent problématiques. D'où l'idée de distinguer les niveaux (magasins et individus) à travers les caractéristiques de chacun en effectuant des analyses par niveaux. Ceci consiste à introduire des quantités décrivant séparément le magasin et les individus afin de mieux fidéliser la mesure de l'indicateur et ainsi atténuer les variations dues aux facteurs de contingences.

Un magasin reste toujours identifié selon sa région ou son pays, donc de la population qui y habite et de la concurrence qui y existe. Par exemple le fait de prendre en compte ces informations, permettra de discerner les régions entre elles et le fait de les inclure permettra de réduire la variation due aux deux facteurs (régions et pays). Pour ce qui est le niveau de l'individu toutes sortes de données sociodémographiques peuvent être utiles pour réduire la variance résiduelle. Le facteur de l'âge traité dans notre recherche constitue un exemple.

Le rôle de l'analyse multi-niveaux dans ce cas consiste à modéliser chaque degré (magasin et individu) en expliquant la variance par l'ensemble des caractéristiques de chacun des niveaux. Ce procédé permettra ainsi d'obtenir des indicateurs (moyenne ou taux) plus représentatifs, c'est-à-dire ne représentant pas d'effets émanant des facteurs de contingence, à condition que ces derniers soient importants.

Au-delà des indicateurs de satisfaction, la modélisation multi-niveaux pointe la nécessité d'utiliser des modélisations plus représentatives du phénomène étudié (le niveau

¹²⁹ Il permet de prendre en compte l'influence du contexte : « Tout ce qui se ressemble s'assemble, ou tous qui s'assemblent ont tendance à se rassembler ».

de satisfaction clients) en raisonnant suivant des structures hiérarchiques qu' il faudra contrôler lorsque la mesure est effectuée.

Section 2 – Vers une recherche ciblée sur la mesure barométrique

Dans notre recherche nous avons fait qu'effleurer le potentiel de recherche qui peut émaner de l'étude des baromètres de satisfaction. Pour distinguer l'intérêt de la recherche académique de celle que nous défendons ici, il faut différencier le but poursuivi dans chaque cas. La recherche académique possède des outils et des procédures avec un cheminement axé sur la compréhension et la modélisation théorique des concepts. Cette approche-là, ne peut pas être valable dans une recherche basée sur les données barométriques. Ceci est dû à la disposition des informations récoltées dans le cadre barométrique, qui ne s'approprient pas à ce type d'approche. Par exemple, la présence d'une mesure barométrique de la satisfaction, de la qualité et de la fidélité, ne permet pas l'étude de ces concepts car la mesure ne répond pas à la validité psychométrique selon le paradigme de Churchill¹³⁰. La mesure barométrique vise à fournir une structure traduisant la satisfaction sous forme de composantes concrètes (les attributs de satisfaction). Ces attributs sont propres à chaque entreprise et sont supposés capter la satisfaction de leurs clients. De plus il est commun, qu'une question de satisfaction globale soit utilisée comme variable dépendante pour permettre de hiérarchiser, après analyse statistique, les poids de ces attributs de satisfaction. D'un point de vue de structure, la divergence est claire, cependant cette discordance ne permet pas à notre sens d'écarter l'utilisation des baromètres du domaine de la recherche. Le désintérêt du domaine académique vient essentiellement des problématiques des échelles utilisées. Des problématiques touchant la façon de formuler les questions et surtout à la composition même des échelles qui est mono-items et non multi-items. Pourtant il existe des avantages à l'utilisation des baromètres, en revanche leur emploi ne peut dépasser le cadre de l'évaluation des indicateurs et ne peut pas concerner directement le concept lui-même. La raison de cette remarque vient du fait que les baromètres sont des mesures effectuées *a posteriori* (disponible, et non modifiable) et leur usage ne peut être considéré que sur des sujets liés à la mesure. D'autre part, suivant les avantages majeurs des baromètres en termes de volume de données et l'utilisation d'un même questionnaire dans le temps, font que leurs

¹³⁰ Gilbert A. Churchill, Jr. A Paradigm for Developing Better Measures of Marketing Constructs, Journal of Marketing Research, Vol. 16, No. 1 (Feb., 1979), pp. 64-73

utilisations en vue de la recherche est plus aisée.

Par conséquent, présenter les bienfaits des baromètres pour la recherche, passe par situer leurs utilisations dans un cadre précis de recherche le tout en discutant les avantages et inconvénients de ces instruments.

I – La question de l'échelle

L'échelle est l'instrument par lequel le sondeur capte un avis de consommateur. Dans une échelle, il y a la question qui sert à capter le concept étudié et le support qui lui est censé discriminer les réponses (capter fidèlement l'information). Le premier souci avec les mesures barométriques, c'est que la formulation des questions ne fait que rarement référence aux antécédents et aux conséquences de la satisfaction. Alors qu'on sait qu'une bonne mesure, dans le cas de la satisfaction, est celle qui permet entre autre de distinguer cette dernière des concepts voisins (la qualité, attitude, performance). Le second ennui dans une mesure barométrique est lié au support des échelles utilisées qui sont mono-items, une seule question pour capter la satisfaction. Ce choix permet de minimiser les non réponses et vise à formuler des items compréhensibles et sans ambiguïté pour les répondants. L'option des échelles mono-items est en contradiction avec la nécessité psychométrique de fiabilité et de validité. Il faut que les échelles soient discriminantes envers le concept étudié.

La qualité psychométrique d'une échelle, initiée par Churchill (1979), a été importée de la littérature en psychologie. Un certain nombre de concepts a été adopté afin de proposer une procédure de construction d'outils de mesure. Cette procédure consiste à vérifier la fiabilité et la validité des échelles selon un protocole bien établi de test et re-test des échelles. L'auteur préconisa dans ses recommandations que les chercheurs sont mieux servis par des mesures multi-items que par des mesures mono-items, lorsqu'ils veulent mesurer de façon fiable et valide le comportement du consommateur. Le paradigme ainsi défini, est fondé sur l'idée qu'il existe un concept dans « l'esprit du consommateur ». Cette technique vise à capter un construit latent à travers des manifestations de ces attributs reflétant le concept étudié, où la somme des attributs constitue l'échelle multi-items. Ces attributs étant associés au construit latent qui est non observable, permettent d'inférer à partir de leur observation le concept étudié. Cette variable latente est supposée, comme équivalent statistique, du concept présent dans

l'esprit du consommateur¹³¹

Pour arriver à capter convenablement le construit latent, la procédure consiste à vérifier que les attributs choisis sont un « bon reflet du concept ». Pour ce faire le paradigme à suivre est :

- a. De spécifier le domaine du construit
- b. générer une liste d'items traduisant les attributs retenus
- c. purifier cette liste en écartant les questions superflues
- d. reformuler, en suite les attributs en étudiant leurs fiabilités et leurs validités jusqu'à ce que l'outil satisfait les critères imposés par le chercheur.

Vérifier la fiabilité signifie que la mesure est exemptée d'erreur, la purification consiste à atteindre une bonne fiabilité, en étudiant la cohérence interne des attributs à représenter le concept étudié. Tous les attributs doivent faire référence au seul construit. La purification consiste donc à enlever ou à ajouter des items et à tester leur cohérence à l'aide d'un indice dit « **Le coefficient alpha de Cronbach** » (Cronbach 1951). Plus ce coefficient est proche de 1 plus l'outil de mesure est fiable. Toutefois la procédure de purification reste conditionnée à la bonne définition du concept. Il se peut que des attributs soient fiables mais pointent un autre concept d'où l'intérêt de bien définir le concept étudié.

Quant à la validité, dans le cas de l'appréciation du construit par un ensemble d'attributs consiste à vérifier si les attributs partagent un maximum de variances avec le construit qu'ils sont censés mesurer. Il s'agit de la validité convergente. Le but étant de bien spécifier la relation attributs et variables latentes. Dans la considération statistique, tester la validité est possible grâce aux « **Rhô de Jöreskog** »¹³². Ainsi la fiabilité consiste à étudier le plan transversal (interne aux attributs) et la validité correspondante qui est à étudier au plan horizontal (construit et attributs).

Dans un cadre barométrique, ces étapes sont souvent omises pour des mesures mono-items. Ce fait est à la source de l'inconsidération de la recherche pour les données

¹³¹ Cette notion est présente dans une lecture critique dans la communication de Galan Jean-Phillipe : une déconstruction des outils de mesure en marketing ; Actes de congrès de L'AFM 2011.

¹³² Nous renvoyons le lecteur à l'article de Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 48, 39–50.

issues des mesures barométriques. Il semblerait à notre sens, que la nature des échelles est le seul argument qui empêche l'utilisation des données barométriques à des fins de recherche. Cette raison est tout à fait louable mais elle n'explique pas l'impossibilité d'étendre la recherche en marketing à d'autres pensées autres que celles touchant la psychométrie sous forme d'une approche hypothético-déductive¹³³.

II – Échelles multi-items et mono-items

Comme cela a été précisé précédemment, dans le domaine du marketing les échelles multi-items règnent en maître quand il s'agit de recherches basées sur des concepts psychologiques. Ce fait est bien sûr lié à la fiabilité de ces instruments pour mesurer le concept. Cependant quelques critiques s'opposent à l'utilisation de la fiabilité comme seul moyen pour juger l'acceptabilité d'une mesure. Par exemple (Bergkvist L. et Rossiter J.R, 2008), en s'appuyant sur des études antérieures comme celles de Gorsuch et McFarland (1972) et celles de Rossiter (2002), estiment que la fiabilité ne doit pas être le seule critère, si la mesure mono-items à une validité prédictive équivalente à la mesure multi-items. Les auteurs précisent aussi que les mesures mono-items sont suffisantes si le construit est concret dans l'esprit du consommateur. L'aspect concret veut signifier que l'objet étudié peut être facilement et uniformément imaginé. Les auteurs proposent aussi ce raisonnement en prenant comme référence les travaux de Drolet et Morrison (2001), qui montrent que dans une mesure multi-items, plus les attributs sont nombreux pour cerner le construit, plus les utilisateurs cherchent à produire des items synonymes, ce qui peut cacher la présence de mauvais items parmi eux. Cette situation est possible quand le mauvais item est corrélé aux bons, ce qui engendre une augmentation du coefficient alpha dissuadant ainsi les chercheurs à éliminer les mauvais items. L.Bergkvist et J.R Rossiter (2008) concluent dans leurs études concernant la publicité qu'aucune différence n'est constatée dans la validité prédictive entre échelle multi-items et mono-items. Cependant leurs études est applicable à des objets concrets et singuliers, à différencier avec les construits abstraits possédant plusieurs dimensions, dont la satisfaction fait partie. Ainsi par ces quelques arguments on constate que les échelles multi-items ne constituent pas l'ultime mode de mesure, dans la recherche marketing.

Il faut rappeler aussi, que l'utilisation des mesures multi-items permet d'éviter les erreurs de formulations, car la purification (test de fiabilité) sert à se prémunir de cela. De

¹³³ Il s'agit de confronter un modèle de mesure (construit sur la base d'une théorie) à des données empiriques.

leur côté les échelles mono-items doivent à tout prix cibler le concept avec une énonciation précise et claire. Ce dernier point est d'autant plus aisé que le concept est concret. Dans la pratique barométrique, la mesure de satisfaction est aussi distinguée par des attributs qui sont cette fois opérationnels. Par exemple la satisfaction est appréciée par diverses composantes : l'accueil, le service après-vente ou encore la qualité du produit. Ces attributs-là sont captés avec une seule question donc une échelle mono-items. Le but de cette mesure n'est pas de sonder le modèle qui lie les attributs à la satisfaction globale, mais plutôt de l'appliquer en supposant qu'il est vrai. Ce modèle traduit ainsi, le fait que les différents attributs sont les sources de la satisfaction globale. De plus comme les attributs captés par les échelles mono-items sont de caractère concret, dans ce cas les échelles mono-items peuvent être satisfaisantes.

En ayant des attributs concrets les praticiens ne cherchent pas à expliciter la satisfaction dans ces dimensions affectives ou cognitives valable pour chaque individu, mais plutôt à construire une résultante agrégée qui traduit l'évaluation globale des clients. L'objectif est de quantifier un avis général. Pour ce faire ils supposent que leurs performances (en terme de satisfaction) est jugé sur des attributs. Une bonne performance sur ces attributs engendrera un niveau de satisfaction élevé dans la population. Le modèle sous-jacent, suppose que le consommateur considère un produit (ou un service) non pas comme un tout, mais comme une somme de plusieurs éléments qui le caractérise. Ainsi on constate que le but des données barométriques est de donner des résultats relatifs à une population, car c'est la somme des avis qui est cherchée. L'intérêt de cette pratique vient de la nécessité de confronter ce que pense l'entreprise à ce que pense le client, pour optimiser le schéma suivant :

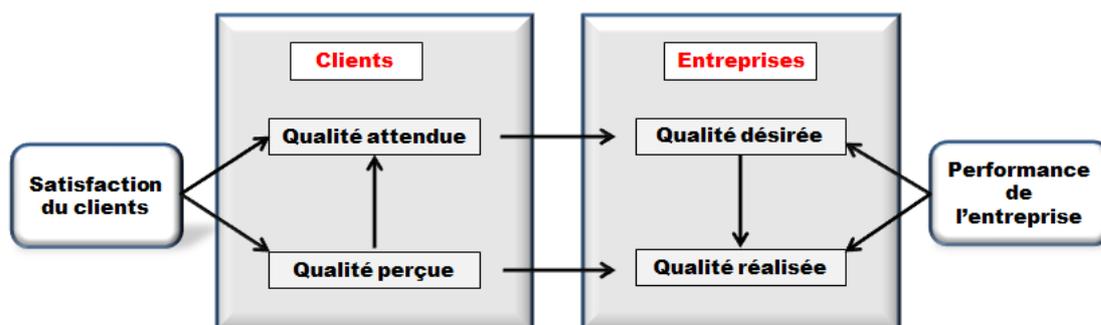


Figure 46 : Relations action entreprise et avis clients. D.RAY (2004)

De ce point de vue, l'approche de la construction des baromètres reste différente de celle préconisée dans la recherche académique. De ce fait, les données barométriques ouvrent une nouvelle voie de recherche qui doit fonctionner suivant le modèle de la démarche barométrique pour résoudre les problématiques qui leur sont propres. Cependant reste à expliciter cette approche pour mieux l'utiliser.

III – Organisation des données barométriques comme des modèles formatifs

Contrairement à la pratique académique où les attributs sont associées aux construits via un modèle qui considère les attributs comme des manifestations du concept présent dans l'esprit du consommateur, la pratique barométrique impose que les attributs sont les causes de ce qui doit être mesuré ici le niveau de satisfaction. Ce niveau de satisfaction cherché est la conséquence des attributs, et de ce fait ce sont les attributs qui forment le concept. Ce schéma est formatif, en opposition aux modèles réflexifs (les attributs reflètent le construit).

En revenant aux conditions de fiabilité, Jean-Phillipe Galan (2011), indique que lorsque le modèle est formatif, la procédure « de purification » des attributs, fondée sur la fiabilité ne peuvent pas s'appliquer. Il donne l'exemple, du statut socio-économique qui lui est la conséquence d'un ensemble d'attribut (l'éducation, le revenu, le prestige de l'activité professionnelle) et que enlever un attribut est inenvisageable car les attributs sont les causes. Ainsi tester la fiabilité en cherchant la cohérence interne (**Le coefficient alpha de Cronbach**), ne devient plus une condition. D'autre part l'auteur précise que le paradigme de Churchill est fondé sur le fait que les attributs sont le reflet du concept. Dans ce cas seulement, les attributs ne varient pas comme les autres peuvent être éliminés. Dans la situation inverse (modèle formatif) éliminer une mesure diminuera la validité du construit, car les attributs représentent les différentes facettes du construit.

C'est pourquoi dans une approche barométrique, les attributs constituent le concept qui est la satisfaction au niveau de la population. De son côté, la mesure académique constitue des mesures de type réflexif qui sont les conséquences d'un construit (la satisfaction) au niveau de l'individu. Ceci est possible car le construit, présent dans l'esprit du consommateur est supposé le même pour chacun. Pour le point de vue

barométrique, la satisfaction individuelle est agrégée vers une résultante globale, et l'hypothèse que le construit est identique pour chaque individu n'est plus nécessaire.

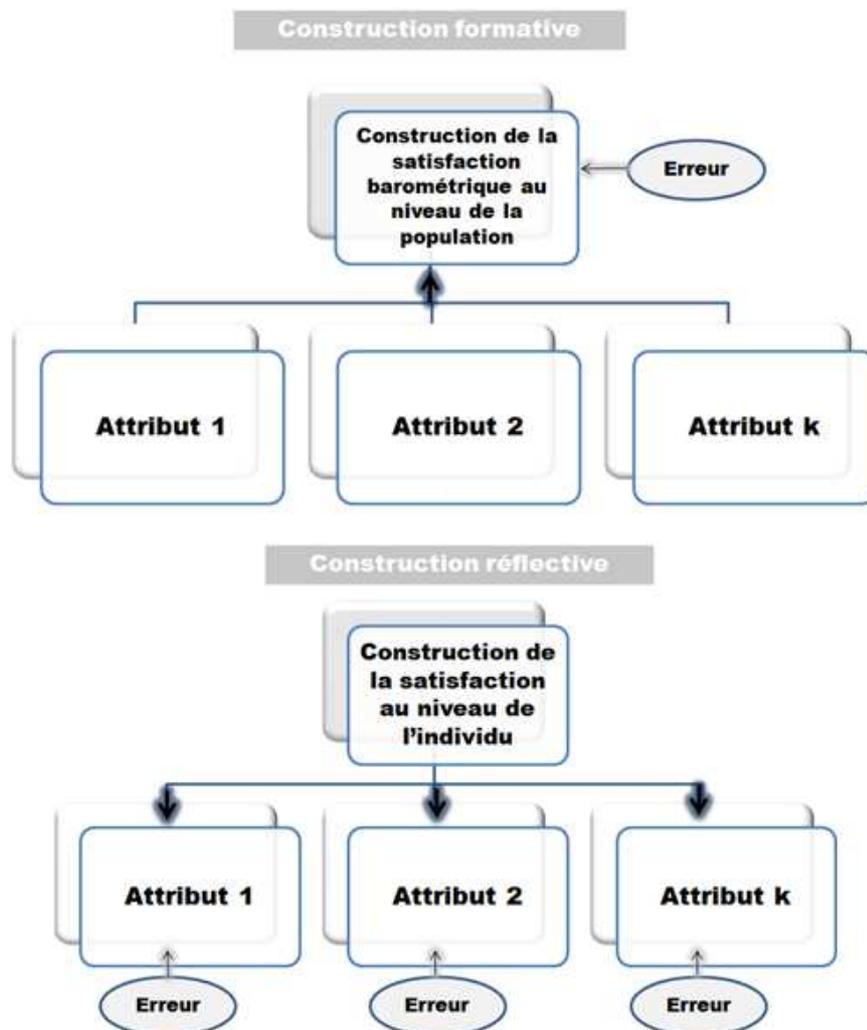


Figure 47 : Les formes de construction: attributs et variables latentes

Conclusion générale

Actuellement le concept de satisfaction est bien circonscrit. Dans un cadre purement académique, la théorisation de la notion et l'explication de sa formation a permis une meilleure opérationnalisation du concept. Les entreprises modernes intègrent la satisfaction clients comme éléments de gestion, à tel point qu'elle est reconnue comme le noyau de la commercialisation. Toutefois son développement et son utilisation dans le champ professionnel ne s'est pas effectué avec les mêmes démarches académiques. Son utilisation comme indicateur de performance a laissé déborder le concept à d'autres dispositions qui ne lui sont pas inhérentes. Nous pensons notamment à sa mesure. A l'exception des méthodologies des baromètres nationaux préconisées par l'approche de Claes Fornell, la mesure de satisfaction clients s'est contentée de produire des questionnaires est de les analyser statistiquement pour construire un indicateur. Ces procédés, même en écartant les biais émanant des questions, ont résumé la mesure de satisfaction en un processus d'agrégation statistique qui est loin de représenter le concept de satisfaction. La satisfaction est un concept psychologique, donc interprétable au niveau de l'individu, et ne peut être défini de la même façon quand on considère un ensemble de clients. Cette satisfaction agrégée, reste néanmoins dépendante du type de calcul et du niveau d'agrégation préconisé (ex : magasins, régions pays), par conséquent les problématiques touchant les indicateurs sont plus attachés à une considération statistique.

Cette thèse constitue un apport sur au moins trois points. Le premier est relatif à l'approche considérée. Nous nous sommes intéressés à l'indicateur de satisfaction dans son fonctionnement comme un outil de comparaison dans le cadre de la mesure barométrique. Chose rarement étudiée comme telle dans la littérature marketing. Le deuxième point est spécifique au cadre d'action des facteurs de contingence. Une définition des facteurs de contingence est apportée, et un diagnostic de leurs effets perturbateurs de la mesure est mis en évidence. Le troisième point est relatif à la méthodologie adoptée. Elle se distingue, d'une part par l'utilisation des outils statistiques rarement employés dans le domaine, et d'autre part par la procédure de test qui permet de quantifier l'importance des effets des facteurs de contingence.

Les problématiques touchant l'indicateur de satisfaction, comme celles que nous avons traitées durant cette thèse, se sont inscrites sur une voie d'investigation des qualités des indicateurs dans leurs fonctions de contrôle, le tout en considérant l'indicateur de satisfaction comme un construit statistique issu d'une agrégation d'un ensemble de mesures individuelles. Le concept de satisfaction est bien sûr toujours présent est inhérent à nos problématiques, toutefois la portée de nos travaux (indicateurs de satisfaction) et les outils mis à disposition (un ensemble de baromètres), nous ont orienté vers une recherche qui vise à examiner la mesure de satisfaction plus que le concept lui-même.

Dans le développement de cette thèse, il a été question de vérifier si des indicateurs de satisfaction (agrégés au niveau magasin) subissent des perturbations dans leurs fonctions de pilotage, lorsqu'on les conditionne par un positionnement géographique (exemple de la région ou le pays) ou lorsqu'on les traite à la lumière des caractéristiques sociodémographiques des individus. Nos conclusions nous indiquent que les niveaux de satisfaction (agrégés par l'unité magasin) sont sensibles à la variation des facteurs géographiques et sociodémographiques, avec des effets notables pour le facteur géographique. Ces faits corroborent la possibilité que les indicateurs de satisfaction ne soient pas indépendants de l'environnement où le sondage est effectué. Par exemple, une différence entre les niveaux de satisfaction entre des unités géographiques dissemblables peut être de sources contingentes. Car les clients de chaque région sont différents et une comparaison directe est déconseillée.

D'autres points ont été soulevés dans ce travail. Ils concernent la stabilité apparente des indicateurs de satisfaction d'une vague à une autre. Outre des effets de certains facteurs dans cette stabilité, il est établi que la stabilité est conditionnée aussi par les niveaux de calcul. Par exemple, une stabilité aux niveaux nationaux (agrégation de tous les clients de l'entreprise) le devient moins lorsqu'on descend au niveau inférieur (ex : magasins de l'entreprise).

L'ensemble de résultats révèlent que les indicateurs de satisfaction subissent des problématiques très différentes de celle qui sont étudiées dans la littérature académique et ouvrent ainsi un champ de recherche qui est celui de la mesure barométrique. Il y a ici un domaine complet à explorer qui se trouve à mi-chemin entre la statistique, le contrôle de gestion et le marketing. De plus, étant donné l'importance de la notion d'indicateur dans le suivi de la performance, le travail sur la mesure barométrique de la satisfaction devient

important pour éclairer les questionnements des managers, qui à la longue deviennent pesants sur le concept de la satisfaction.

Le concept de satisfaction a été longuement étudié, mais son opérationnalisation peine à être encadrée. L'émergence des nouvelles problématiques d'ordre barométrique, justifie pleinement la nécessité d'étendre le travail de recherche pour non seulement expliquer, mais surtout afin d'établir des processus de mesure à la lumière des baromètres nationaux. Ce travail doit être accompagné avec des procédures statistiques appropriées et des outils adaptés aux problématiques des managers.

A la lumière de nos résultats, une des positions que nous défendons est celle qui préconise un étalonnage des mesures de satisfaction pour chaque baromètre, afin de comparer seulement ce qui est comparable.

Étant donné la limite de notre étude de cas, seule une partie des facteurs est testée. Par définition ces facteurs sont situationnels et peuvent donc se manifester avec des importances différentes suivant le cas d'étude. Ce qu'il faut retenir dans nos résultats est le fait que leurs actions peuvent handicaper les fonctions des indicateurs de satisfaction quand ils sont utilisés pour comparer les performances entre points de ventes.

Les facteurs de contingence traités dans notre recherche sont spécifiques à notre étude de cas, d'autres peuvent subsister selon le secteur ou le type de baromètre. Nous pouvons penser à l'ensemble des facteurs sociodémographiques, l'éducation, l'expertise du client ou encore son ancienneté.

D'une façon générale, les facteurs de contingence sont vus comme des quantités qui conditionnent les indicateurs de satisfaction. Toutes procédures de comparaison doit tenir compte de ces facteurs, à condition qu'ils soient établis comme notables par des procédures statistiques comme celles que nous avons exposé dans notre travail.

Dès le départ, nous avons fait la distinction entre la procédure classique de la recherche, et celle que nous discutons. Notre démarche vise à donner des arguments pour appréhender de nouvelles problématiques liées à la satisfaction mais qui concernent un autre champ, qui est celui des indicateurs de satisfaction. Ces derniers sont, par construction, les seuls produits des baromètres car ils sont des quantités relatives à une population et non aux individus. A partir de cette considération, la problématique barométrique sera située au niveau de la population. Ceci signifie que les problématiques traitées ne sont plus d'ordre psychologique mais bien statistique ou encore socio-

psychologique. Un des sujets qui peut concerner l'indicateur est la façon avec laquelle il est évalué, c'est-à-dire le calcul statistique qui le produit. À notre sens les problématiques touchant le sujet sont importantes et aujourd'hui, il y a une nécessité à les traiter.

Nous avons signalé que dans les mesures barométriques, les échelles utilisées sont mono-items. Ces échelles peuvent représenter fidèlement l'information, à condition que les attributs fassent référence à des objets concrets pour les consommateurs et que les questions soient bien formulées. Si on rajoute à cela, l'avantages du nombre de données dans les études barométriques, cela nous conforte à supposer que ces données ont une aptitude à servir la recherche.

Par cela, la recherche basée sur les données barométriques devient complémentaire en étendant le champ d'application à des nouveaux sujets qui une fois résolus, contribueront à mieux comprendre le concept de satisfaction dans son utilisation comme indicateur de performance. De plus, il est aujourd'hui important de rappeler que d'un point de vue conceptuel la satisfaction a atteint son apogée. Le construit est très bien compris. Ainsi, une recherche étendue de cette façon donne des nouvelles perspectives qui seront profitables au domaine académique.

En somme, le problème des échelles mono-items ne paraît pas constituer un obstacle insurmontable. Les données barométriques par leurs nombre de données, et l'avantage temporel (même questionnaire répété sur plusieurs vagues) peuvent apporter des éclaircissements sur la dynamique de la satisfaction ou encore sur les différences culturelles ou par secteur. Toutefois, un travail préliminaire sur le processus de mesure doit être effectué; l'exemple de l'analyse multi-niveaux¹³⁴ constitue un axe d'application

La jonction entre la recherche sur les baromètres et le développement de la modélisation multi niveaux, peut apporter une meilleure compréhension des phénomènes en marketing comme la satisfaction clients. Cette jonction permet ainsi d'apporter des éléments de réponses aux problématiques des indicateurs barométriques. Une meilleure mesure assurera une décision optimale, d'où l'intérêt d'inclure les données barométriques dans les processus de la recherche académique.

¹³⁴ L'approche multiniveaux est à privilégier dans le cas où la mesure est effectuée sur plusieurs unités : comme l'exemple des magasins. Ces unités constituent des niveaux influents qu'il faudra distinguer lors de l'analyse statistique.

Bibliographie

ANDREASEN, A.(1977) « A Taxonomy of Consumer Satisfaction, Dissatisfaction Measures, in Conceptualisation and Measurement of consumer Satisfaction and Dissatisfaction ». Cambridge, Mass, Marketing Science institute.

ANDERSON E., WEITZ B. (1989), Determinants of continuity in conventional industrial channel dyads, *Marketing Science*, 8, 4, 310-323.

ANDERSON, EUGENE W. (1994), "Cross-Category Variation in Customer Satisfaction and Repurchase Likelihood," *Marketing Letters*, 5 (Winter), 19-30.

ANDERSON E.W., FORNELL C. et LEHAMN D.R. (1994), Customer satisfaction, market share and profitability : findings from Sweeden, *Journal of Marketing*, 58, 53-66.

ANDERSON (1996) "Customer Satisfaction and Price Tolerance," *Marketing Letters*, vol. 7(3), p. 19-30

ANDERSON E.W., FORNELL C. et RUST R.T. (1997), Customer satisfaction, productivity and profitability : differences between goods and services , *Marketing Science*, 16, 2, 129-145.

ANDERSON R.E. (1973), Consumer dissatisfaction : the effects of disconfirmed expectancy on perceived product performance, *Journal of Marketing Research*, 10, 38-44.

ANDERSON M ET SULLIVAN (1993), The antecedents and consequences of customer satisfaction for firms, *Marketing Science*, 12, 125-143.

ANDERSON, FORNELL, MAZVACHERYL (2004) "Customer Satisfaction and Sharholder Value," *Journal of Marketing*, vol. 68, October, p. 172-185

ADERSON, E.W, FORNELL C (2000); Foundations of the American Customer Satisfaction Index ; Total Quality Management , Vol. 11 Issue 7, p 869, 14

ARNOULD E.J., PRICE L.P. (1993), River magic : extraordinary experience and the extended service encounter, *Journal of Consumer Research*, 20, 24-45.

AURIER P., EVRARD Y. (1998), Elaboration et validation d'une échelle de mesure de la satisfaction des consommateurs, Actes du 14ème Congrès International de l'Association Française du Marketing, 51- 71.

AUDRAIN A-F (2000), Proposition d'un cadre conceptuel des déterminants des fonctions de contribution des attributs à la satisfaction, Actes du 16ème Congrès International de l'Association Française du Marketing, 2, 905-918.

AUDRAIN A-F., Evrard Y., (2001), « Précision et discriminance du concept de satisfaction des consommateurs. », Actes du 17ème Congrès International de l'Association Française du Marketing, Deauville, CD Rom

AUDRAIN-PONTEVIA A-F(2003) étude de la dynamique de la satisfaction : bilan et perspectives de recherche, Actes du Congrès International de l'AFM 2003

AURIER P., EVRARD Y., N'GOALA G. (2000), Valeur de consommation et valeur d'usage : une application au cas de la consommation cinématographique, Actes du Congrès International de l'Association Française du Marketing, tome 1, 151-162.

AURIER P., EVRARD Y., N'GOALA G. (1998), La valeur du produit du point de vue du consommateur, Valeur marché et organisation, Actes des 15èmes Journées Nationales des I.A.E. Nantes, tome 1, 199-212.

ANTHONY, ROBERT NEWTON (1965). Planning and control systems : a framework for analysis. Harvard University Press.

ANTHONY, R. N. (1988). The Management control Function (Harvard University Press.).

BAGOZZI R.P., GOPINATH M., NYER P.U. (1999), The role of emotions in marketing, Journal of the Academy of Marketing Science, 27,2, 184-206.

BANKER, RAJIV D, GORDON POTTER, ET DHINU SRINIVASAN. An empirical investigation of an incentive plan that includes nonfinancial performance measures. The Accounting Review 75, no 1 (2000): 65-92.

BERNARD, JEAN-MARC, ET CAMILO CHARRON. « L'analyse implicite bayésienne, une méthode pour l'étude des dépendances orientées ». Mathématiques et sciences humaines no 134 (1996): 5-38.

BERGKVIST .L ET ROSSITER J.R (2008) : Comparaison des validités prédictives des mesures d'un même construit des échelles mono-item et des échelles multi-items, Recherche et applications en Marketing, vol 23, N°1

BEARDEN W.O., TEEL J.E. (1983), « Selected determinants of consumer satisfaction and complaint reports. », Journal of Marketing Research, 20, 2, pp. 21 - 28.

BLATTBERG R. ET DEIGHTON J. (1996), Manage marketing by the customer equity », Harvard Business Review, 4, 134-144.

BERLAND, NICOLAS. Mesurer et piloter la performance, 2009. <http://www.management.free.fr/>.

BRESSOUX.P Modélisation statistique appliquée aux sciences sociales (2010), Edition de Boeck.

BOUQUIN, HENRI. Contrôle de gestion. Presses Universitaires de France, 1998.

BITNER M-J. (1990), Evaluating service encounters : the effects of physical surroundings and employee responses, Journal of Marketing, 54, 69-82.

BOLTON R.N. et DREW J.H. (1991), A multistage model of customer's assessment of service quality and value, Journal of Consumer Research, 17, 375-384.

BOLTON R.N. et DREW J.H. (1994), Linking customer satisfaction to service operations and outcomes, in Services Quality, new directions in theory and practice, Rust R. Oliver R.L., Sage Publications, London.

BOLTON R.N. (1998), « A dynamic model of the duration of customer's relation with a continuous service provider : the role of satisfaction. », *Marketing Science*, 17,1, 45-65

BOLTON R.N. ET LEMON K.N. (1999), « A dynamic model of customers' usage of services : usage as an antecedent and consequence of satisfaction. », *Journal of Marketing Research*, 36, 5, 171-186.

BOLTON, KANNAN, BRAMLETT (2000) "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value, *Journal of the Academy of Marketing Science*, vol. 28(1), p. 95–108

BURLAUD.A et SIMON C.L (2013), « Le contrôle de gestion » Edition la Découverte.

CARDOZO R- N. (1965), « An experimental study of customer effort, expectation, and satisfaction. », *Journal of Marketing Research*, 2, 244-249. Cardozo, R.N. « Customer satisfaction: laboratory study and Marketing Action, ». *Consumer Behavior* (1973)..

CARLSON, J. ., VINCENT, LESLIE H, HARDESTY, DAVID M, & BEARDEN, WILLIAM O. (Feb2009). Objective and Subjective Knowledge Relationships: A Quantitative Analysis of Consumer Research Findings. *Journal of Consumer Research*, p. 864-876.

CARRILLAT, F. A., JARAMILLO, F., & MULKI, J. P. (2009, Spring). EXAMINING THE IMPACT OF SERVICE QUALITY: A META-ANALYSIS OF EMPIRICAL EVIDENCE. *Journal of Marketing Theory & Practice*, p. 95-110.

CARMAN J. M. (1990), Consumer perceptions of service quality : an assessment of SERVQUAL dimensions, *Journal of Retailing*, 33 - 55.

CHURCHILL G.A, (1979). A Paradigm for Developing Better Measures of Marketing Constructs, *Journal of Marketing Research*, Vol. 16, No. 1 , pp. 64-73

CHURCHILL G.A. et SURRENANT C. (1982), An investigation into the determinants of customer satisfaction, *Journal of Marketing Research*, 19, 491-504.

COHEN, J. (1992, juillet). A Power Primer. *Psychological Bulletin*, p. 155-159.

CHOFFEL, Denis, et François MEYSSONNIER. « DIX ANS DE DEBATS AUTOUR DU BALANCED SCORECARD ». *Comptabilité et Connaissances*, 2005.

CONCHAR, M. P., CRASK, M. R., & ZINKHAN, G. M. (2005). Market Valuation Models of the Effect of Advertising and Promotional Spending: A Review and Meta-Analysis. *Journal of the Academy of Marketing Science*.

CONSTABILE M. (1998), Customer satisfaction and trust into the ressource-based perspective. Research propositions endorsing the confirmation / disconfirmation paradigm, *Actes de la Conférence de l'EMAC*.

COOIL B., KEININGHAM T.L, AKSOY L., & HSU.H (2007) « A Longitudinal Analysis of Customer Satisfaction and Share of Wallet: Investigating the Moderating Effect of Customer Characteristics », *Journal of Marketing*

COURGEAU D. ET BACCAÏNI B. (1997), Analyse multi-niveaux en sciences sociales, Population, 4, 831-864.

COURGEAU D. (2004), Du groupe à l'individu : synthèse multi-niveaux, Paris, PUF.

CHRONBACH L.J (1961), Essentials of psycholog 2éd, NewYork, Harper & Row.

CROSNO, J. L., & DAHLSTROM, R. (2010, septembre). Examining the Nomological Network of Opportunism: A Meta-Analysis., p. 177-190.

CUCHERAT, M., BOISSEL, J. P., & LEIZOROVICZ, A. (s. d.). Manuel pratique de méta-analyse des essais thérapeutiques. CHU Lyon. Consulté à l'adresse <http://www.spc.univ-lyon1.fr/livreMA/frame.htm>

CUGINI ANTONELLA, CARU` ANTONELLA et ZERBINI FABRIZIO (2007) The Cost of Customer Satisfaction A Framework for Strategic Cost Management in Service Industries, European Accounting Review Vol. 16, No. 3, 499-530

DABHOLKAR P., (1995), The convergence of customer satisfaction and service quality evaluations with increasing customer patronage, Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior, 8, 32-43.

DAY R.L. (1990), How satisfactory is research on consumer satisfaction ? Advances in Consumer Research, 7, 593-597.

DATO-ON, M. C., & DAHLSTROM, R. (2003). A Meta-Analytic Investigation of Contrast Effects in Decision Making. Psychology & Marketing, p. 707-731.

DERBAIX et PHAM (1989), Pour un développement des mesures de l'affectif en marketing : synthèse et prérequis, , 4, 4/89, 71-87.

DUFER, JEAN, ET JEAN- Recherche et Applications en Marketing LOUIS MOULINS. « La relation entre la satisfaction du consommateur et sa fidélité à la marque: un examen critique ». Recherche et Applications en Marketing IV, no 2 (1989): 89.

DROLET A.L ET MORRISON D.G (2001), Do we really need multiple-item measures in service research ?,Journal of service Research,3,p 196-204.

EVARD Y. (1989), From involvement to satisfaction in media consumption, Journal of consumer satisfaction, dissatisfaction and complaining Behaviour, 2, 70-74.

EVARD Y. (1993), La satisfaction des consommateurs : état des recherches, Revue Française du Marketing, 144/145, 53-65.

EVARD Y., AURIER (1994), The influence of emotions on satisfaction with movie consumption, Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behaviour, 7, 119-125.

EISEND, M. (2009, Summer). A meta-analysis of humor in advertising. Journal of the Academy of Marketing Science, p. 191-203.

ERRAMI, YOUSSEF. « Les apports du Balanced Scorecard à la recherche performances », 2007.

FARLEY, U. J., LEHMANN, R. D., & SAWYER, A. (1995). Empirical Marketing generalisation using meta-analysis. *Marketing science*. vol 14 N°3.

FERN, E. F., & MONROE, K. B. (1996). Effect-Size Estimates: Issues and Problems in Interpretation. *Journal of Consumer Research*, p. 89-105.

FINNETI B. *Theory of probability*. New York Wiley, 1974

FIELD, A. P. (2001). Meta-analysis of correlation coefficients : A Monte Carlo comparison of fixed-and random effects methods. *Psychological Methods*, p. 161-180.

FORNELL, C., & LARCKER, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 48, 39-50.

FORNELL C. (1992), A National Customer Satisfaction Barometer : the Swedish experience, *Journal of Marketing*, vol. 56, 6-21.

FORNELL, C. (1995), "The Quality of Economic Output: Empirical Generalizations About Its Distribution and Relationship to Market Share," *Marketing Science*, 14 (3), 203-211

FORNELL (C.), ITTNER (C.D.), LARCKER (D. F.) (1995), Understanding and Using the American customer satisfaction index (A.C.S.I), assessing the financial impact of quality initiatives, an institute's Conference on managing for total quality, p 14

FORNELL (C.), JOHNSON (M. D.), ANDERSON (E. W.), CHA (J.), BRUYAN (B. E.), The American customer satisfaction index : nature, purpose and findings, *Journal of marketing*, 1996, (60), p. 7-18.

FOURNIER S. et GLEN MICK D. (1999), Rediscovering satisfaction, *Journal of Marketing*, 63, 5-23.

FRANKE, G. R. (2001). Applications of Meta-Analysis for Marketing and Public Policy: A Review. *Journal of Public Policy & Marketing*.

FRISOU J. (1998), Vers une théorie éclectique de la fidélité des clients : le cas des services de télécommunication, Actes du 14ème Congrès International de l'Association Française du Marketing.

GALAN JEAN-PHILIPPE (2011), Une déconstruction des outils de mesure en marketing, *Actes de congrès de L'AFM*

GLOWA LAWSON (2000) "Satisfaction measurement: Is it worth it?" *Quirk's Marketing Research Review*, www.quirks.com, No 0618

GARBARINO E., JOHNSON M.S. (1999), The different roles of satisfaction, trust, and commitment in customer relationships, *Journal of Marketing*, 63, 70-87.

GEYSKENS I., STEENKAMP J-B. et KUMAR N. (1998), Generalizations about trust in marketing channel relationships using meta-analysis, *International Journal of Research in Marketing*, 15, 3, 224-248.

GOODE, MARK AND MOUTINHO LUIZ (1996), "The Effects of Consumers Age on Overall Satisfaction: An Application to Financial Services," *Journal of Professional Services Marketing*, 13 (2), 93-105.

GOTTELAND D., HAON C., RAY D. ET BOULE J.M., 2008. La perception de l'environnement : quels effets sur la performance de l'entreprise ?. *Finance Contrôle Stratégie*, 11, 1: 155-183.

GRAFF R., PERRIEN J., RICARD J., LANDRY C., (1999), La confiance : son statut et sa valeur normative, *Actes du 15ème Congrès International de l'Association Française du Marketing*, 1, 261-282.

GREWAL, D., KAVANOOR, S., FERN, E., COSTLEY, C., & BARNES, J. (1998). Publicité comparative versus non comparative: méta-analyse. *Recherche et application en marketing*.

GORSUCH R.L ET MCFARLAND S.G (1972), Single versus multiple-item scales for measuring religious values, *Journal for the Scientific Study Of religion* , 11,1,53-64.

GOUNARIS, STATHAKOPOULOS (2001) "Behavioral responses to customer satisfaction: an empirical study," *European Journal of Marketing*, vol. 35(5/6), p. 687-670

GRIGOROUDIS E., NIKOLOPOULOU.G AND ZOPOUNIDIS.C (2008), Customer satisfaction barometers and economic development: an explorative ordinal regression analysis,*Total Quality Management*Vol. 19, No. 5, May 2008, 441-460

GUIBERT N. (1999), La confiance en marketing : fondements et applications, *Recherche et Applications en Marketing*, 14, 1, 1-19.

HARTING (1998) "Zadowolenie klientów, lecz nie nadmiernym kosztem," (trans. Customer satisfaction but not an excessive cost) *Zarządzanie na Świecie*, vol. 4, p. 46-47

HEDGES, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, p. 107-128.

HELGESEN (2006) "Are loyal customers profitable? Customer satisfaction, customer (action) loyalty and customer profitability at the individual level," *Journal of Marketing Management*, vol. 22(3), p. 245-266

HOLBROOK M.B. (1994), The nature of customer value : an axiology of services in the consumption experience, in *Service quality : new directions in theory and practice*, Ed Rust R. et Oliver R.L., Sage publications.

HOLBROOK M.B. et HIRSCHMAN E (1982), The experiential aspects of consumption : consumer fantasies, feelings, and fun, *Journal of Consumer Research*, 11, 728-739.

HONGSIK, J. C., CHANG-HOAN, C., & SUTHERLAND, J. (2007). A Meta-Analysis of Studies on the Determinants of Standardization and Localization of International Marketing and Advertising Strategies. *Journal of International Consumer Marketing*, p. 109-147.

HOMBURG, CHRISTIAN AND ANNETTE GIERING (2001), "Personal Characteristics as Moderators of the Relationship Between Customer Satisfaction and Loyalty: An Empirical Analysis," *Psychology & Marketing*, 18 (1), 43–66.

HOMBURG C., GIERING A. & MENON A. (2003) Relationship Characteristics as Moderators of the Satisfaction-Loyalty Link: Findings in Business-to-Business Context *Journal of Business-to-Business Marketing*, Vol. 10(3)

HOMBURG C., KOSCHATE N., & HOYER W.D (2006) The Role of Cognition and Affect in the Formation of Customer Satisfaction: A Dynamic Perspective *Journal of Marketing* Vol. 70 21-31

HOWARD, STEH (1969), *The theory of buying behavior*, New-York, John Wiley and sons.

HUNT H. (1977), *Customer satisfaction / dissatisfaction. Overview and future directions. Conceptualization and measurement of customer satisfaction and dissatisfaction*, H.K. Hunt ed., Cambridge, mass., MSI.

ITTNER C.D. ET LARKER D.F. (1998a), "Innovations in performance measurement : trends and research implications", *Journal of management accounting research*, Vol. 10. 205-238.

ITTNER C.D. ET LARKER D.F. (1998b), "Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction", *Journal of Accounting research*, Vol. 36, Supplement, pp. 1-35

JOHNSON, M. D., HERMANN A. ET GUSTAFSSON A. (2001). The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, v. 22, n. 2, p. 217-245, Apr.

JOHNSON M.D., HERRMANN A., ET GUSTAFSSON A. (2002) ,Comparing customer satisfaction across industries and countries , *Journal of Economic Psychology* 23 749–769

JEFFREYS H. (1961), *Theory of Probability* .Oxford university Press.

KANO, SERAKU, TAKAHASHI, TSUJI (1984) "Attractive quality and must-be quality," *The Journal of Japanese Society for Quality Control*, vol. 14(2), p. 39-48

KAPLAN R. S. , NORTON D. P. (1998), *Le tableau de bord prospectif*, Les Editions d'Organisation, Paris.

KAPLAN R.S. ET NORTON D.P. (2001a), "Transforming the Balanced Scorecard from performance measurement to strategic management: part 1", *Accounting Horizons*, Vol. 15, N°1, March, 87-104.

KOBYLANSKI A., PAWLOWSKA B. ET STRYCHALSKA-RUDZEWICZ A. (2011) The role of customer satisfaction in the quality management systems: a cross cultural study *international journal of management and marketing research* vol ,3

LA BARBERA P.A. ET MAZURSKY D. (1983), « A longitudinal assessment of consumer satisfaction / dissatisfaction : the dynamic aspect of the cognitive process. », *Journal of Marketing Research*, 20., 4, 393-404

LAI A.W.(1995), CUSTOMER VALUES, product benefits and customer value : a consumption behavior approach, *Advances in Consumer Research*, 381-388.

LATOUR S.A. et PEAT N.C. (1979), Conceptual and methodological issues in consumer satisfaction research, *Advances in Consumer Research*, 431-437.

LADHARI, RIADH.(2005) « La satisfaction du consommateur, ses déterminants et ses conséquences ». *Revue de l'université de Moncton* 36, no 2 17 41-201.

LECOUTRE.B (1984) Réinterprétation fiducio-bayésienne du test F de l'analyse de la variance. *L'Année de la psychologie*.

LECOUTRE.B,ROUANET H (1981) Deux structures statistiques fondamentales en analyse de la variance univariée et multivariée,*Mathématiques et Science Humaines*, 75,p 71-82

LECOUTRE.B (1984) L'.analyse Bayésienne des comparaisons. Presses universitaires de Lille.

LINDA G. ET OLIVER R.L. et (1980), « Effects of satisfaction and its antecedents on consumer preference and intention. », *Advances in Consumer Research*, 8, 88-93

LAMBERT A.RICHARD (1998), Customer Satisfaction and Future Financial Performance Discussion of Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction *Journal of Accounting Research* Vol. 36 Supplement

LLOSA S. (1996), Contributions à l'étude de la satisfaction dans les services, Thèse pour le doctorat en sciences de gestion, Institut d'administration des entreprises, Université d'Aix-Marseille, décembre.

LOEVINGER, J. (1947). « A systematic Approach to the Construction and Evaluation of test of Ability ». *Psychological Monographs* 61, no 4: 1-49.

LORINO PHILIPPE (2001), le balanced scorecard revisité : dynamique stratégique et pilotage de performance exemple d'une entreprise énergétique, 22ÈME CONGRES DE L'AFIC, France

LORINO PHILLIPE (2001) : Méthodes et pratiques de la performance , Editions d'Organisation, Paris, 1997, 2001.

MACINTOSH, NORMAN (1994). *Management accounting and control systems*. Wiley.

MAZURSKY D. ET GEVA A., (1989), "Temporal decay in satisfaction – purchase intention relationship. », *Psychology and Marketing*, 6, 3, 211-227.

MOORE, HOPKINS & HOPKINS (1998) "Quality and empowerment programs: dual paths to customer satisfaction," *Managing Service Quality*, vol. 8(2), p. 133-141

MORGAN, STRATI, MCGUINNESS, DENNIS (2000) "Determinants of customer brand switching behavior," *AM 2000 Conference Proceedings of Academy of Marketing Annual Conference*, Derby

MORGAN R.M. et HUNT S.D. (1994), The commitment – trust theory of relationship marketing, *Journal of Marketing*, 58, 20-38.

MILLS, J. E. ., & STEPCHENKOVA, S. (2010). Destination Image: A Meta-Analysis of 2000-2007 Research., p. 575-609.

MILLER G.A. (1956), « The magical number seven, plus or minus two: some limits on our capacity for processing information », *Psychology*, volume 63, pp.81–97

MITTAL V., ROSS W.T., BALDASARE P.M. (1998), « The Assymetric impact of negative and positive attribute level performance on overall satisfaction and repurchase intentions. », *Journal of Marketing*, 62, 1, 33-47.

MITTAL V., KUMAR P. ET TSIROS M. (1999), « Attribute – level performance, satisfaction, and behavioral intentions over time : a consumption-system approach. », *Journal of Marketing*, 4, 88-101.

MUFFATTO, PANIZZOLO (1995) “A process-based view for customer satisfaction,” *International Journal of Quality & Reliability Management*, vol. 12(9), p. 154-169

NAUMANN, HOISINGTON (2001) *Customer Centered Six Sigma: Linking Customers, Process Improvement, and Financial Results*. Milwaukee: ASQ Quality Press

NGOALA G. (1988), « Epistémologie et théorie du marketing relationnel », *Actes du 17ème Congrès International de l'Association Française du Marketing*, 473-495.

NGOBO P-V, (1997), *Les standards de comparaison dans les modèles de satisfaction des consommateurs : Structure, dynamique et conséquences*, Thèse de doctorat de sciences de gestion, Université de Montpellier II.

.NGOBO PAUL VALENTIN (1998), "Les standards de comparaison dans la gestion de la satisfaction des clients" *Décisions Marketing* n° 13 (Janvier - Avril), pp. 57-65

NGOBO PAUL VALENTIN & RAMAROSSON ANDY (2005), "Satisfaction et performance de l'entreprise : synthèse et extension", *Décisions Marketing*, oct-Dec, pp. 75-84

NGOBO P-V. (1997), *Qualité perçue et satisfaction des consommateurs : un état des recherches*, *Revue Française du Marketing*, 163, 67-79.

NGOBO P-V. (2000), *Satisfaction des clients et part de marché de l'entreprise : un réexamen au regard de récentes avancées théoriques*, *Recherche et Applications en Marketing*, 15, 2, 21-41.

NGUYEN N. (1992), *Un modèle explicatif de l'évaluation de la qualité d'un service : une étude empirique*, *Recherche et Applications en Marketing*, 6,2 83-98.

NYER (2000) “An investigation into whether complaining can cause increased customer satisfaction, *Journal of Consumer Marketing*, vol. 17(1), p. 9-19

MITTAL V., KUMAR P. ET TSIROS M. (1999), « Attribute – level performance, satisfaction, and behavioral intentions over time: a consumption-system approach. », *Journal of Marketing*, 4, 88-

MOUTINHO, LUIZ AND GOODE, MARK (1995), “Gender Effects to the Formulation of Overall Product Satisfaction: A Multivariate Approach,” *Journal of International Consumer Marketing*, 8 (1), 71-83.

O’SULLIVAN, DON, MARK C. HUTCHINSON, ET VINCENT O’CONNELL. (2009). Empirical evidence of the stock market's (mis) pricing of customer satisfaction. *International Journal of Research in Marketing* 26, no. 2 p154-161..

O’SULLIVAN, DON, ET JOHN MCCALLIG.(2009). Does customer satisfaction influence the relationship between earnings and firm value? *Marketing Letters* 20, no 4 , p337-35.

OLIVER RL. (1980), A cognitive Model of the Antecedents and Consequences of Satisfaction Decision, *Journal of Marketing Research*, 17, 460 - 467.

OLIVER R.L., LINDA G. (1980), Effect of satisfaction and its antecedants on consumer preference and intention, *Advances in Consumer Research*, 8, 88-93.

OLIVER RL. (1981), Measurement and Evaluation of Satisfaction Process in retail Store, *Journal of Retailing*, 57, 3, 25-48.

OLIVER R.L. ET WESTBROOK R.A. (1993), Profiles of consumer emotions and satisfaction in ownership and usage, *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 6, 12-27.

OLIVER (1993), Cognitive, affective, and attribute bases of satisfaction response, *Journal of Consumer Research*, 20, 418-430.

OLIVER (1994), Conceptual issues in the structural analysis of consumption emotion, satisfaction, and quality : evidence in service setting, *Advances in Consumer Research*,. 21, 16-22.

OLIVER R.L. (1997), Satisfaction : a behavioral perspective on the consumer, *Mc Graw-Hill International Editions*.

OLIVER R.L. (1989), Processing of the satisfaction response in consumption, *Journal of Consumer Satisfaction/dissatisfaction and Complaining Behavior*, 2, 1-26.

OLSHAVSKY R. W., MILLER J.A. (1972), Consumer expectations, products performance, and perceived product quality, *Journal of Marketing Research*, 9, 19 - 21.

OLSON J.C. et DOVER P.A. (1976), Effects of expectation creation and disconfirmation on belief elements of cognitive structure, *Advances in Consumer Research*, 168-175.

OLSON J-C. et DOVER P.A. (1979), Disconfirmation of consumer expectations through product trial, *Journal of Applied Psychology*, 64, 2, 179-189.

OUCHI, WILLIAM. (1979). « A conceptual framework for the design of organizational control mechanisms ». *Management Science* 25, no 9 (1979): 833-848.

PARASURAMAN A., ZEITHAML V.A. et BERRY L.L. (1985), A conceptual model of service quality and its implications for future research , *Journal of Marketing*, 49, 41-50.

PARASURAMAN A., ZEITHAML V.A. et BERRY L.L. (1988), SERVQUAL : a multi-item scale for measuring consumer perceptions of service quality, *Journal of Retailing*, 64,1, 13-37.

PALMATIER, R. W., DANT, R., GREWAL, D., & EVANS, K. R. (2007, mars). Les facteurs qui influencent l'efficacité du marketing relationnel, une méta-analyse. *Recherche et application en marketing*.

PETTY R.E., UNNAVA R.H. et STRATHMAN A.J. (1991), Theories of attitude change, *Handbook of Consumer Behavior*, Prentice Hall, 241-280.

PLICHON V. (1998), La nécessité d'intégrer les états affectifs à l'explication du processus de satisfaction du consommateur, *Actes du 15^{ème} Congrès International de l'Association Française du Marketing*, 671-694.

PRIM I. (1998), Qualité de service et satisfaction : une perspective de long terme pour une approche réconciliatrice, *Actes du 15^{ème} Congrès International de l'Association Française du Marketing*, 957-972.

PFUFF, M. ET BLIVICE, S. (1977), "Socio-Economic Correlates of Consumer and Citizen Dissatisfaction, and Activism," in R.L. Day (Ed.), *Consumer Satisfaction, Dissatisfaction, and Complaining Behavior*, 115-123. Bloomington, IN: Indiana

PETERSON, ROBERT A., AND WILLIAM R. WILSON (1992). "Measuring Customer Satisfaction: Fact and Artifact," *Journal of the Academy of Marketing Science*, 20 (1), 61-71.

PHILEMON OYEWOLE (2001): Consumer's Socio-Demographic Characteristics and Satisfaction with Services in the Airline Industry, *Services Marketing Quarterly*, 23(2)

POINCELOT E. ET WEGMANN G. (2005) "Utilisation des critères non financiers pour évaluer ou piloter la performance : analyse théorique », *Comptabilité Contrôle Audit*, Tome 11, Vol. 2, pp. 109-125.

RAY DANIEL, (2004). *Mesure de la Satisfaction Clients: pratiques et besoins actuels des entreprises, positionnement des offres et principaux enjeux méthodologiques*. *Revue Française du Marketing*, 197, 2/5: 1-23.

RAY DANIEL ET DAVID GOTTELAND, 2005. *Mesurer l'asymétrie des impacts des attributs sur la satisfaction : comparaison de la validité convergente de cinq méthodes*. *Recherche et Applications en Marketing*, 20,1: 1-19.

. RAY J.-J., RAY D., 2008. *Modéliser les phénomènes multiniveaux en marketing*. *Recherche et Applications en Marketing*, Vol 23, n°1: 55-80.

REICHHELD (1996) "The Loyalty Effect," Harvard Business School Press, Boston,

-
- REICHHELD F. SASSER W.E. (1990), Zero defections : quality comes to services, *Harvard Business Review*, 68, 105-111
- ROBINSON W.S (1950), Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, p351-357.
- ROBERT C.P (2005), *Le choix bayésien* .Edition Springer
- ROTTER (1967), A new scale for the measurement of interpersonal trust, *Journal of Personality*, 35, 651-665.
- ROSENTHAL, R., & DIMATTEO, M (2001). META-ANALYSIS: Recent Developments in Quantitative Methods for Literature Reviews. *Annual Review of Psychology*, p. 52-59
- ROSSITER J.R (2002), The C-OAR-SE procedure for scale development in marketing, *International Journal of Research in Marketing*,305-335.
- ROUANET H., LEPINE D (1976a). Structures linéaires et analyse des comparaisons, *Mathématiques et Sciences Humaines*, 56, 5-46
- ROUANET H., LEPINE D., PELNARD-CONSIDERE J (1976b). Bayes-fiducial procedures as practical substitutes for misplaced significance testing, in De Gruiter & Van der Kamp (Eds), *Advances in Psychological and Educational Measurement*, pp. 33-50, New York, Wiley
- RUST R.T. ET ZAHORIK A. (1993), Customer satisfaction, customer retention and market share, *Journal of Retailing*, 69, 2, 193-216.
- RUST R.T., ZAHORIK A.J., KEININGHAM T.L. (1995), Return on quality (ROQ) : Making service quality financially accountable, *Journal of Marketing*, 59, 58-70.
- RUST R.T., ZAHORIK A.J. ET KEININGHAM T.L. (1996), *Service marketing*, New York, Harper Collins.
- SAKO M. (1992), *Price, quality and trust. Inter firms relations in Britain and Japan*. Cambridge, Cambridge University Press
- SCHURR P.H.et OZANE J.L. (1985), Influences on exchange processes : buyer's preconception of a seller's trustworthiness and bargaining toughness, *Journal of Consumer Research*, 11, 939-953.
- SELNES F. (1993), An examination of the effect of product performance on brand reputation, satisfaction and loyalty, *European Journal of Marketing*, 27, 9, 19 - 35.
- SINGH J., SIRDESHMUKH D., (2000), Agency and trust mechanisms in consumer satisfaction and loyalty judgments, *Journal of the Academy of Marketing Science*, 28, 1, 150-167.
- S. CEM BAHADIR, BHARADWA, S., & PARZEN, M. (2010). A meta-analysis of the determinants of organic sales growth. *Journal of Research in Marketing*, p. 87-89.

SZYMANSKI, D. M., & HENARD, D. H. (2001). Customer Satisfaction: A Meta-Analysis of the Empirical Evidence. *Journal of the Academy of Marketing Science*, p. 16-20.

SONGTING DONG, MIN DING, RAJDEEP GREWAL , PING ZHAO (2009) : Functional forms of the satisfaction–loyalty relationship, *Intern. J. of Research in Marketing* 28 38–50

SOLEY, LAWRENCE C. AND LEONARD N. REID (1983), “Satisfaction with the Information Value of Magazine and Television Advertising,” *Journal of Advertising*, 12 (3), 27-32.

SIRIEX L., DUBOIS P-L., (1999), Vers un modèle qualité-satisfaction intégrant la confiance? *Recherche et Applications en Marketing*, 14, 3, 1-22.

SWAN J.E. ET COMBS L.J. (1976), Product performance and consumer satisfaction: a new concept, *Journal of Marketing*, 40, 25-33.

SWAN J.E. et TRAWICK I.F. (1981), Disconfirmation of expectations and satisfaction with a retail service, *Journal of Retailing*, 57, 3, 49-67.

TAX S.S., BROWN S.W. ET CHANDRASHEKARAN M., (1998), « Customer evaluations of service complaint experiences : implications for relationship marketing. » *Journal of Marketing*, 62, 4, 60-76.

TRAPPEY, C. (1997). Une méta-analyse du choix de consommateur et de la publicité subliminale. *Recherche et application en marketing*

TÉROUANNETÉROUANNE E.(1998) (corrélation entre variables nominales, ordinales, métriques ou numériques *Math. Inf. Sci. hum*, n°142, , pp.5-16)

TSE David K., NICOSIA Franco M., WILTON Peter C., (1990), Consumer satisfaction as a process, *Psychology and Marketing*, 7, 3, 177-193.

TSE et WILTON (1988), Models of consumer satisfaction formation : an extension, *Journal of Marketing Research*, 25, 204-212.

UELTSCHY, LINDA C. KRAMPF, ROBERT F. (2001), « Cultural Sensitivity to satisfaction and service quality measures », *Journal of Marketing Theory & Practice*; Summer2001, Vol. 9 Issue 3, p14, 18p

VANHAMME J., (1998), La surprise et son influence sur la satisfaciton des consommateurs, du 15ème Congrès de l’Association Française du Marketing, 2, 809-837.

VANHAMME J. (2001), L’influence de la surprise sur la satisfaction des consommateurs : étude exploratoire par journal de bord, *Recherche et Applications en Marketing*, 16, 2, 1-32.

WESTBRROK R.A. (1983), Sources of consumer satisfaction with retail outlets, *Journal of Retailing*, 68-85.

WESTBROOK R. ET REILLY M.D., (1983), « Value-percept disparity : an alternative to the disconfirmation of expectations theory of consumer satisfaction. », *Advances in Consumer Research*, 10, 256-261.

WESTBROK, OLIVER (1991) "The dimensionality of consumption emotion patterns and consumer satisfaction," *Journal of Consumer Research*, vol. 18(1), p. 84-91
73

WHITE J. CHRISTOPHER (2010) The impact of emotion on service quality, satisfaction, and positive word-of-mouth intentions over time *Journal of Marketing Management* Vol 26, N° 5-6 p 381-394

WOODRUFF, GARDIAL, (1996) "Know Your Customer. New Approaches to Understanding Customer Value and Satisfaction," Blackwell Publishers Inc

WOODRUFF (1997) "Customer value: the next source for competitive advantage," *Journal of the Academy of Marketing Science*, vol. 25(2), Spring, p. 139-153

Yi Y. (1990), A critical Review of Customer satisfaction, in *Review of Marketing*, V.A. Zeithaml (ed.), American Marketing Association, Chicago, Il., 68 – 123.

ZAIRI (1996) "Effective Benchmarking: Learning from the Best," Champ and Hall, London

ZAIRI (2000) "Managing customer satisfaction: a best practice perspective," *The TQM Magazine*, vol.12(6), p. 389-394

ZALTMAN G. et MOORMAN C., (1988), The importance of personal trust in the use of research, *Journal of Advertising Research*, 16-24.

ZEITHALM V. (1981), How consumer evaluation process differ between goods and services, *American Marketing Association Proceedings*, 186-190

ZEITHAML (1988), Consumer perceptions of price, quality and value : a means-end model synthesis of evidence, *Journal of Marketing*, 52, 2-22.

ZEITHAML, BITNER (2000) "Service Marketing: Integrating Customer Focus Across the Firm," Irwin McGraw-Hill, Boston, MA

Liste des figures et tables

Figure 1 L'évolution des indices ACSI.....	12
Figure 2 Évolution de l'ACSI par firme suivant 10 secteurs	14
Figure 3: Le Balanced Scorcard de Kaplan et Norton	24
Figure 4 : Liaison concept et indicateur	27
Figure 5 : Situation de concordance.....	28
Figure 6: Situation discordante	28
Figure 7: Modèle de disconfirmation	43
Figure 8: Le modèle A.C.S.I, source article de D. Ray (2004)	53
Figure 9: Différence dans les niveaux de satisfaction : âge des répondants	69
Figure 10 : Différence dans les niveaux de satisfaction : le sexe des répondants.....	70
Figure 11: Sensibilité des critères de satisfaction sur 16 vagues d'enquêtes	73
Figure 12:Le rôle des deux erreurs systématique et aléatoire	90
Figure 13: Actions des facteurs de contingence	95
Figure 14 : Répartition géographique des magasins du baromètre B1	106
Figure 15: Distribution des notes moyennes de satisfaction par magasin	107
Figure 16: Répartition géographique des notes moyennes de satisfaction par magasin	107
Figure 17: Répartition géographique des magasins du baromètre	110
Figure 18: Dynamiques des items du baromètre B2.....	111
Figure 19 : Extrait du Baromètre B1	112
Figure 20: Extrait du baromètre B2.....	112
Figure 21 : conjugaison de la méta-analyse et procédures fiducio-bayésiennes	134
Figure 22 : Répartition géographique des magasins du baromètre B1 en fonction de la satisfaction moyenne.....	141
Figure 23: Variation du coefficient de corrélation multiple, sur l'ensemble des 221 magasins.....	145
Figure 24 : Variations des coefficients des quatre facteurs par magasin	147
Figure 25: Répartition géographique des valeurs des coefficients de corrélation multiple suivant les magasins.....	148
Figure 26 : Distribution des valeurs de l'importance de l'effet de la fréquentation	152

Figure 27: Distribution des valeurs de l'importance de l'effet du Pays	154
Figure 28 : Dispersion des valeurs des variances intra-magasin sur l'ensemble des 221 magasins	160
Figure 29 : Répartition géographique des variances selon les magasins	161
Figure 30 : Distribution des valeurs de l'importance de l'effet de l'âge sur la variance intra magasin	163
Figure 31 : Distribution des valeurs de l'importance de l'effet de la fréquentation sur la variance intra magasin	164
Figure 32 : Distribution des valeurs de l'importance de l'effet du Pays sur la variance intra magasin	165
Figure 33 : Répartition géographique des magasins du baromètre B2	170
Figure 34 : Évolution des indices de distorsion relative au niveau agrégé (ensemble du parc) suivant les 16 vagues	171
Figure 35 : Valeurs des coefficients de variation des indices de distorsion relative, sur les 16 vagues pour chaque magasin	172
Figure 36 : Importance des effets région (à gauche) et vague (à droite)	174
Figure 37 Évolutions des scores de satisfaction enregistrés par magasin	178
Figure 38 : Structure hiérarchique testée	179
Figure 39:Extrait des données d'analyse	179
Figure 40: Benchmarks 2012 du NPS par industrie, source site web SATMETRIX	188
Figure 41 : Benchmarks 2012 par industrie des scores ACSI, source site web ACSI	188
Figure 42:Les deux composantes de la croissance	190
Figure 43 : Schéma de liaison entre la satisfaction et la recommandation au niveau d'un individu	193
Figure 44: Les sens d'inclusion (d'implication) entre les populations des satisfaits insatisfaits et promoteurs détracteurs	197
Figure 45: Dépendance de l'indicateur de satisfaction, aux niveaux d'analyse.	203
Figure 46 : Relations action entreprise et avis clients. D.RAY (2004)	210
Figure 47 : Les formes de construction: attributs et variables latentes	212
Tableau 1 Informations disponibles par secteurs	68
Tableau 2 : Table des effectifs par secteurs	68

Tableau 3: Résultats des analyses de variance par secteurs	71
Tableau 4 : Items de satisfactions baromètre B1	104
Tableau 5 : Ensemble des informations or notes de satisfaction	105
Tableau 6: Distribution des notes moyennes de satisfaction par magasin	106
Tableau 7: Items de satisfaction baromètre B2	109
Tableau 8: Dispersion des valeurs du coefficient de corrélation multiple	144
Tableau 9: Effets de l'âge des répondants sur le coefficient de corrélation multiple entre magasins	150
Tableau 10 : Effets du sexe des répondants sur le coefficient de corrélation multiple entre magasins	150
Tableau 11 : Effets de la fréquentation du magasin sur le coefficient de corrélation multiples entre magasins	151
Tableau 12 Effets du pays sur la variation du coefficient de corrélation multiple entre magasins	153
Tableau 13 Effets de la région sur la variation du coefficient de corrélation multiple entre magasins	154
Tableau 14 Classement par taux ou moyenne exemple A	159
Tableau 15 Classement par taux ou moyenne exemple B	159
Tableau 16 : Effets de l'âge des répondants sur la variance	162
Tableau 17: Effets du sexe des répondants sur la variance	163
Tableau 18: Effets de la fréquentation des répondants sur la variance	164
Tableau 19: Effets du pays des répondants sur la variance	165
Tableau 20: Effets de la région et du pays des répondants sur la variance	166
Tableau 21 : Un test de saisonnalité des λp	171
Tableau 22 : Effet de la région sur la sensibilité de λp	174
Tableau 23 : Effet de la région sur la sensibilité de λp	175
Tableau 24: Modèle hiérarchique avec niveau magasin.	181
Tableau 25 : Modèle hiérarchique avec niveau magasin. La région variable explicative.	182
Tableau 26 : Modèle hiérarchique avec niveau magasin et région	184
Tableau 27 : Structure de croisement NPS et indicateur de satisfaction	194
Tableau 28 Résultats d'implication NPS et Taux de satisfaction	196