

1 2 F N 1 7 8 1 2 F N 1 7 8 1 2
A 7 9 C 1 1 2 F N 1 7 8 1 2 4 M 1 2
8 5 7 B A 7 9 C 1 3 1 4 M 1 2 5 3
1 1 1 2 7 9 1 1 3 7 2 Y 8 5 3 1 7 8 1 2
A 7 9 C 1 1 1 1 1 1 1 1 1 1 4 M 1 2
8 5 7 B 1 1 1 1 1 1 1 1 1 1 5 3
1 1 1 3 7 2 Y 8 1 1 1 1 1 1 1 1 1 1 J
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 2 F N 1 7 8 1 2
A 7 9 C 1 3 1 4 M 1 2
8 5 7 B 1 3 1 4 M 1 2 5 3
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

L'utilisation des méthodes PLS dans les études de satisfaction



L'utilisation des méthodes PLS dans les études de satisfaction

Introduction

Introduction : le contexte

■ Les données issues de questionnaires de satisfaction ont certaines particularités :

- Il y a souvent beaucoup de questions très liées entre elles (corrélations importantes)
- Les questions sont souvent regroupées par thème ; chaque thème exprimant un concept . On cherche alors à simultanément décrire le contenu de chaque thème et explorer les liaisons entre les thèmes, en particulier celui qui reflète la satisfaction. Il faut alors disposer de modèles à plusieurs équations.
- La notion de satisfaction elle-même n'est pas souvent réductible à une seule variable mesurable

Introduction : le contexte

- La description de ce type de données ne pose pas de problèmes particuliers (analyses factorielles avec choix des thèmes actifs et illustratifs, typologies ...)
- Les méthodes prédictives classiques s'appliquent mal ou difficilement à ces données

Sommaire

Présentation de deux exemples d'application avec :

- La régression PLS
- L'approche PLS (*utilisation de la méthode PLS pour la modélisation de relations structurelles sur variables latentes*)

Conclusion

Les objectifs

■ La régression PLS répond à deux objectifs :

- *Prédiction* : Relier un bloc de variables à expliquer Y à un bloc de variables explicatives X.
- *Description* : Décrire simultanément les relations entre les variables composant le bloc des X, les variables composant le bloc des Y, et les liaisons entre X et Y

Données d'exemple

- Le questionnaire porte sur 6 items, plus une question de satisfaction globale (les questions sont sur une échelle de 1 à 10).
- Les items peuvent être séparés en deux groupes
 - Un groupe 'IMAGE' (Manufacturer image, Salesforce image)
 - Un groupe 'STRATEGY' (Delivery Speed, Price flexibility, Service, Usage Level)
- Le but est de relier ces items à la satisfaction globale

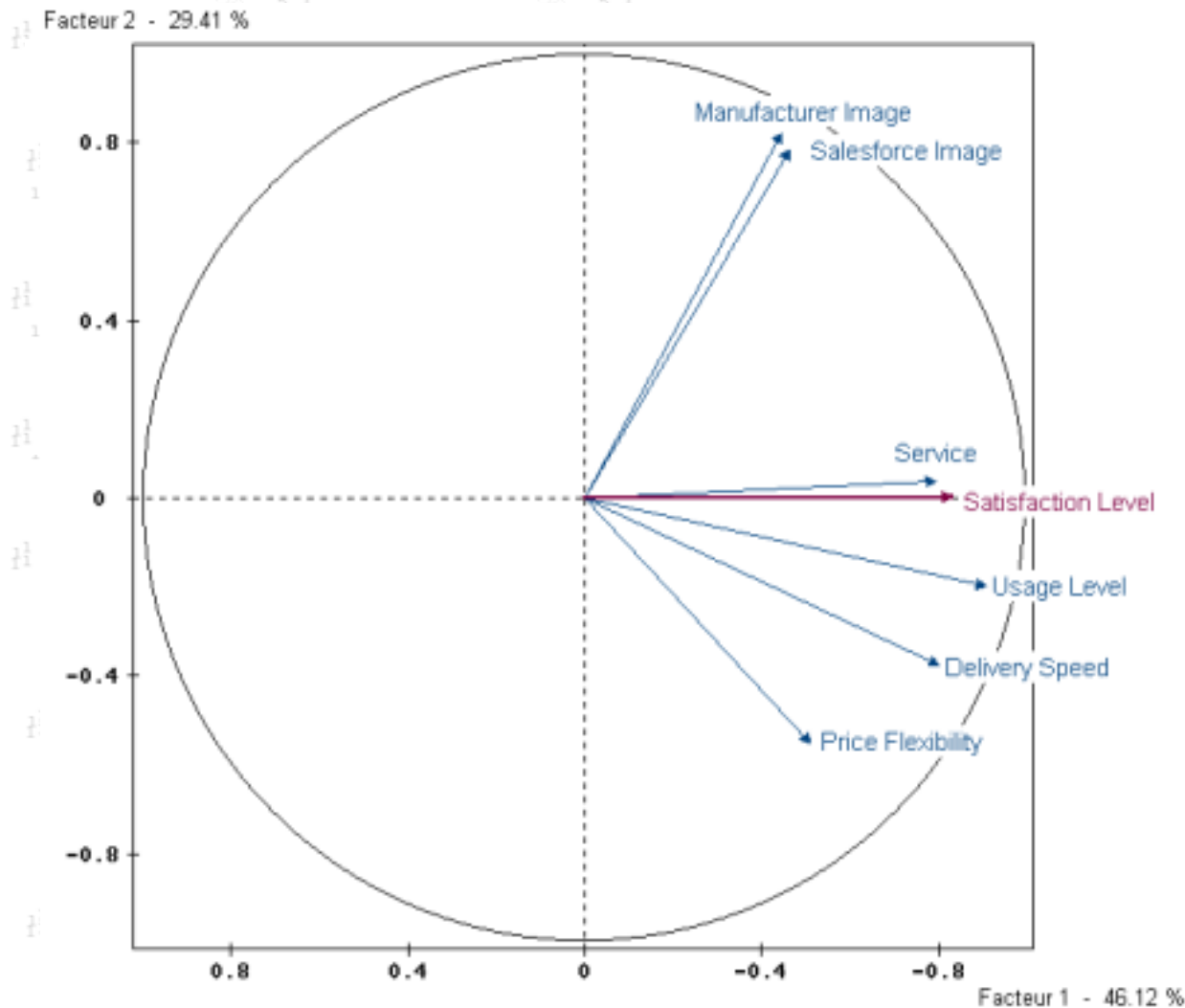
Delivery Speed	Price Flexibility	Manufacturer Image	Service	Salesforce Image	Usage Level	Satisfaction Level
4.1	6.9	4.7	2.4	2.3	3.2	4.2
1.8	6.3	6.6	2.5	4.0	4.3	4.3
3.4	5.7	6.0	4.3	2.7	4.8	5.2
2.7	7.1	5.9	1.8	2.3	3.2	3.9
6.0	9.6	7.8	3.4	4.6	5.8	6.8
1.9	7.9	4.8	2.6	1.9	4.5	4.4
4.6	9.5	6.6	3.5	4.5	4.6	5.8
1.3	6.2	5.1	2.8	2.2	4.4	4.3
5.5	9.4	4.7	3.5	3.0	6.3	5.4
4.0	6.5	6.0	3.7	3.2	5.4	5.4
2.4	8.8	4.8	2.0	2.8	3.2	4.3
3.9	9.1	4.6	3.0	2.5	4.7	5.0
2.8	8.1	3.8	2.1	1.4	3.9	4.4
3.7	8.6	5.7	2.7	3.7	3.8	5.0
4.7	9.9	6.7	3.0	2.6	5.4	5.9

Corrélations entre la satisfaction globale et les items

- Toutes les corrélations sont significatives
- Les corrélations nous donnent la hiérarchie des liaisons des items avec la satisfaction globale
- Les items du groupe 'strategy' sont plus liés à la satisfaction que ceux du groupe 'image'.

Pearson Correlation Coefficients, N = 100 Prob > r under H0: Rho=0	
	X10
X9 Usage Level	0.71070 <.0001
X1 Delivery Speed	0.65063 <.0001
X5 Service	0.63123 <.0001
X3 Price Flexibility	0.52481 <.0001
X4 Manufacturer Image	0.47593 <.0001
X6 Salesforce Image	0.34091 0.0005

L'analyse en composantes principales



- La premier plan reconstitue 75% de l'inertie
- On voit clairement les deux groupes de variables 'Image' et 'Stratégie'
- La satisfaction est plus liée à la stratégie qu'à l'image.

Les résultats de la régression multiple

$$R^2 = 0.795, F = 60.37 \text{ Sig.} = 0.0001$$

Parameter Estimates							
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Standardized Estimate
Intercept	Intercept	1	-0.84270	0.35373	-2.38	0.0192	0
X1	Delivery Speed	1	0.10704	0.04806	2.23	0.0284	0.16524
X3	Price Flexibility	1	0.30258	0.04727	6.40	<.0001	0.49036
X4	Manufacturer Image	1	0.41940	0.05959	7.04	<.0001	0.55463
X5	Service	1	0.45278	0.10239	4.42	<.0001	0.39758
X6	Salesforce Image	1	-0.19944	0.08593	-2.32	0.0225	-0.17969
X9	Usage Level	1	-0.03058	0.09275	-0.33	0.7424	-0.03213

- La variable 'Usage level' n'est pas significative alors que c'est la plus corrélée à la satisfaction globale.
- Les coefficients pour 'Salesforce image' et 'Usage Level' sont négatifs alors que ces variables sont corrélées positivement à la satisfaction globale.

Les résultats de la régression PLS

(modèle à une seule composante)

1. Équation liant la satisfaction aux items (données centrées réduites) :

$$R^2 = 0.710$$

$$satisfaction = \left(\begin{array}{l} 0.237 \text{ speed} + 0.191 \text{ price flexibility} + 0.173 \text{ image} \\ + 0.230 \text{ service} + 0.124 \text{ salesforce image} + 0.259 \text{ usage level} \end{array} \right)$$

- L'équation qui lie la satisfaction aux items est cohérente dans les signes et les valeurs relatives des coefficients (proportionnels aux corrélations, puisque on a une seule composante PLS)
- Le R^2 du modèle n'est pas trop loin de celui de la régression (0.71 contre 0.79).

Les résultats de la régression PLS

2. L'équation liant la satisfaction aux items exprimée sur les données d'origine donne :

$$satisfaction = \left(\begin{array}{l} 0.154 \text{ speed} + 0.118 \text{ price flexibility} + 0.131 \text{ image} \\ + 0.262 \text{ service} + 0.138 \text{ salesforce image} + 0.247 \text{ usage level} + 0.3295 \end{array} \right)$$

- On peut l'utiliser pour prévoir la satisfaction d'un individu à partir de ses notes
- Ses valeurs ne sont pas très différentes de l'équation sur les données centrées réduites, car les échelles et les dispersions dans les notes sont voisines

Les résultats de la régression PLS

3. Description des questions formant le bloc 'X' (R² des items sur la première composante)

Number of Extracted Factors	Percent Variation Accounted for by Partial Least Squares Factors										
	Model Effects							Dependent Variables			
	X1	X3	X4	X5	X6	X9	Current	Total	X10	Current	Total
1	63.300	28.368	21.292	59.987	22.011	81.353	46.052	46.052	71.102	71.102	71.102

- La première composante PLS reconstitue 63% de X1 (delivery speed), 28% de X3 (price flexibility), 21% de X4 (manufacturer image) ...
- Les items du bloc 'image' (X4 et X6) sont moins bien reconstitués que ceux du bloc 'strategy' (X1 X3 X5 X9)
- Le modèle reconstitue 46.05% de l'inertie totale du bloc des 'X'

Conclusion : la régression PLS

- La régression multiple est orientée vers l' "explication" de y et cela peut se faire au détriment de certaines variables (coefficients non significatifs, signes aberrants ...)
- L'ACP réalise la meilleure description du bloc des X , mais cela se fait indépendamment de y
- La régression PLS est une méthode orientée simultanément vers la description des X et l' "explication" de Y ; l'utilisateur obtient ainsi :
 - Un bon ajustement à y
 - Une meilleure cohérence dans les signes et les importances relatives des coefficients
 - Une description satisfaisante du bloc des X .

Modélisation de relations structurelles sur variables latentes

■ La modélisation de relations structurelles sur variables latentes met en jeu 3 composantes :

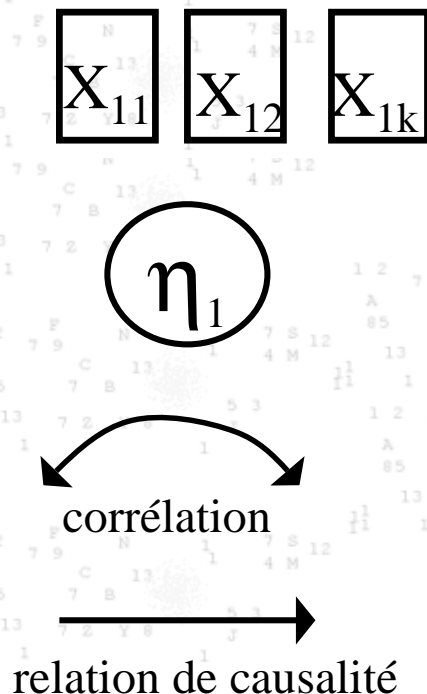
- Les variables manifestes (ou indicateurs, ou variables mesurées)
- Les variables latentes (ou concepts, facteurs)
- Les relations entres variables (corrélations, covariances)

Modélisation de relations structurelles sur variables latentes

- **En combinant ces 3 composantes, l'utilisateur peut avec ce type de modèles :**
 - Spécifier simultanément plusieurs relations entre les variables (les modèles consistent en une estimation de séries de régressions séparées mais interdépendantes). L'utilisateur spécifie ces relations dans le modèle structurel
 - Incorporer des variables qu'il n'est pas possible de mesurer directement. Une variable latente est une variable ou un concept que l'on peut seulement approcher par des variables mesurées

Modélisation de relations structurelles sur variables latentes

- La représentation des modèles se fait traditionnellement sous forme de schémas structurels :



- Les variables manifestes sont représentées sous forme de carrés (pour des études basées sur des questionnaires, chaque indicateur représente une question)
- Les variables latentes sont normalement représentées comme des cercles
- Les relations sont représentées avec des flèches ; des flèches simples pour indiquer une causalité, des flèches courbes indiquant les deux extrémités pour des corrélations.

L'approche PLS

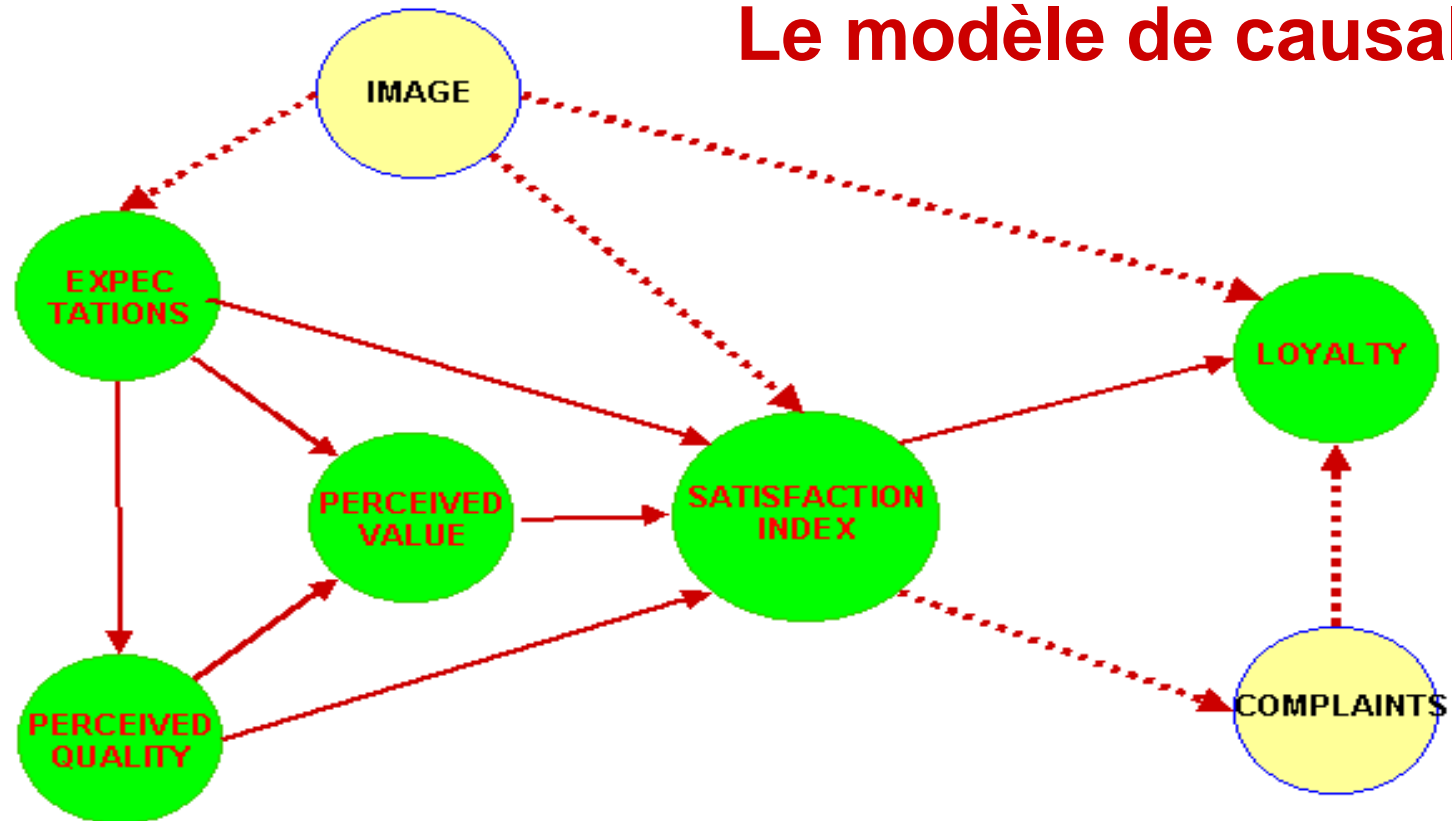
- **L'approche PLS partage ces caractéristiques avec les autres techniques de modélisation SEM classiques.**
- **Par rapport aux autres techniques, elle permet de travailler sur des données individuelles et donc faire de la prévision (les techniques classiques comme LISREL travaillent sur des matrices de covariance et l'on a plus accès aux données individuelles)**
- **On peut travailler avec beaucoup de variables (compatible avec une approche exploratoire)**

Estimation de l'indice ECSI

(cette partie reprend les résultats de MP Bayol, C Tellier, M Tenenhaus, CISIA•CERESTA, HEC – 2001)

- L'indice européen ECSI est un baromètre de satisfaction des consommateurs. Cet indice est inspiré d'un modèle suédois et est compatible avec l'indice américain de Claes FORNELL.
- Il met en œuvre 7 variables latentes liées entre elles. Le modèle est basé sur des théories du comportement des consommateurs et est applicable à de nombreux secteurs.
- Parmi ces 7 variables 5 sont obligatoires et 2 optionnelles

Le modèle de causalité



- Les 5 concepts obligatoires sont en vert, les 2 optionnels sont en jaunes, les flèches représentent les relations de causalité
- La structure globale constitue le *modèle ECSI*

Les équations du modèle

1. Relations entre les variables manifestes et les variables latentes :

$$x_{jh} = \lambda_{jh0} + \lambda_{jh} \xi_j + \varepsilon_{jh}$$

- Chaque variable manifeste est liée à sa latente par une régression simple (chaque item reflète une partie de sa latente)

2. Relations entre les variables latentes :

$$\text{Customer Expectation} = \beta_{20} + \beta_{21} \text{Image} + \zeta_2$$

$$\text{Perceived Quality} = \beta_{30} + \beta_{32} \text{Customer Expectation} + \zeta_3$$

$$\text{Perceived Value} = \beta_{40} + \beta_{42} \text{Customer Expectation} + \beta_{43} \text{Perceived Quality} + \zeta_4$$

$$\begin{aligned} \text{ECSI} = & \beta_{50} + \beta_{51} \text{Image} + \beta_{52} \text{Customer Expectation} \\ & + \beta_{53} \text{Perceived Quality} + \beta_{54} \text{Perceived Value} + \zeta_5 \end{aligned}$$

$$\text{Customer Complaint} = \beta_{60} + \beta_{65} \text{ECSI} + \zeta_6$$

$$\text{Customer Loyalty} = \beta_{70} + \beta_{71} \text{Image} + \beta_{75} \text{ECSI} + \beta_{76} \text{Customer Complaint} + \zeta_7$$

Corrélations entre les latentes et les manifestes

($\rho > 0.5$)

	Image	Customer expectation	Perceived quality	Perceived value	Customer satisfaction	Complaint	Loyalty
Image1	.717		.571		.539		
Image2	.565						
Image3	.657						
Image4	.791		.571		.543		
Image5	.698		.544		.500		
C_exp1		.689					
C_exp2		.644					
C_exp3		.724					
P_qual1	.622	.537	.778		.661		
P_qual2	.		.651				
P_qual3	.621		.801		.651		
P_qual4	.		.760		.587		
P_qual5	.599		.732		.516		
P_qual6	.551		.766		.539		
P_qual7	.596		.803	.547	.707		
P_val1				.933			
P_val2	.541		.594	.911	.631		.524
C_sat1	.558		.638		.711		
C_sat2	.524		.672		.872		
C_sat3	.613		.684	.588	.884	.547	.610
Complaint			.537		.540	1	
Loyalty1							.854
Loyalty2							
Loyalty3	.528		.537		.659		.869

1/ Relations entre les manifestes et les latentes :

Outer Model

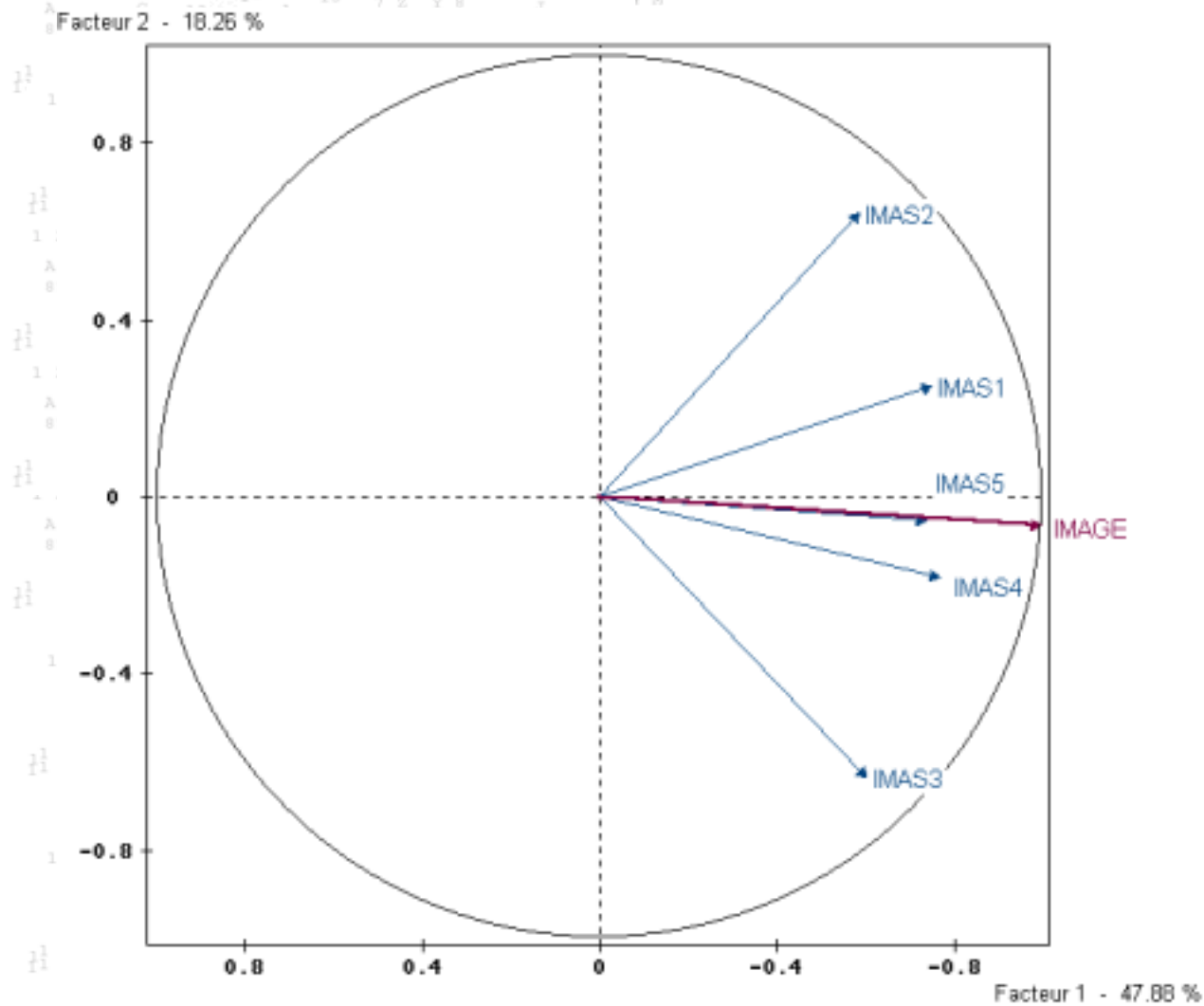
Variable	Weight	Communal
IMAGE outward		
IMAS1	.1610	.5137
IMAS2	.1406	.3200
IMAS3	.1515	.4326
IMAS4	.1960	.6267
IMAS5	.1597	.4869
CUS_EXP outward		
CUEX1	.2566	.4714
CUEX2	.2492	.4152
CUEX3	.2807	.5266
PER_QUAL outward		
PERQ1	.1091	.6052
PERQ2	.0948	.4234
PERQ3	.1306	.6414
PERQ4	.1049	.5781
PERQ5	.0932	.5362
PERQ6	.1055	.5871
PERQ7	.1428	.6444

- On donne ici les poids reliant les latentes standardisées aux manifestes
- Par exemple la variable latente **IMAGE** est estimée par la formule :

$$\begin{aligned}
 \text{IMAGE} = & 0.1610 \text{ IMAS1} \\
 & + 0.1406 \text{ IMAS2} \\
 & + 0.1515 \text{ IMAS3} \\
 & + 0.1590 \text{ IMAS4} \\
 & + 0.1597 \text{ IMAS5}
 \end{aligned}$$

- La communalité entre **IMAGE** et **IMAS1** (i.e la partie que **IMAGE** a en commun avec **IMAS1**) est de **51.37%** [la communalité est l'analogie de la mesure \cos^2 en ACP]
- **IMAS4** est l'item le plus proche du concept latent **IMAGE**

ACP du groupe 'Image'



- Le groupe **IMAGE** est homogène

1/ Relations entre les manifestes et les latentes (suite) :

PER_VAL	outward	
PERV1	.2656	.8698
PERV2	.2749	.8292

ECSI	outward	
CUSA1	.1759	.5050
CUSA2	.2568	.7598
CUSA3	.2928	.7827

CUS_COMP	outward	
CUSCO	.4406	1.0000

CUS_LOY	outward	
CUSL1	.2058	.7307
CUSL2	.0677	.0748
CUSL3	.2499	.7546
=====		

- Pour ECSI, l'estimation donne :

$$ECSI = 0.1759 CUSA1$$

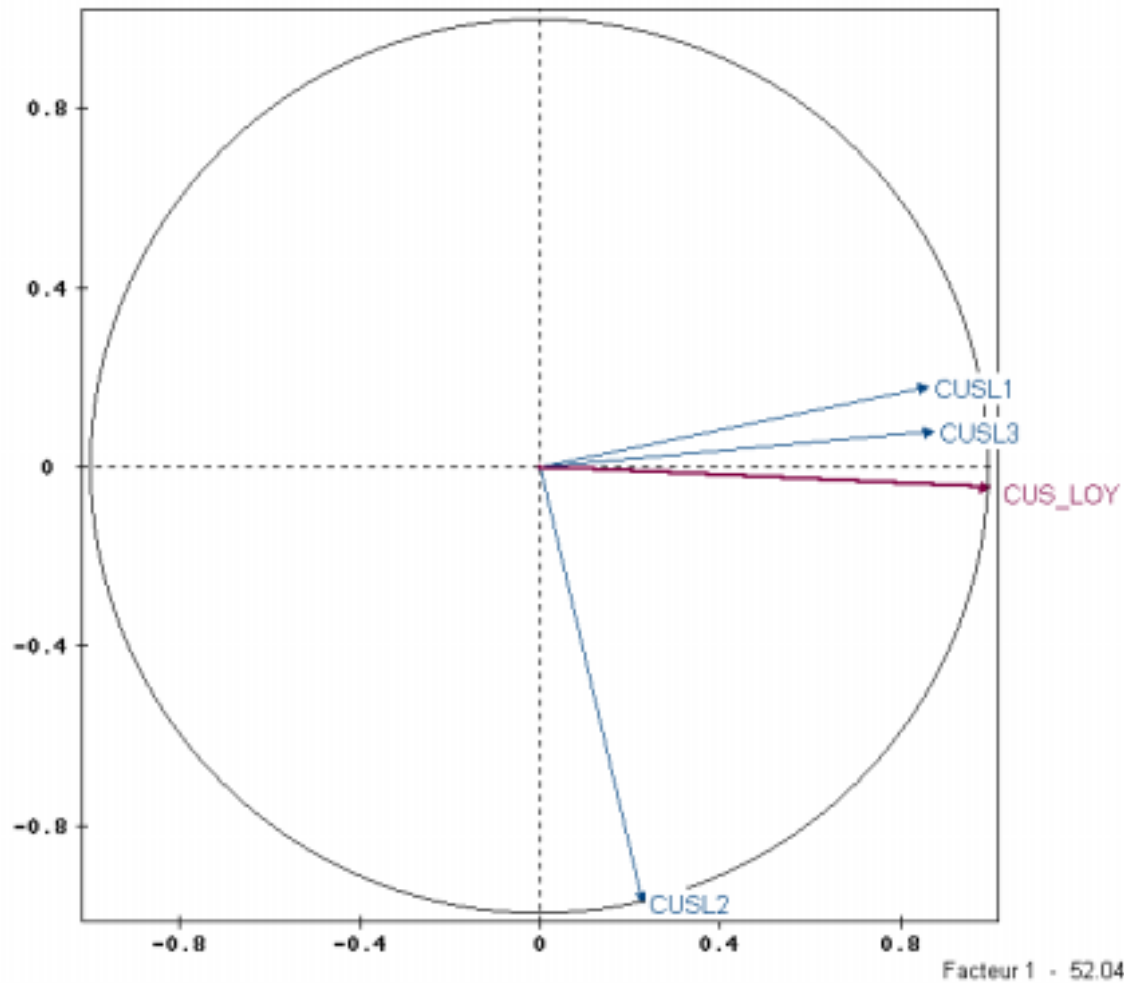
$$+ 0.2568 CUSA2$$

$$+ 0.2928 CUSA3$$

- L'item CUSL2 ne reflète pas le concept latent CUS_LOY (sa communalité est de 7.48% contre plus de 70% pour les deux autres items)

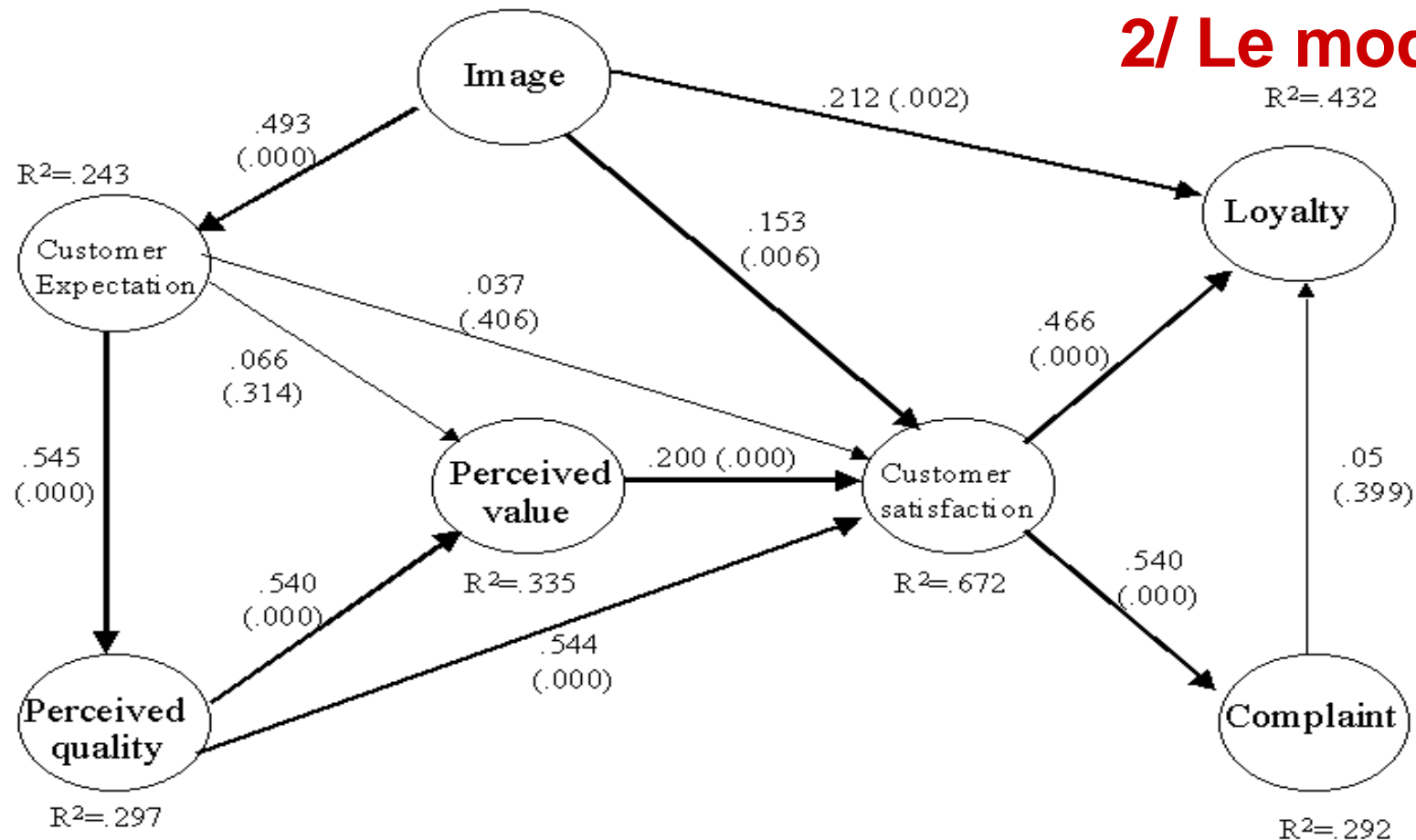
ACP du groupe 'Customer Loyalty'

Facteur 2 - 32.78 %



- Le groupe n'est pas homogène
- La variable CUSL2 ne reflète pas sa latente

2/ Le modèle



- Le modèle de causalité ci-dessus résume les principales liaisons du modèle ECSI. Les coefficients entre deux latentes sont les coefficients des régressions simples et multiples entre les variables. On donne également le R^2 de ces régressions. Ces deux informations figurent dans les sorties du logiciel LVPLS.
- Les niveaux de significativité de ces coefficients sont donnés à côté des coefficients. Dans le cas où le coefficient est significativement différent de 0, les flèches sont en gras.

Commentaires

- Les concepts "Image", "Perceived value" et "Perceived quality" ont un impact significatif sur "Customer Satisfaction". C'est la qualité perçue qui a l'impact le plus important sur la satisfaction : 0.544. (L'image et la valeur perçue ont moins d'impact : 0.200 et 0.153). Il n'est pas étonnant que la qualité réelle des téléphones mobiles soit plus importante pour les utilisateurs que des caractéristiques marketing plus ou moins abstraites.
- Les attentes du client n'ont pas d'impact direct sur la satisfaction.
- La fidélité est un facteur très important dans le secteur de la téléphonie mobile : elle dépend essentiellement de la satisfaction (0.466) et dans une moindre mesure de l'image (0.212).
- On peut noter que les réclamations dépendent de la satisfaction mais n'ont pas d'impact direct sur la fidélité.
- Il faut cependant rester prudent quand à l'interprétation des coefficients non significatifs : cela peut venir d'un problème d'interdépendance entre variables
- Pour pallier ce problème, il serait judicieux d'utiliser la régression PLS (Martens & Næs, 1989, Tenenhaus, 1998) pour évaluer les impacts entre les latentes plutôt que des régressions multiples "classiques".

Validation des résultats

■ La validation des résultats se fait

- Par Validation croisée
- Par Jackknife (c'est la méthode disponible par défaut dans LVPLS)

Conclusion : l'approche PLS

- L'approche PLS est une technique très puissante pour la modélisation des relations structurelles sur variables latentes
- Possibilité d'utiliser les méthodes de type jackknife ou validation croisée pour tester les coefficients
- Possibilité de travailler sur des données individuelles et donc de faire de la prévision
- Ces caractéristiques en font une méthode qui facilite l'analyse des modèles complexes dans des circonstances qui font échouer les autres méthodes (non normalité, multicollinéarité, données qualitatives ...)

Conditions suggérant l'utilisation de la régression PLS

- Avez vous des problèmes de multicollinéarité avec des variables très corrélées entre elles ?
- Avez-vous beaucoup de variables ?

Conditions suggérant l'utilisation de l'approche PLS

- Travaillez-vous avec des modèles mettant en œuvre des variables latentes ?
- Avez vous des problèmes de multicollinéarité avec des variables très corrélées entre elles ?
- Avez-vous beaucoup de variables ?

Références

- **La régression PLS, théorie et Pratique** *Michel TENENHAUS, technip 1999*
- **ECSI '2000 : Use of PLS Path Modelling to estimate the European Consumer Satisfaction Index (ECSI) model** *MP Bayol, C Tellier, M Tenenhaus* (www.cisia.com/cisia/download/PLS_CISIA.doc)
- **LVPLS Program Manual, Version 1.8** LOHMÖLLER JB Zentralarchiv für Empirische Sozialforschung, Köln 1987
- **PLS' 2001 : 2ème Symposium International sur les Méthodes PLS** *Capri Palace, Île de Capri (Naples, Italie), du 1er au 3 octobre 2001*
Co-organisé par le CISIA·CERESTA, HEC et l'Université de Naples "Federico II"

1 2 F N 1 7 8 1 2 F N 1 7 8 1 2
A 7 9 C 1 1 2 F N 1 7 8 1 2 4 M 1 2
8 5 7 B A 7 9 C 1 3 7 8 1 2 5 3
1 1 1 2 7 9 1 1 3 7 2 Y 8 5 3 1 7 8 1 2
A 7 9 C 1 1 1 1 1 1 1 1 1 1 4 M 1 2
8 5 7 B 1 1 1 1 1 1 1 1 1 1 5 3
1 1 1 3 7 2 Y 8 1 1 1 1 1 1 1 1 1 1 J
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 2 F N 1 7 8 1 2
A 7 9 C 1 3 7 8 1 2
8 5 7 B 1 3 7 8 1 2 5 3
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

L'utilisation des méthodes PLS dans les études de satisfaction

Annexes

Les questions du thème 'attentes' telles qu'elles apparaissent dans le questionnaire

Q5 EXPECTATIONS OF THE OVERALL QUALITY OF YOUR BANK

Q5Ba) Now I will ask you to think back about **your expectations** for the overall quality of "*your bank*" (six months ago or at the moment you became customer of this bank). Please give me a rating on a 10 point scale where "1" means your expectations were "very low" and "10" means your expectations were "very high." How would you rate your expectations of the overall quality of "*your bank*" ?

1 2 3 4 5 6 7 8 9 10 99

Q5Bb) Using a 10 point scale on which "1" means "very badly" and "10" means "very well," please tell me to what extent you expected "*your bank*" to provide products and services to meet your personal needs ?

1 2 3 4 5 6 7 8 9 10 99

Q5c) You probably thought about how often things could go wrong with "*your bank*". Using a 10 point scale, on which "1" means "very often" and "10" means "very rarely" how often did you expect that things could go wrong at "*your bank*" ?

1 2 3 4 5 6 7 8 9 10 99

Le questionnaire ECSI

Image (ξ_1)

- a) It can be trusted in what it says and does
- b) It is stable and firmly established
- c) It has a social contribution for the society
- d) It is concerned with customers
- e) It is innovative and forward looking

Customer Expectations of the overall quality (ξ_2)

- a) Expectations for the overall quality of “your mobile phone provider” at the moment you became customer of this provider
- b) Expectations for “your mobile phone provider” to provide products and services to meet your personal need
- c) How often did you expect that things could go wrong at “your mobile phone provider”

Perceived Quality (ξ_3)

- a) Overall perceived quality
- b) Technical quality of the network
- c) Customer service and personal advice offered
- d) Quality of the services you use
- e) Range of services and products offered
- f) Reliability and accuracy of the products and services provided
- g) Clarity and transparency of information provided

Le questionnaire ECSI

Perceived Value (ξ_4)

- a) Given the quality of the **products and services offered by “your mobile phone provider”** how would you rate the **fees and prices that you pay for them?**
- b) Given the **fees and prices that you pay for “your mobile phone provider”** how would you rate the **quality of the products and services offered by “your mobile phone provider”?**

Customer Satisfaction (ξ_5)

- a) Overall satisfaction
- b) Fulfilment of expectations
- c) How well do you think “your mobile phone provider” compares with your ideal mobile phone provider ?

Le questionnaire ECSI

Customer Complaints (ξ_6)

- a) You complained about “your mobile phone provider” last year. How well, or poorly, was your most recent complaint handled
- or
- b) You did not complain about “your mobile phone provider” last year. Imagine you have to complain to “your mobile phone provider” because of a bad quality of service or product. To what extent do you think that “your mobile phone provider” will care about your complaint?

Customer Loyalty (ξ_7)

- a) If you would need to choose a new mobile phone provider how likely is it that you would choose “your provider” again?
- b) Let us now suppose that other mobile phone providers decide to lower their fees and prices, but “your mobile phone provider” stays at the same level as today. At which level of difference (in %) would you choose another mobile phone provider?
- c) If a friend or colleague asks you for advice, how likely is it that you would recommend “your mobile phone provider”?