

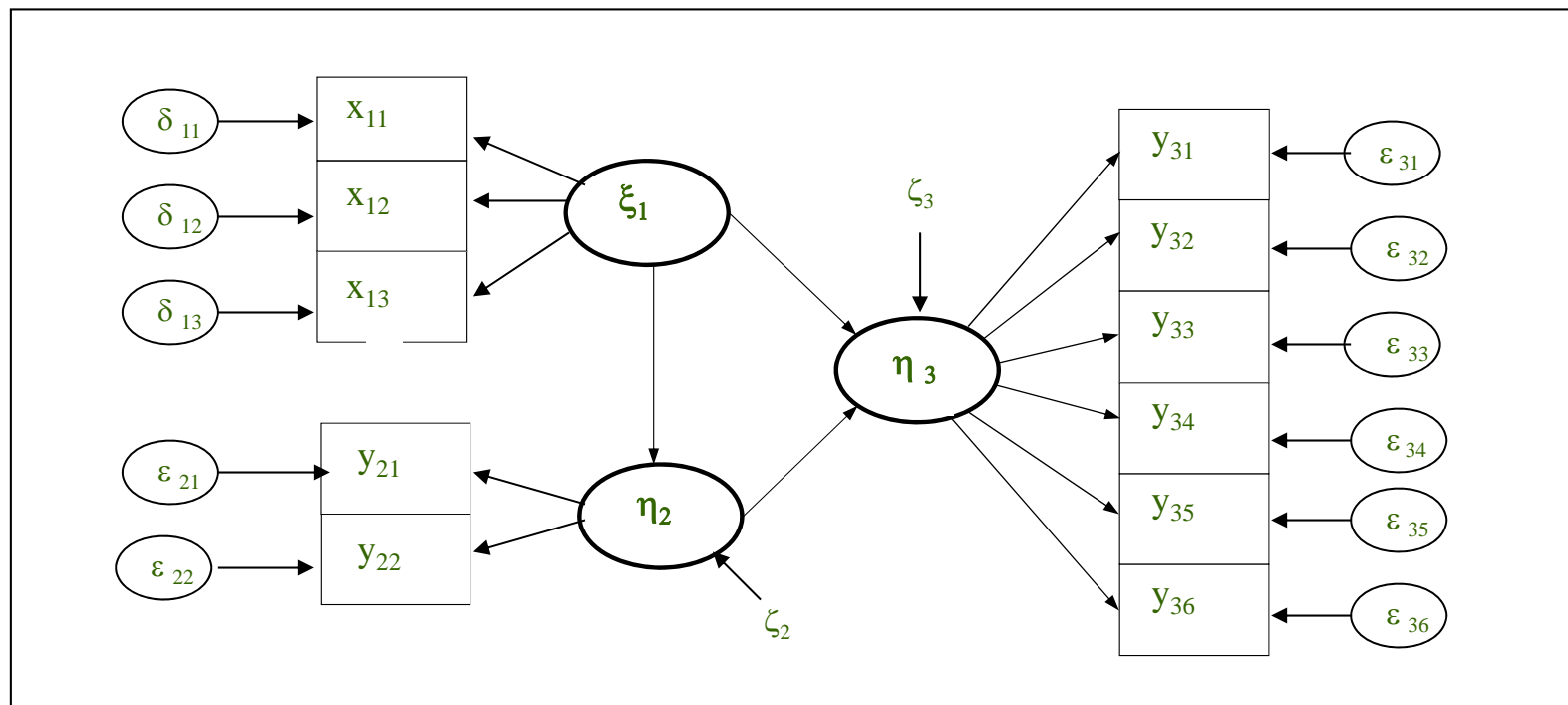


***Les modèles d'équations structurelles à variables
latentes
Applications et exercices***

Emmanuel Jakobowicz
Addinsoft
XLSTAT

1^{er} février 2012

Le modèle structurel



$$\eta_{(m \times 1)} = B_{(m \times m)} * \eta_{(m \times 1)} + \Gamma_{(m \times n)} * \xi_{(n \times 1)} + \zeta_{(m \times 1)}$$

$$y_{(p \times 1)} = \Lambda_{y(p \times m)} * \eta_{(m \times 1)} + \varepsilon_{(p \times 1)}$$

$$x_{(q \times 1)} = \Lambda_{x(q \times n)} * \xi_{(n \times 1)} + \delta_{(q \times 1)}$$

} Eq. du modèle structurel

} Eq. du modèle de mesure

Démonstration: La méthode LISREL

Nous utiliserons les logiciels AMOS (aujourd'hui SPSS) et LISREL

Ce sont les 2 principaux logiciels de traitement de la méthode LISREL

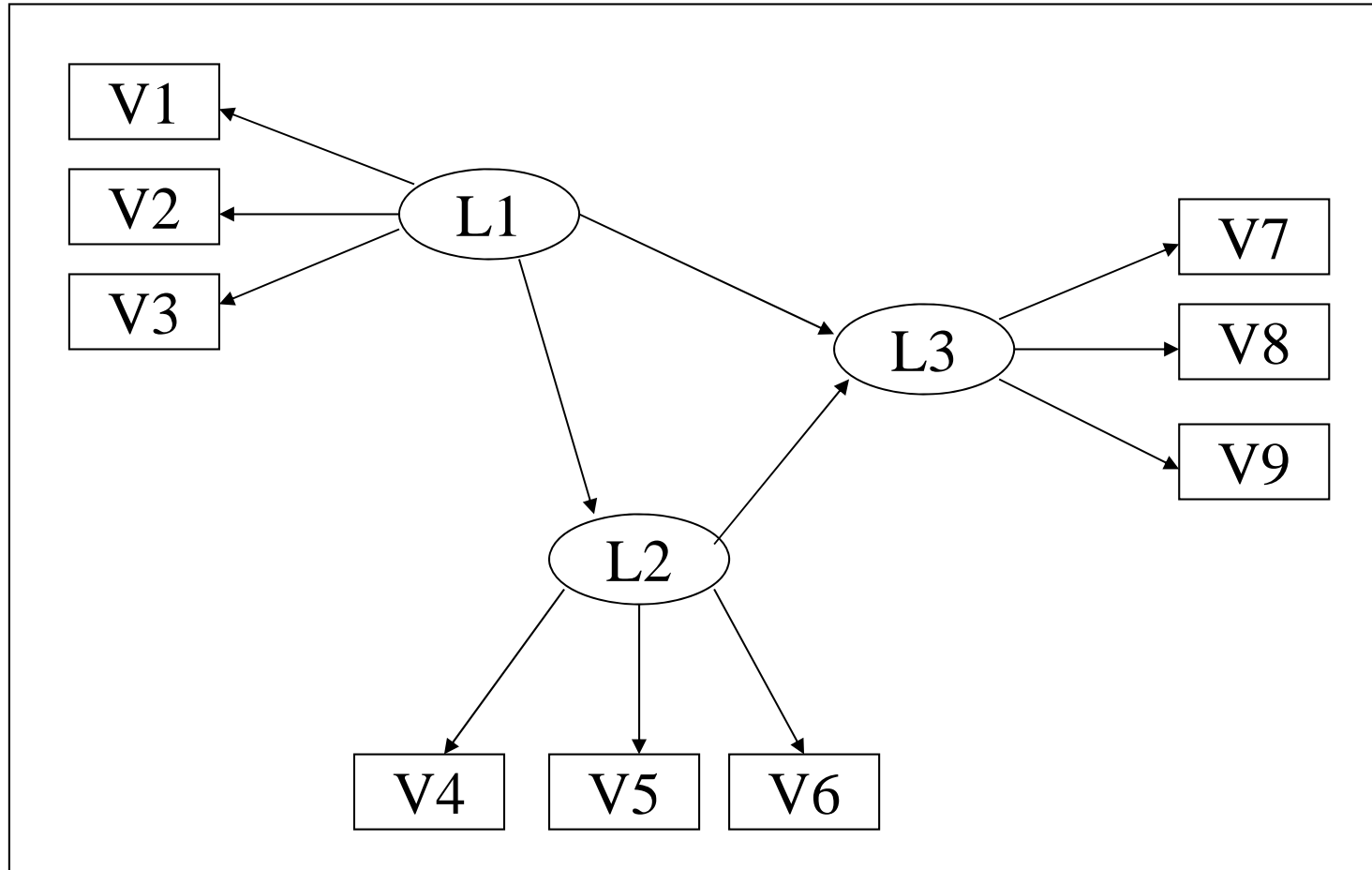
Nous utiliserons un modèle très simple (3 Variables Latentes et 9 Variables Manifestes)

AMOS 19.0 est disponible avec IBM-SPSS.

LISREL 8.8 est disponible à l'essai 15 jours ou il existe une version étudiants gratuite avec une limite dans le taille du modèle

www.ssicentral.com/lisrel/downloads.html

Le modèle



AMOS: Les étapes

Étapes dans l'étude d'un modèle structurel avec Amos Graphics:

1. *Obtenir un jeu de données sans données manquantes avec si possible des données multinormales*
2. *Vérifier l'unidimensionnalité des blocs de variables*
3. *Charger les données* (Files – Data Files...)
4. *Dessiner le modèle LISREL avec les variables latentes (ellipses), les variables manifestes (rectangles) et les erreurs de mesure (cercles)*
5. *Modifier les paramètres d'estimation si nécessaire* (View/Set – Analysis Properties...)
6. *Estimer le modèle en utilisant la commande* (Model Fit – Calculate Estimates)
7. *Étudier les coefficients obtenus sur le modèle et les tableaux de sortie dans* (View/Set – Table output)

LISREL: Les étapes

Étapes dans l'étude d'un modèle structurel avec LISREL:

1. *Obtenir un jeu de données avec si possible des données multinormales*
2. *Compléter les données manquantes et obtenir la matrice de covariance (*.cov) ou les données en format (*.psf) grâce au logiciel PRELIS intégré dans LISREL*
3. *Vérifier l'unidimensionnalité des blocs de variables*
4. *Paramétriser le modèle LISREL en utilisant le langage SIMPLIS*
5. *Donner les paramètres d'estimation si nécessaire*
6. *Estimer le modèle en utilisant la commande L*
7. *Étudier les coefficients obtenus sur le modèle et les tableaux de sortie (transformation possible en LaTeX)*

LISREL: Exemple de commande

Analyse de la structure de covariance du modèle à 3 variables manifestes

Observed Variables

V1 V2 V3 V4 V5 V6 V7 V8 V9 ←

Définition des variables manifestes

Raw data from file '...\exemple.psf' ←

Définition du fichier de données

Sample Size = 1000

Latent Variables L1 L2 L3 ←

Définition des variables latentes

Relationships

V1 V2 V3 = L1

V4 V5 V6 = L2 ←

Relations du modèle de mesure

V7 V8 V9 = L3

L2 = L1 ←

Relations du modèle interne

L3 = L1 L2

Path Diagram ←

Affichage du modèle graphique

End of Problem

LISREL: Exemple de sortie

LISREL Estimates (Maximum Likelihood) [\(back to TOC\)](#)

Measurement Equations

```

V4 = 0.78*L2, Errorvar.= 0.085 , R² = 0.88
                        (0.0058)
                        14.72
V5 = 0.78*L2, Errorvar.= 0.079 , R² = 0.88
      (0.014)          (0.0056)
      57.10           14.12
V6 = 0.79*L2, Errorvar.= 0.084 , R² = 0.88
      (0.014)          (0.0058)
      56.75           14.38
V7 = 0.87*L3, Errorvar.= 0.075 , R² = 0.91
                        (0.0054)
                        13.82
V8 = 0.88*L3, Errorvar.= 0.077 , R² = 0.91
      (0.013)          (0.0055)
      67.58           14.07
V9 = 0.88*L3, Errorvar.= 0.088 , R² = 0.90
      (0.013)          (0.0058)
      65.68           15.13

V1 = 0.71*L1, Errorvar.= 0.082 , R² = 0.86
      (0.019)          (0.0058)
      38.29           14.13
V2 = 0.70*L1, Errorvar.= 0.076 , R² = 0.87
      (0.018)          (0.0055)
      38.45           13.84
V3 = 0.71*L1, Errorvar.= 0.092 , R² = 0.84
      (0.019)          (0.0060)
      37.68           15.18
  
```

Structural Equations

```

L2 = 0.48*L1, Errorvar.= 0.77 , R² = 0.23
      (0.032)          (0.040)
      15.28           19.07
L3 = 0.42*L2 + 0.30*L1, Errorvar.= 0.62 , R² = 0.38
      (0.031)  (0.031)          (0.032)
      13.49           9.39           19.60
  
```

Goodness of Fit Statistics [\(back to TOC\)](#)

```

Degrees of Freedom = 24
Minimum Fit Function Chi-Square = 27.43 (P = 0.28)
Normal Theory Weighted Least Squares Chi-Square = 26.76 (P = 0.32)
Estimated Non-centrality Parameter (NCP) = 2.76
90 Percent Confidence Interval for NCP = (0.0 ; 19.66)
Population Discrepancy Function Value (FO) = 0.0028
90 Percent Confidence Interval for FO = (0.0 ; 0.020)
Root Mean Square Error of Approximation (RMSEA) = 0.011
90 Percent Confidence Interval for RMSEA = (0.0 ; 0.029)
P-Value for Test of Close Fit (RMSEA ≤ 0.05) = 1.00
90 Percent Confidence Interval for ECVI = (0.066 ; 0.086)
ECVI for Saturated Model = 0.090
ECVI for Independence Model = 12.59
Independence AIC = 12578.32
Model AIC = 68.76
Saturated AIC = 90.00
Independence CAIC = 12631.49
Model CAIC = 192.82
Saturated CAIC = 355.85
Non-Normed Fit Index (NNFI) = 1.00
Parsimony Normed Fit Index (PNFI) = 0.67
Comparative Fit Index (CFI) = 1.00
Incremental Fit Index (IFI) = 1.00
Relative Fit Index (RFI) = 1.00

Root Mean Square Residual (RMR) = 0.0048
Standardized RMR = 0.0071
Goodness of Fit Index (GFI) = 0.99
Adjusted Goodness of Fit Index (AGFI) = 0.99
Parsimony Goodness of Fit Index (PGFI) = 0.53
  
```

Les données manquantes en LISREL

Délétion par listes, Délétion par paires, Imputation de la moyenne, Imputation multiple par algorithme EM (Expectation Maximization)

Une méthode spécifique: Full Information Maximum Likelihood

Méthode prenant en compte uniquement les données disponibles sans mécanisme de complétion dans le calcul de la matrice de covariance.

Cette méthode maximise la vraisemblance casewise des données observées.

C'est la méthode qui obtient en général les meilleurs résultats en terme de biais des estimations.

Démonstration: L'approche PLS

Nous utiliserons le logiciel XLSTAT 2011

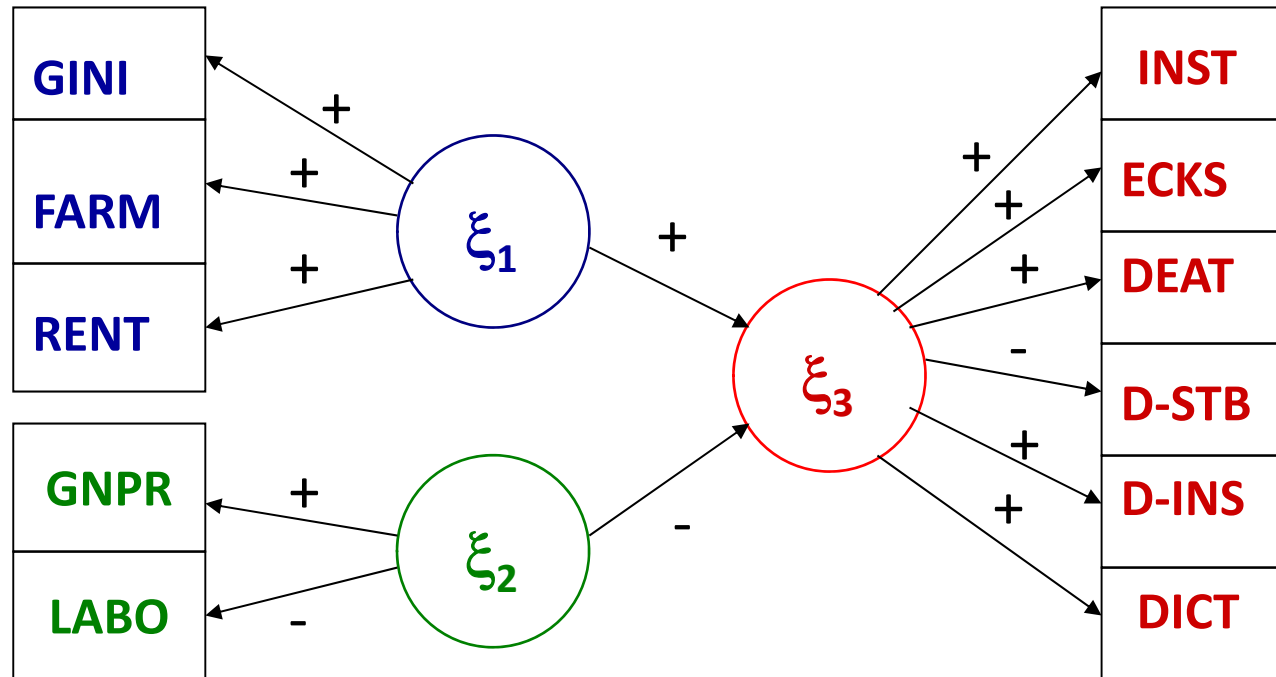
Nous utiliserons les données de Russett du dernier cours (Instabilité Politique)

XLSTAT est disponible à l'essai durant 30 jours sur www.xlstat.com .

Le module PLSPM est le plus abouti du marché en terme d'application de l'approche PLS.

Le modèle

Inégalité agricole (ξ_1)



Développement industriel (ξ_2)

Instabilité politique (ξ_3)

XLSTAT PLSPM: Les étapes

Étapes dans l'étude d'un modèle structurel avec XLSTAT-PLSPM :

- 1. Ouvrir les données dans Excel.***
- 2. Charger le logiciel.***
- 3. Créer un nouveau projet PLSPM.***
- 4. Copier les données dans la feuille D1.***
- 5. Dessiner le modèle PLS avec les variables latentes (ellipses) dans la feuille PLSPMGraph.***
- 6. Cliquer sur les variables latentes et sélectionner les variables manifestes à partir des données.***
- 7. Lancer les calculs.***
- 8. Étudier les sorties dans la feuille PLSPM générée.***

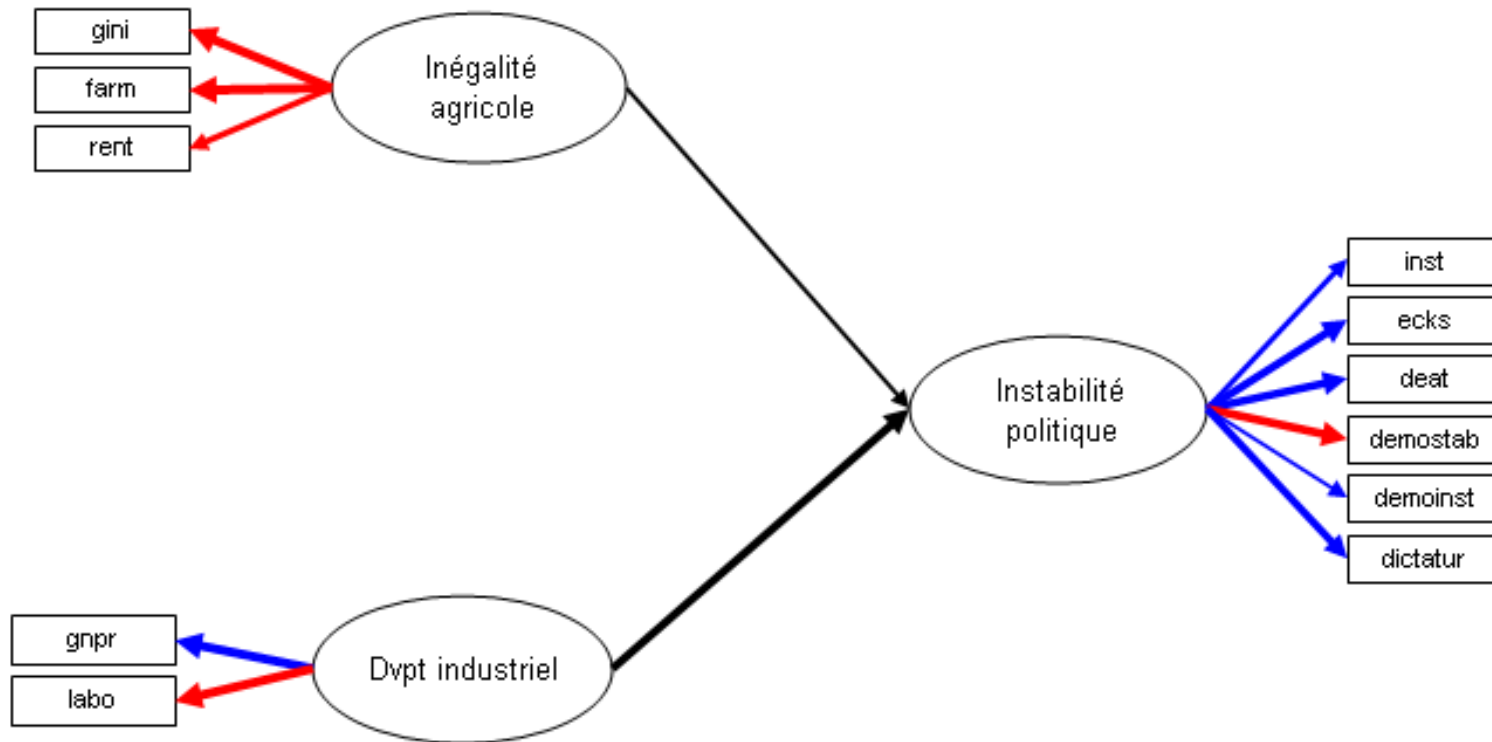
XLSTAT-PLSPM : La standardisation des VM

Plusieurs méthodes de standardisation des variables manifestes existent:

Options	Conditions			Options de standardisation			
	<i>Comparables</i>	<i>Moy. interprétables</i>	<i>Var. reflètent l'imp.</i>	<i>Stand.</i>	<i>Moy.</i>	<i>Var.</i>	<i>Remise à l'échelle</i>
1	Non	-	-	Oui	0	1	Non
2	Oui	Non	Non	Oui	0	1	Oui
3	Oui	Oui	Non	Oui	Initiale	1	Oui
4	Oui	Oui	Oui	Non	Initiale	Initiale	-

TAB. 3.2 – Options de standardisation présentes dans les logiciels appliquant l'approche PLS

XLSTAT-PLSPM : Les sorties

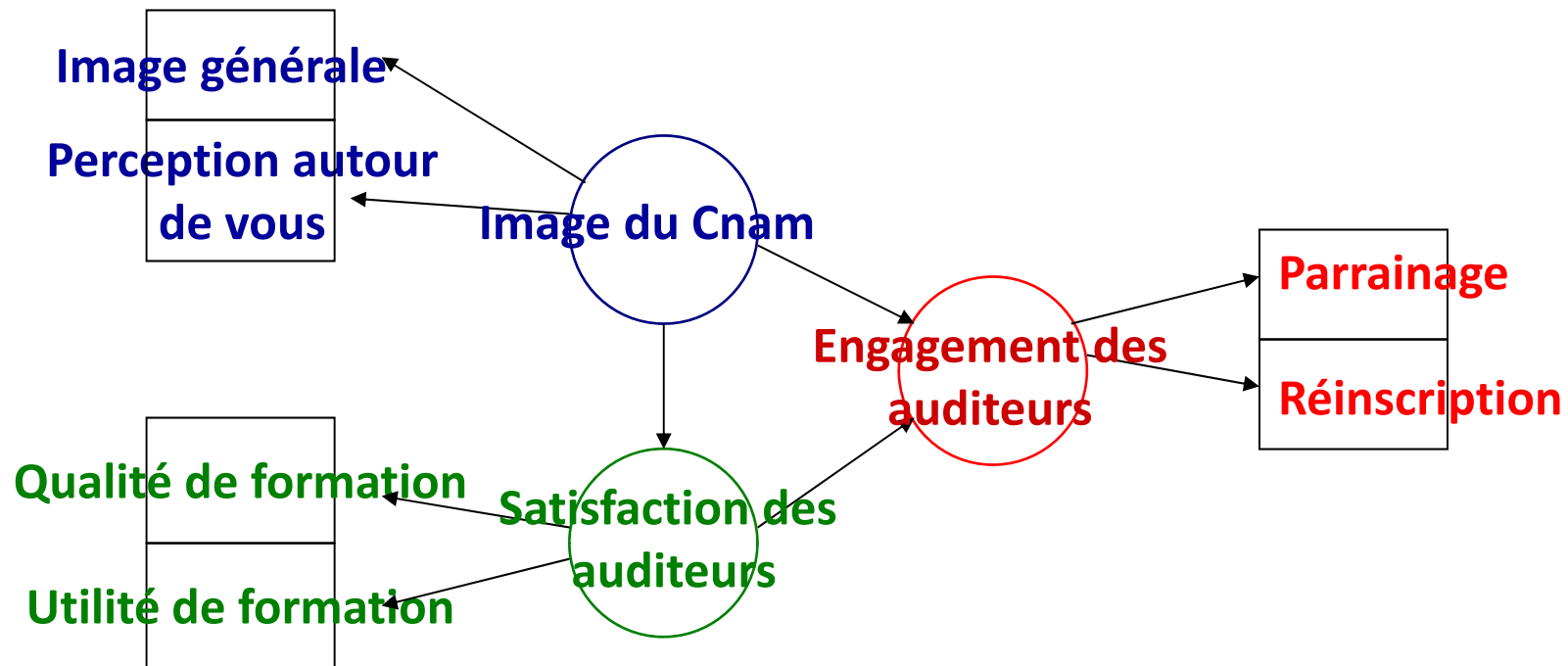


Étude du petit questionnaire distribué

Donnez une note de 1 à 10 selon que vous êtes tout à fait d'accord (10) ou pas du tout d'accord (1) avec les affirmations suivantes :

Image	<i>Le CNAM est une université... reconnue autour de vous qui a une bonne image</i>	1	2	3	4	5	6	7	8	9	10
		1	2	3	4	5	6	7	8	9	10
Satisfaction	<i>Les formations du CNAM sont... de bonne qualité utiles dans la vie active</i>	1	2	3	4	5	6	7	8	9	10
		1	2	3	4	5	6	7	8	9	10
Engagement	<i>Le CNAM offre des formations... que vous conseillerez à vos collègues que vous seriez prêt(e) à suivre (une fois celle-ci terminée)</i>	1	2	3	4	5	6	7	8	9	10
		1	2	3	4	5	6	7	8	9	10
	<i>Etes-vous satisfait(e) par les formations que vous suivez (avez suivi) au CNAM?</i>	Pas du tout	Plutôt pas	Plutôt	Très						
Variables de groupe	Est-ce la première formation que vous suivez au CNAM?		OUI	NON							

Étude du petit questionnaire distribué



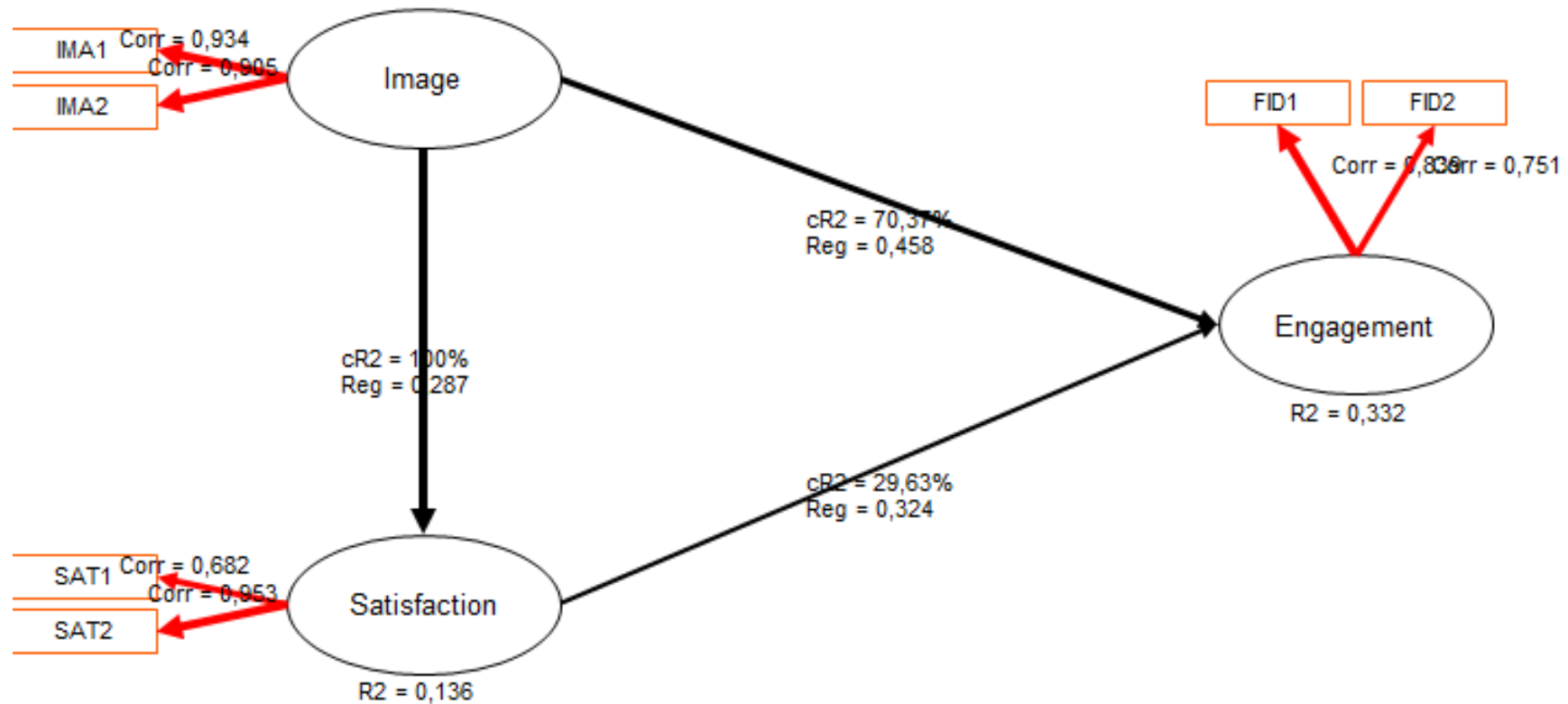
Étude du petit questionnaire distribué

*Aucune identification n'est possible avec la méthode LISREL
En raison de la faible taille de l'échantillon.*

*Avec l'approche PLS, l'algorithme converge et on peut ainsi
analyser les résultats.*

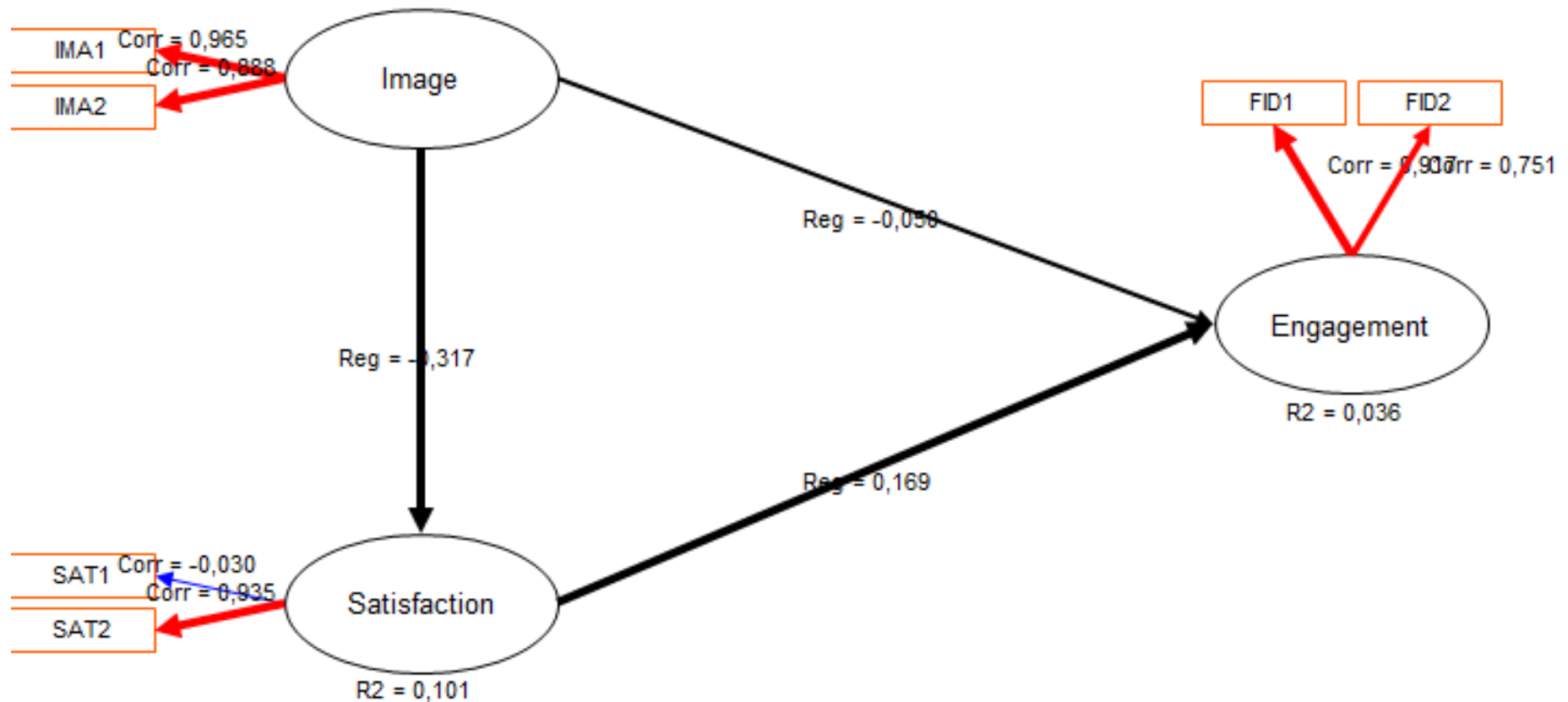
Étude du petit questionnaire distribué 2011-2012

N = 18



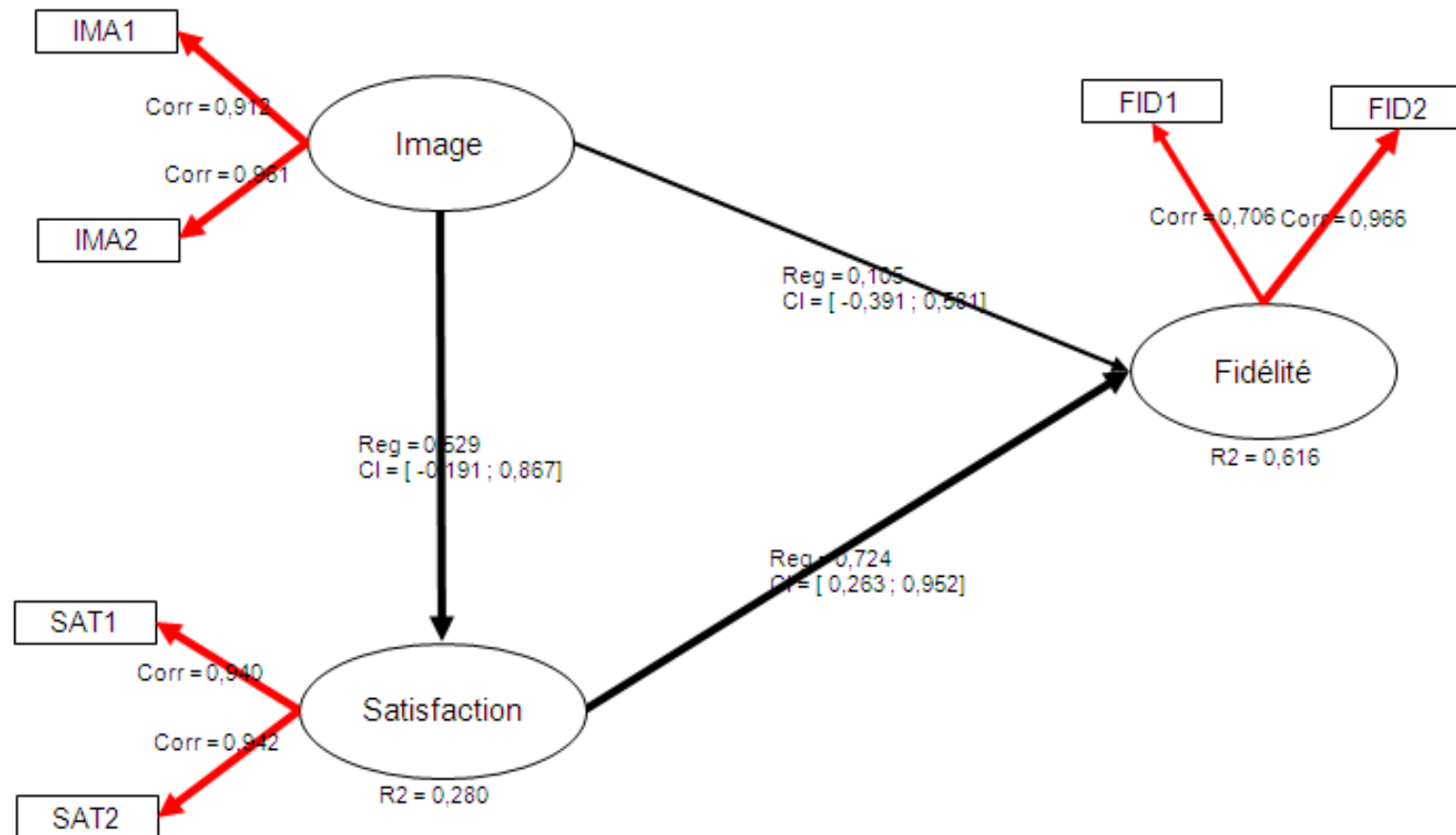
Étude du petit questionnaire distribué 2010-2011

N = 26



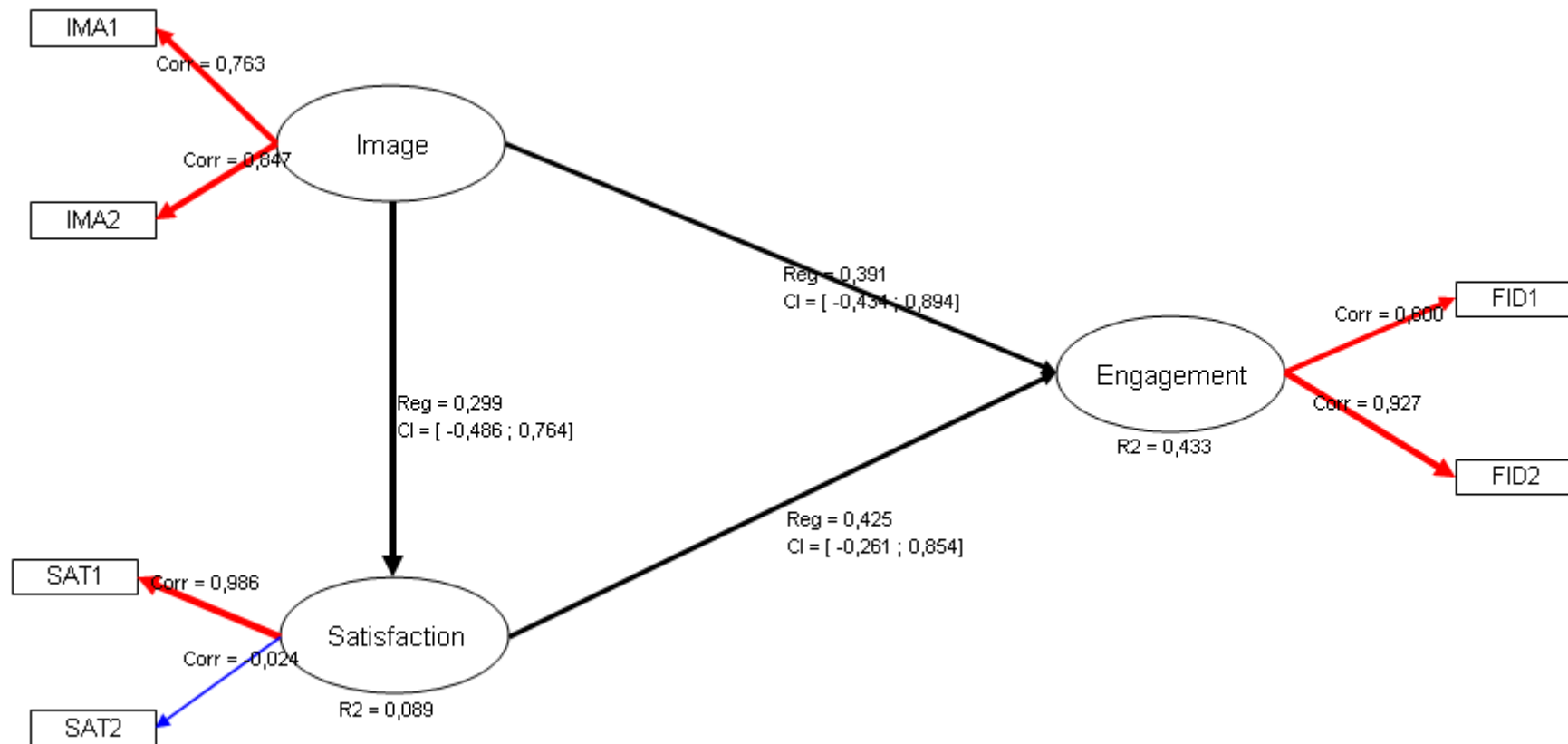
Étude du petit questionnaire distribué 2009-2010

N = 17



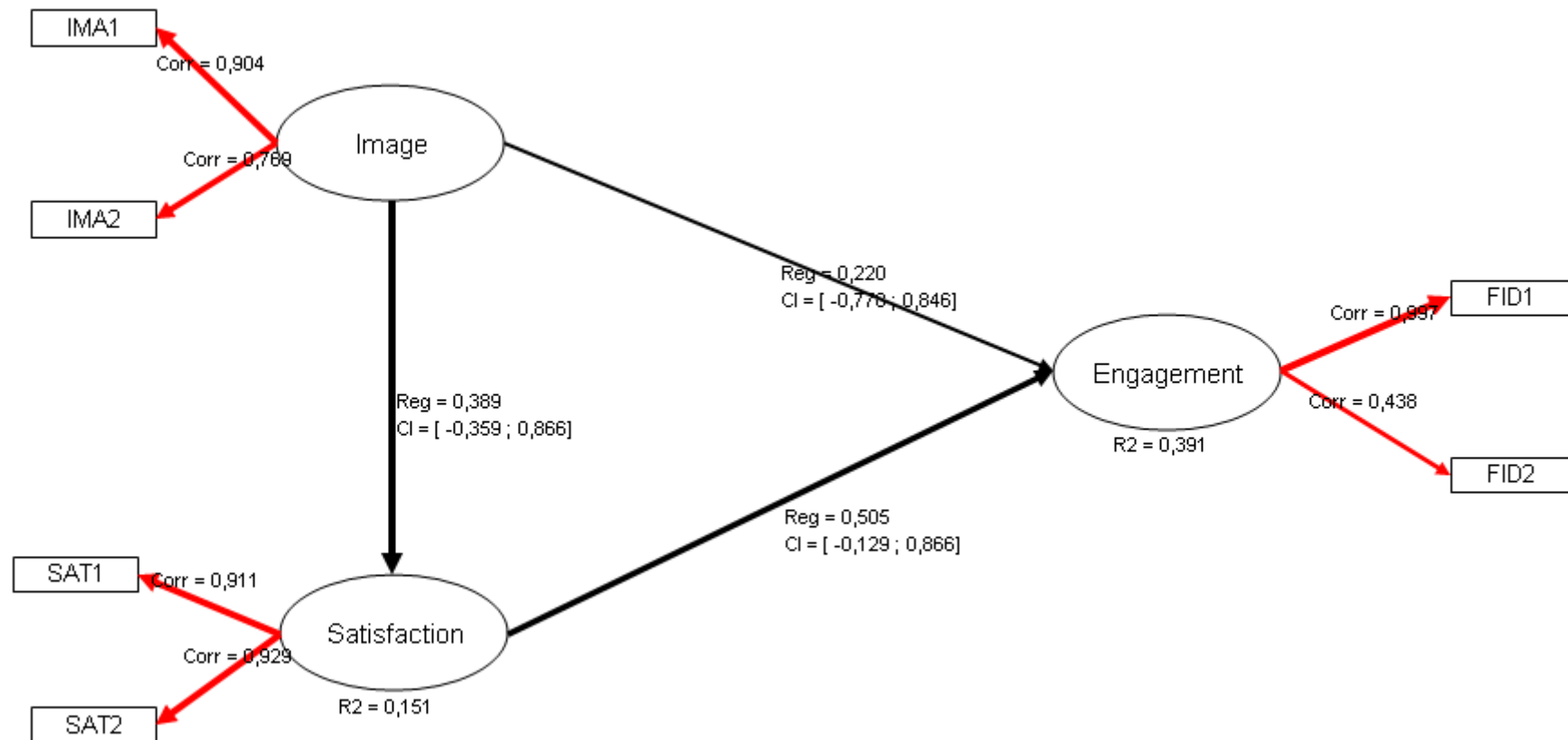
Étude du petit questionnaire distribué 2007-2008

N = 19



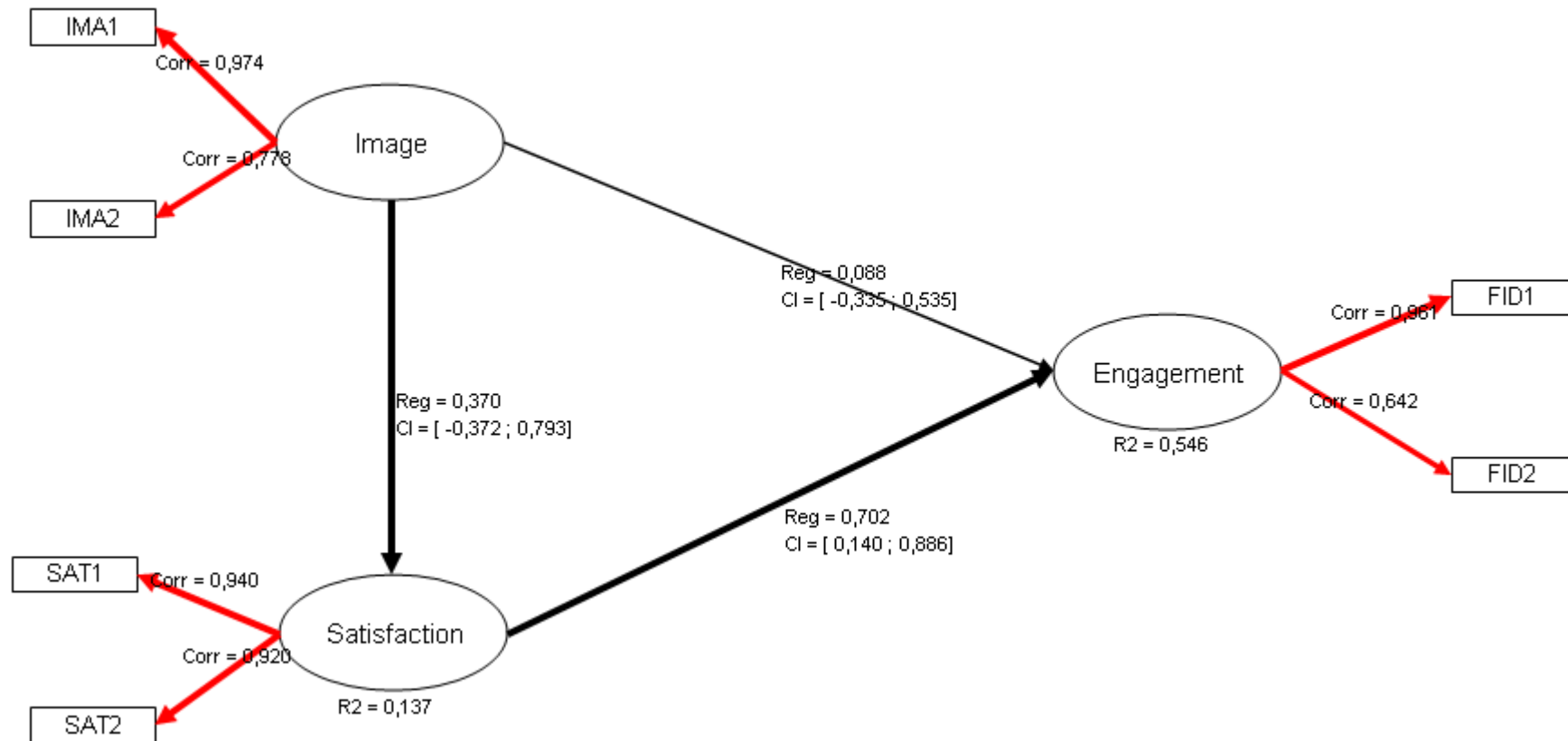
Étude du petit questionnaire distribué 2006-2007

N = 20



Étude du petit questionnaire distribué 2005-2006

N = 24



Étude du petit questionnaire distribué

L'utilisation de méthodes de bootstrap nous permettent d'obtenir des intervalles de confiance. Ainsi, en 2011-2012, mise à part la relation image – fidélité, les autres relations ne sont pas significatives.

Les R^2 pour 2011-2012 sont bons sachant qu'il n'y a que 18 observations. Il y a de faibles différences entre 2011 et les 4 années précédentes.

Ces résultats montre une forte influence de l'image sur l'engagement des auditeurs. La satisfaction a une influence non significative sur l'engagement.

Ces résultats nous permettent de conseiller au CNAM d'améliorer son image afin d'obtenir plus de nouveaux inscrits.

Comment améliorer ces résultats:

- augmenter la taille de l'échantillon,*
- ajouter des variables manifestes afin d'améliorer le modèle de mesure.*

Quelques problèmes associés aux modèles d'équations structurelles à variables latentes

Résultats des exercices distribués

Modélisation graphique - LISREL - 1

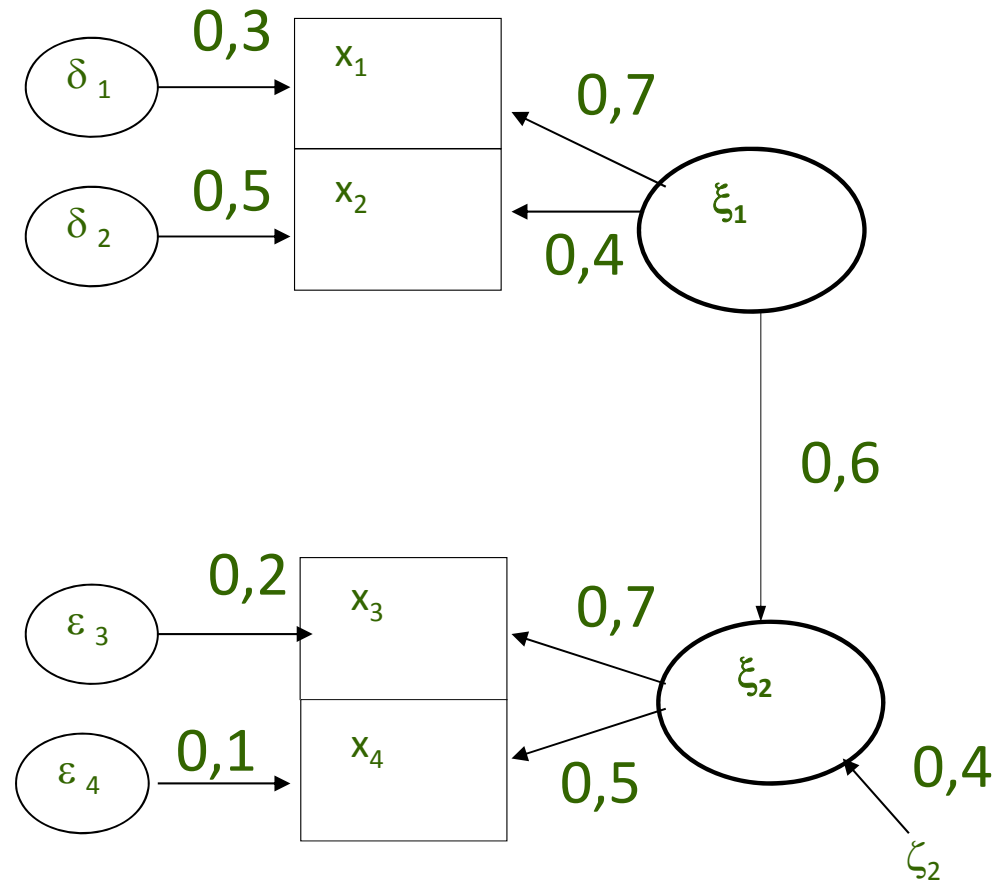
$$x_1 = 0,7\xi_1 + 0,3$$

$$x_2 = 0,4\xi_1 + 0,5$$

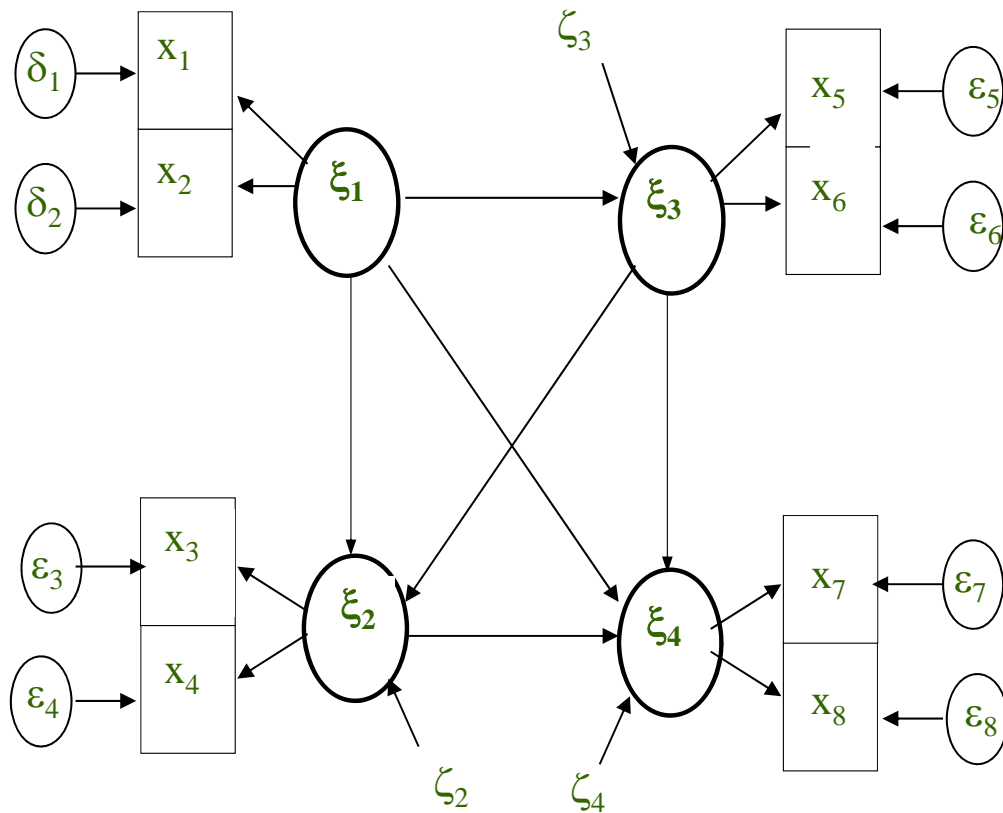
$$x_3 = 0,7\xi_2 + 0,2$$

$$x_4 = 0,5\xi_2 + 0,1$$

$$\xi_2 = 0,6\xi_1 + 0,4$$



Modélisation graphique - LISREL - 2



$$x_1 = \pi_1 \xi_1 + \delta_1$$

$$x_2 = \pi_2 \xi_1 + \delta_2$$

$$x_3 = \pi_3 \xi_2 + \varepsilon_3$$

$$x_4 = \pi_4 \xi_2 + \varepsilon_4$$

$$x_5 = \pi_5 \xi_3 + \varepsilon_5$$

...

$$\xi_2 = \beta_{12} \xi_1 + \beta_{32} \xi_3 + \zeta_2$$

$$\xi_3 = \beta_{13} \xi_1 + \zeta_3$$

$$\xi_4 = \beta_{14} \xi_1 + \beta_{24} \xi_2 + \beta_{34} \xi_3 + \zeta_4$$

Modélisation graphique - PLS - 1

$$x_1 = 0,7\xi_1 + 0,3$$

$$x_2 = 0,4\xi_1 + 0,5$$

$$x_3 = 0,7\xi_2 + 0,2$$

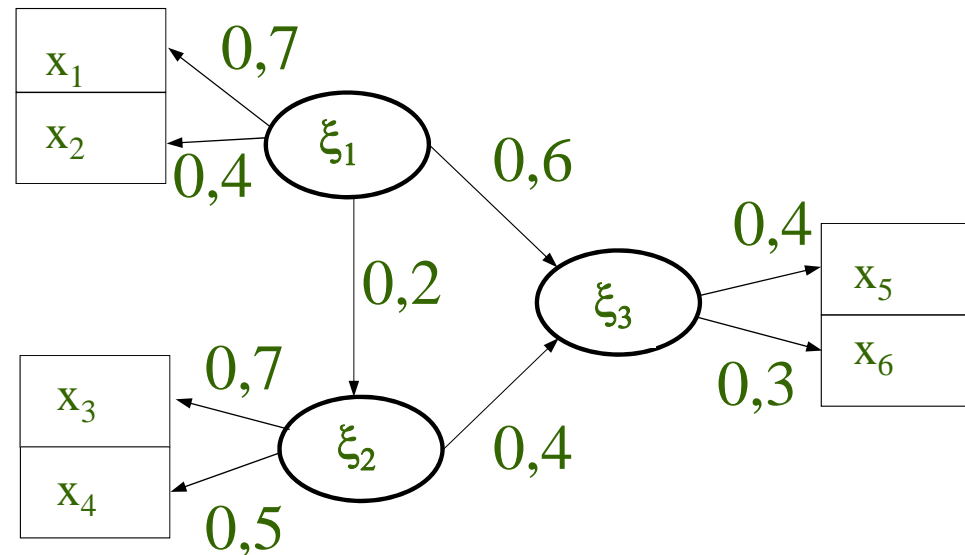
$$x_4 = 0,5\xi_2 + 0,1$$

$$x_5 = 0,4\xi_3 + 0,2$$

$$x_6 = 0,3\xi_3 + 0,3$$

$$\xi_2 = 0,2\xi_1 + 0,4$$

$$\xi_3 = 0,6\xi_1 + 0,4\xi_2 + 0,1$$



Modélisation graphique - PLS - 2

$$\xi_1 = 0,7x_1 + 0,3x_3 + 0,6x_4 + 0,4$$

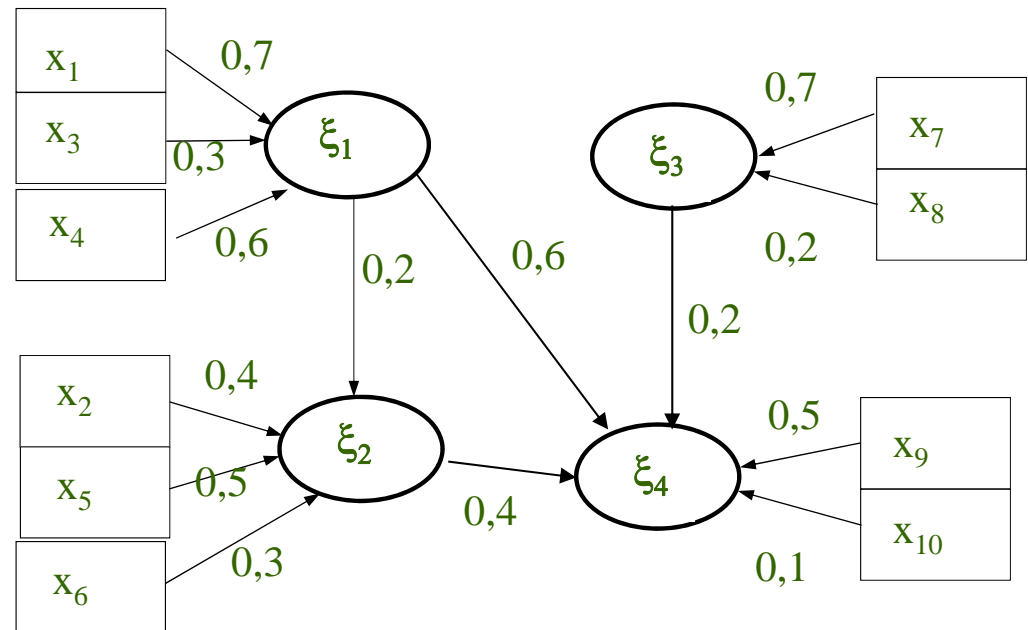
$$\xi_2 = 0,4x_2 + 0,5x_5 + 0,3x_6 + 0,1$$

$$\xi_3 = 0,7x_7 + 0,2x_8 + 0,2$$

$$\xi_4 = 0,5x_9 + 0,1x_{10} + 0,1$$

$$\xi_2 = 0,2\xi_1 + 0,1$$

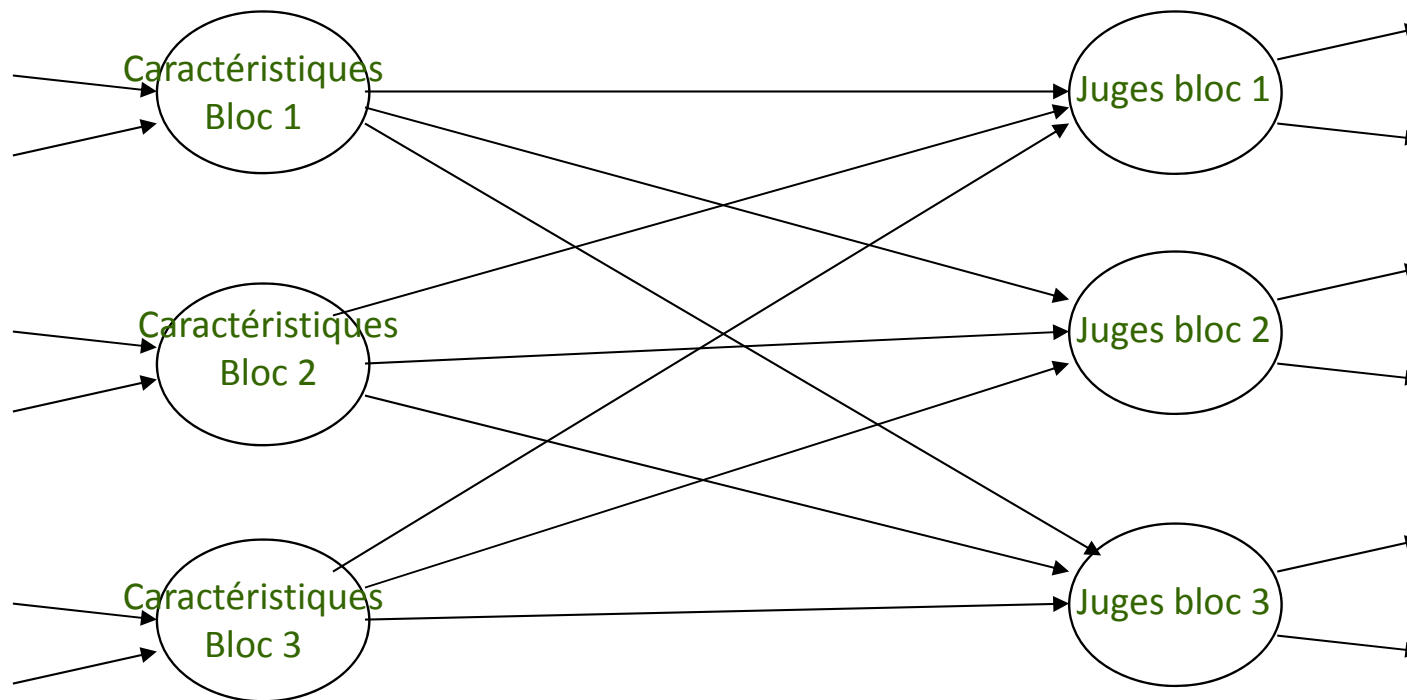
$$\xi_4 = 0,6\xi_1 + 0,4\xi_2 + 0,2\xi_3 + 0,1$$



Quelques problèmes associés aux modèles d'équations structurelles à variables latentes

Quelques problèmes associés aux modèles d'équations structurelles à variables latentes – Cas1 – Analyse sensorielle

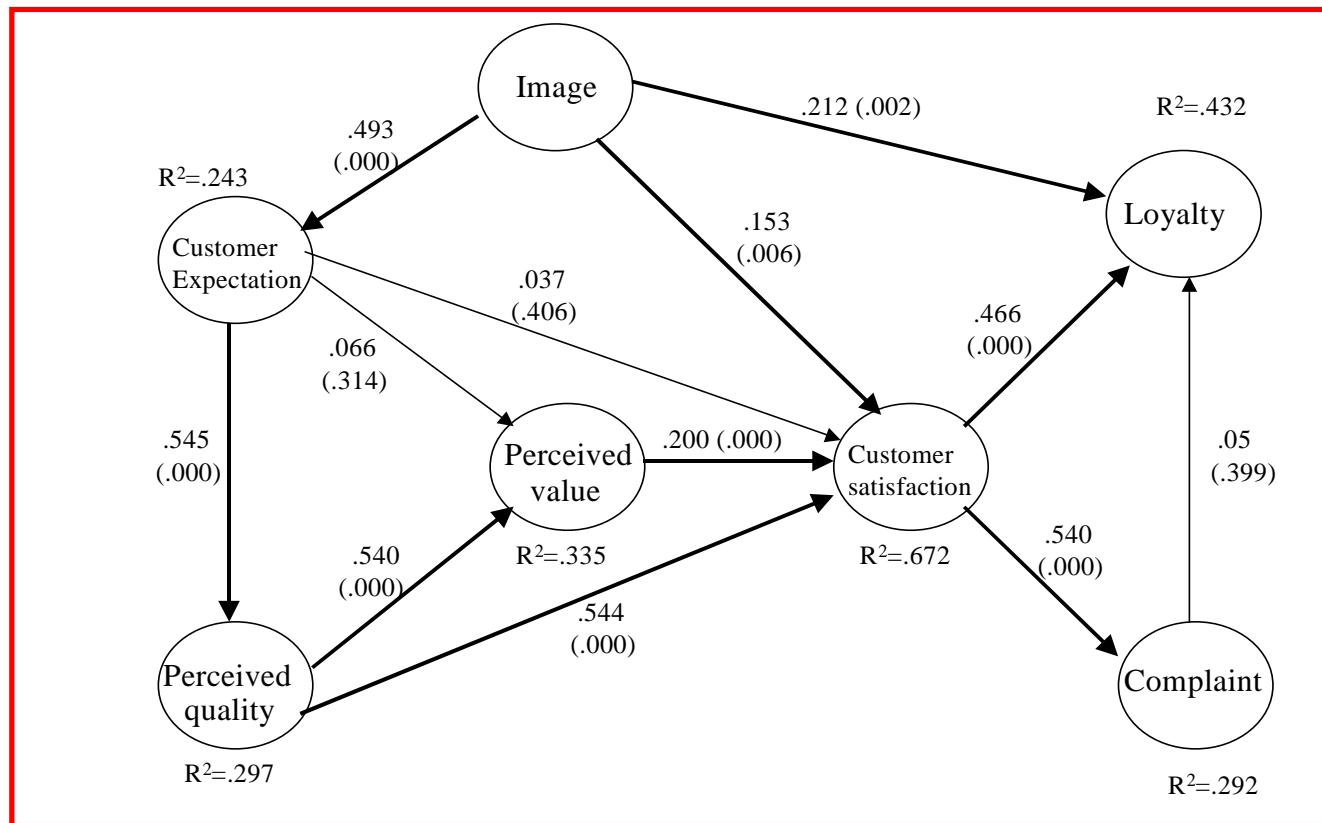
Une solution possible serait la suivante:



avec l'approche PLS et certaines variables formatives et d'autres réflexives

Quelques problèmes associés aux modèles d'équations structurelles à variables latentes –

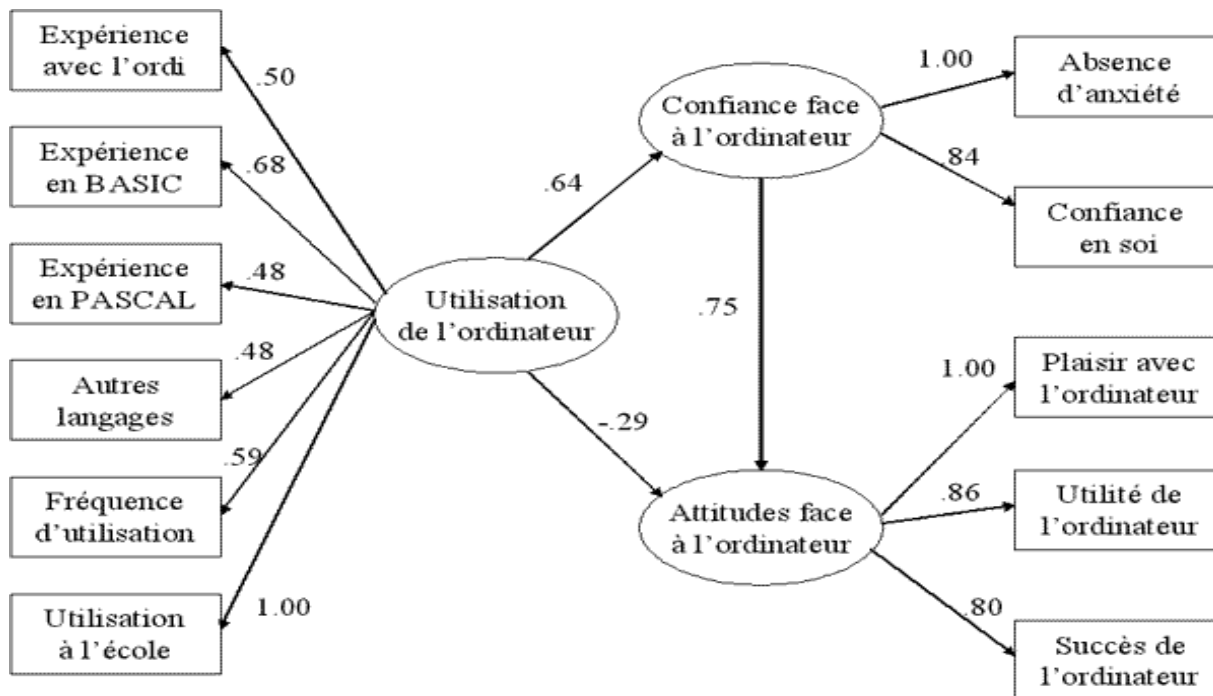
Cas 2 – Satisfaction



Les valeurs entre parenthèses sont les p-valeurs (le poids associé est significatif si la p-valeur < 0,05)

Le principal levier de la fidélité (loyalty) est la satisfaction des clients. Il faudra donc appuyer sur cette satisfaction et sur la qualité perçue qui est le principal levier de la satisfaction.

Quelques problèmes associés aux modèles d'équations structurelles à variables latentes – Cas 3



Ce modèle est bien défini et chaque coefficient est significatif, ainsi chaque relation peut être étudiée. On se rend compte que l'utilisation de l'ordinateur a un effet négatif sur l'attitude face à l'ordinateur alors qu'elle a un effet positif sur la confiance. D'autres remarques simples peuvent être faites en "lisant" le modèle.

Rappel: PLS vs LISREL

PLS est lié à LISREL de la même façon que l'ACP à l'analyse factorielle

- Orienté vers la prédiction des variables manifestes et latentes (basé sur la variance)
- Modèle externe réflectif et formatif
- Sans distribution
- Les observations peuvent être dépendantes
- Chaque VL est une combinaison linéaire de ces propres VM
- Consistance au sens large
- Prédiction optimale
- Évaluation de la qualité de prévision par jackknife (Q^2)
- Meilleur modèle de mesure car les VL sont estimées dans l'espace de leur propres VM

Cours de Statistique Multivariée Approfondie

- Orienté vers l'estimation des paramètres (basé sur la covariance)
- Modèle externe réflectif
- Hypothèses de distribution
- Les observations doivent être indépendantes
- Les facteurs ne sont pas estimés
- L'estimation des VL est faite sur l'ensemble des VM
- Estimations consistantes
- Paramètres optimaux
- Évaluation du modèle par tests d'hypothèse (N doit être grand)
- Meilleur modèle structurel car les VL sont estimées sans restriction d'espace