


Corso Modelli Causali

Prof. N. Carlo Lauro

I Modelli ad Equazioni Strutturali ed il PLS Path Modeling

(slide a cura di Laura Trinchera e Carlo Lauro)

Dipartimento di Matematica e Statistica
Università degli Studi di Napoli "Federico II"



Introduzione

I Modelli ad Equazioni Strutturali (Structural Equation Models, SEM)* sono modelli in grado di modellizzare complesse strutture di relazioni di **CAUSALITA'** tra concetti latenti (le *Variabili Latenti, VL*) a partire da un insieme di indicatori reali di solito definite come *Variabili Manifeste (VM)*.

Il termine SEM è da riferirsi a una famiglia molto vasta di metodi statistici in grado di analizzare le relazioni statistiche di causa-effetto e non ad un'unica tecnica.

* Per uno studio approfondito fare riferimento a Bollen 1989, Kaplan 2000 e Tenenhaus et al. 2005

Dr. Laura Trinchera, Università degli Studi di Napoli "Federico II"

Introduzione

I parametri di ogni Modello ad Equazioni Strutturali possono essere stimati seguendo due diversi approcci:

- l'approccio basato sulla covarianza tra le VM, i cosiddetti *covariance-based methods*
- l'approccio basato sulla ricerca di particolari componenti latenti, i cosiddetti *component-based methods*

Un po' di storia...chi ha inventato i SEM??

Diverse discipline hanno avuto un ruolo nello sviluppo dei Modelli ad Equazioni Strutturali:

- L'"Econometria" di Sewall Wright e la Path Analysis
 - La Psicometria e l'Analisi Fattoriale
 - Economisti e Sociologi alla riscoperta dei Modelli Causali
- I Modelli ad Equazioni Strutturali e Karl Jöreskog

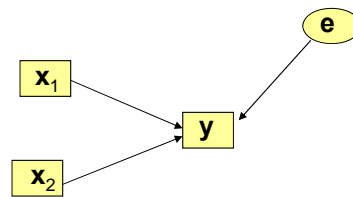
Il Path Diagram

Il Path Diagram è una rappresentazione grafica delle relazioni esistenti tra le variabili sotto esame

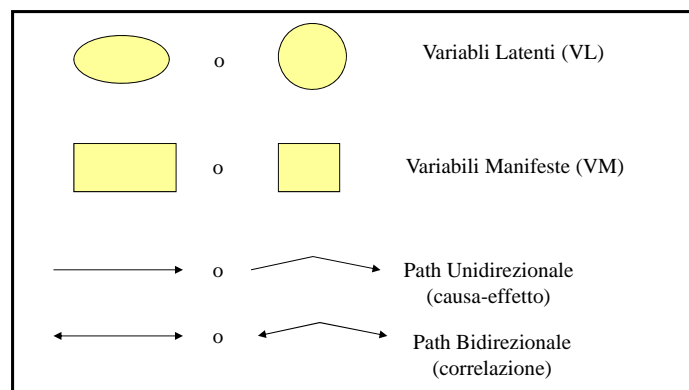
Un semplice modello di regressione multipla del tipo:

$$y = ax_1 + bx_2 + e$$

può essere rappresentato graficamente attraverso un semplice Path Diagram:



Il Path Diagram: i simboli



Modelli ad Equazioni Strutturali: le nozioni di base

The diagram illustrates a Structural Equation Model (SEM) with three latent variables (ξ_1, ξ_2, ξ_3) and their corresponding manifest variables (x_{ij}). The model is divided into two parts: the **Modello di Misura o modello esterno** (Measurement Model) and the **Modello Strutturale o modello interno** (Structural Model). The Measurement Model shows the relationship between latent variables and their indicators, with path coefficients β_1 and β_2 connecting ξ_1 and ξ_2 to ξ_3 . The Structural Model shows the relationship between manifest variables, with external weights w_{12} and w_{22} connecting x_{12} and x_{22} to ξ_2 . The manifest variables are grouped into blocks: x_{11}, x_{21}, x_{31} for ξ_1 ; x_{12}, x_{22}, x_{32} for ξ_2 ; and $x_{13}, x_{23}, x_{33}, x_{43}, x_{53}$ for ξ_3 .

- P Variabili Manifeste (VM) osservate su N unità $\rightarrow x_{pq}$ generica MV
- Q Variabili Latenti (VL) $\rightarrow \xi_q$ generica LV
- Q blocchi di variabili, ognuno composto da una VL e dalle corrispondenti VM
- Ogni q-esimo blocco è composto da p_q VM x_{pq} , con $\sum_{q=1}^Q p_q = P$

Modelli ad Equazioni Strutturali: il modello di misura

Indicatori Riflessivi

es. Intelligenza

Indicatori Formativi

es. Status Sociale

Modo A

- Le VL generano le corrispondenti VM (la VL sottostante ogni blocco è unica \rightarrow blocco unidimensionale)
- Le VL sono antecedenti rispetto alle VM
- Gli indicatori riflessivi di uno stesso blocco devono essere fortemente legati tra loro, in altre parole devono covariare
- La consistenza Interna al blocco deve essere verificata (ad esempio usando l'alpha di Cronbach)

Modo B

- La VL è una combinazione lineare delle corrispondenti VM (ogni VL è quindi un costrutto multidimensionale)
- Non è detto che gli indicatori formativi di uno stesso blocco debbano covariare tra loro.
- La consistenza Interna non necessita di essere verificata.

Modelli ad Equazioni Strutturali: il modello di misura

Indicatori Riflessivi

Indicatori Formativi

Modo A

$$\mathbf{x}_{pq} = \lambda_{pq}\xi_q + \varepsilon_{pq}$$

- **Regressione OLS semplice** → nessun problema legato alla multicollinearità tra gli indicatori di ogni blocco.
- La VL può essere considerata come la prima componente principale del proprio blocco di VM sotto il vincolo di spiegare al meglio le relazioni tra VL adiacenti.

Modo B

$$\xi_q = \sum_{p=1}^{P_q} w_{pq} \mathbf{x}_{pq} + \delta_q$$

- **Regressione OLS multipla** → la multicollinearità crea problemi!!
- **Ogni VL è il migliore predittore delle sue VM** sotto il vincolo di spiegare al meglio le relazioni tra le VL adiacenti.

Modelli ad Equazioni Strutturali: Indicatori Riflessivi

Verificare l'unidimensionalità del blocco:

Un modello di misura di tipo **riflessivo** assume che ogni blocco di variabili manifeste sia unidimensionale, in altre parole che un solo concetto latente si rifletta in più indicatori.

La **consistenza interna** di un blocco di variabili manifeste può essere verificata soprattutto attraverso tre indici:

- Sulla base dei risultati di un' **Analisi in Componenti Principali**:
 - Un solo autovalore > 1
 - **Loading Plot (Correlazioni Fattori/Variabili)** per identificare eventuali sottogruppi unidimensionali
- Utilizzando l' **alpha di Cronbach**:
$$\alpha_q = \frac{\sum_{p=p'}^{P_q} cor(\mathbf{x}_{pq}, \mathbf{x}_{pq'})}{P_q + \sum_{p=p'}^{P_q} cor(\mathbf{x}_{pq}, \mathbf{x}_{pq'})} \times \frac{P_q}{P_q - 1}$$
- Utilizzando il **rho di Dillon-Goldstein**:
$$\rho_q = \frac{\left(\sum_{p=1}^{P_q} \lambda_{pq} \right)^2}{\left(\sum_{p=1}^{P_q} \lambda_{pq} \right)^2 + \sum_{p=1}^{P_q} (1 - \lambda_{pq}^2)}$$

Un blocco è considerato unidimensionale se questi ultimi 2 indici sono almeno uguali a 0,7

Modelli ad Equazioni Strutturali:

Indicatori iflessivi

E se il blocco non è unidimensionale?



Le soluzioni possibili:

- **Rimuovere** le variabili manifeste che rendono il blocco non unidimensionale
- Dividere il blocco multidimensionale in **sotto-blocchi unidimensionali**
- Cambiare il tipo di modello di misura e passare al formativo...ma **ATTENZIONE** all'INTERPRETAZIONE della variabile latente!!!

Modelli ad Equazioni Strutturali:

Indicatori Formativi



I pesi esterni (w_{pq}) sono i coefficienti di regressione OLS ottenuti da una regressione multipla di ogni Variabile Latente sul corrispondente blocco di Variabili Manifeste:

$$\mathbf{w}_q = (\mathbf{X}'_q \mathbf{X}_q)^{-1} \mathbf{X}'_q \boldsymbol{\xi}_q$$

e se le variabili manifeste sono molto correlate tra loro (come risolvere il problema della multicollinearità tra le Variabili Manifeste) ???

- utilizzare i **loadings** al posto dei weights
- utilizzare i coefficienti di una **regressione PLS**



Le Variabili Latenti sono ottenute come **combinazione lineare** delle Variabili Manifeste:

$$\boldsymbol{\xi}_q = \sum_{p=1}^{P_q} w_{pq} \mathbf{x}_{pq} + \delta_q$$

quindi i pesi devono essere tutti dello stesso segno!!

Modelli ad Equazioni Strutturali: Modello Strutturale

Il modello strutturale descrive le relazioni di causalità tra le variabili latenti.

Per ogni variabile latente endogena nel modello il modello strutturale può essere scritto come:

$$\xi_j = \sum_{m=1}^M \beta_{mj} \xi_m + \zeta_j$$

dove:

- β_{mj} è il path-coefficient che lega l'*m-essima* VL alla *j-esima* VL endogena
- M è il numero di variabili latenti esplicative che impattano sulla generica variabile latente endogena ξ_j

Modelli ad Equazioni Strutturali: i parametri del modello

Per concludere, in ogni Modello ad Equazioni Strutturali i seguenti parametri devono essere stimati:

* i pesi esterni (outer weights) $\rightarrow w_{pq}$

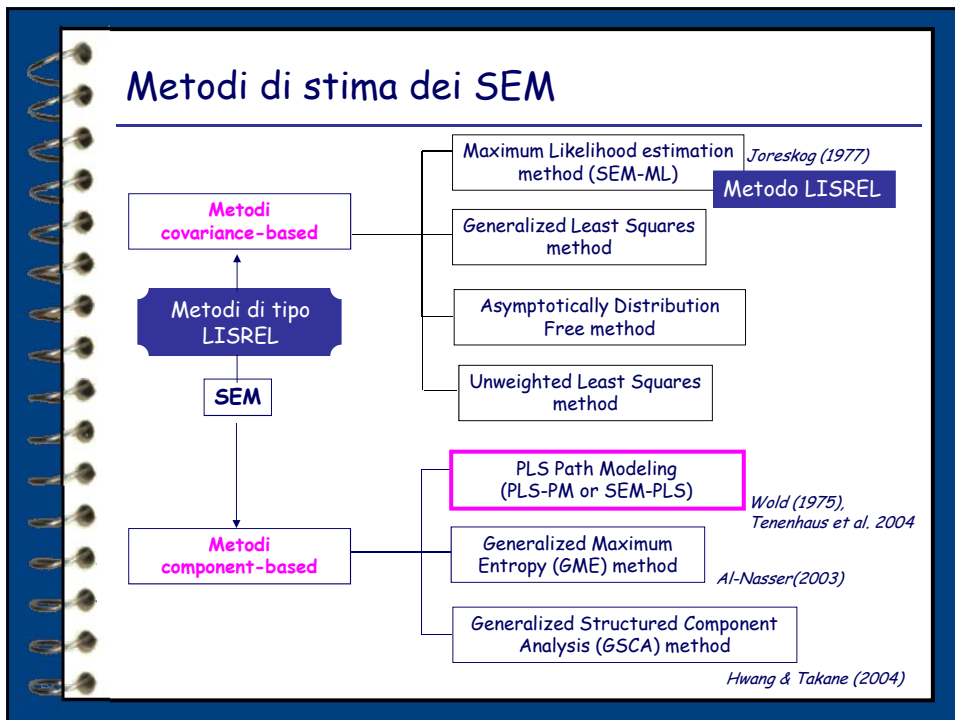
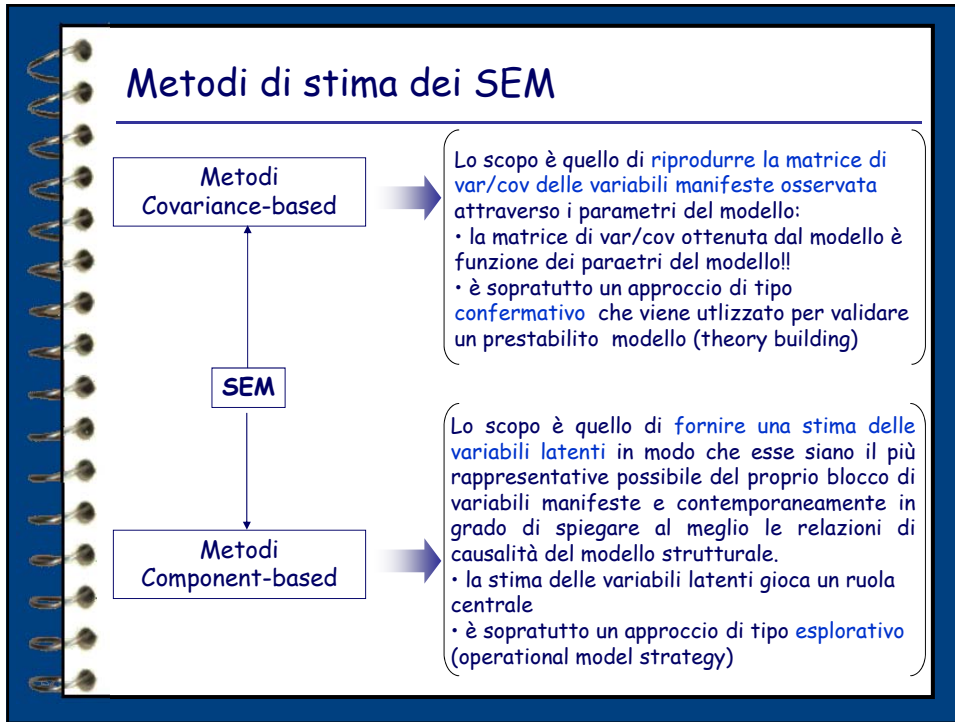
* i loadings $\rightarrow \lambda_{pq}$

* i path-coefficients $\rightarrow \beta_{mj}$

Così come i residui:

* del modello di misura $\rightarrow \delta_q$ (nel caso di indicatori **formativi**)
 ε_{pq} (nel caso di indicatori **riflessivi**)

* del modello strutturale $\rightarrow \zeta_j$



Il PLS Path Modeling o SEM-PLS

È un algoritmo iterativo che permette di stimare le relazioni esistenti tra i diversi blocchi di variabili attraverso un sistema di **equazioni lineari**.



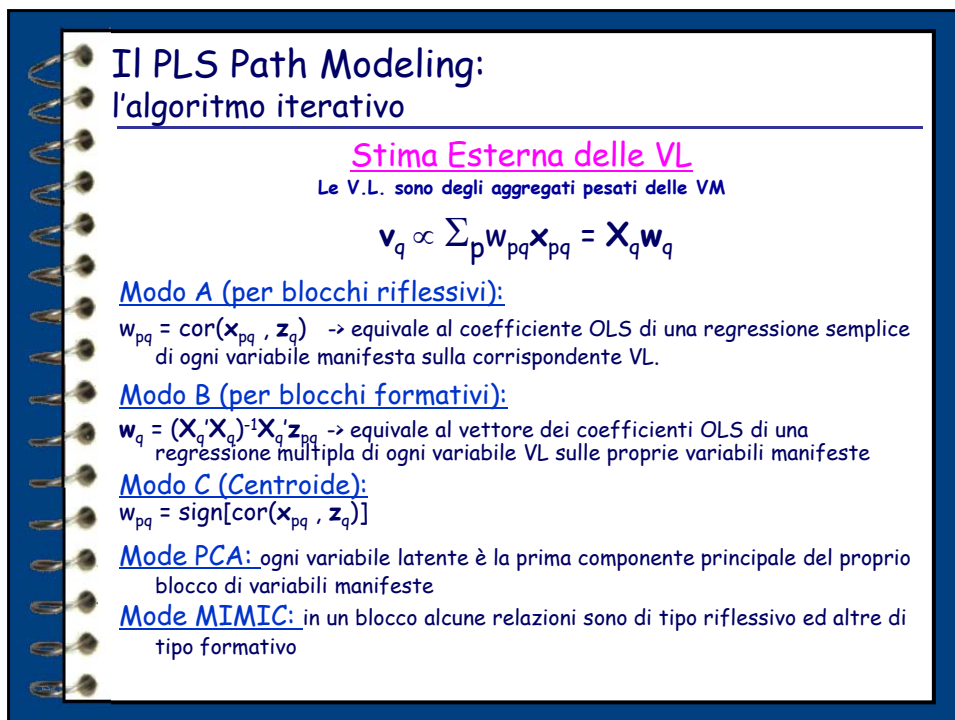
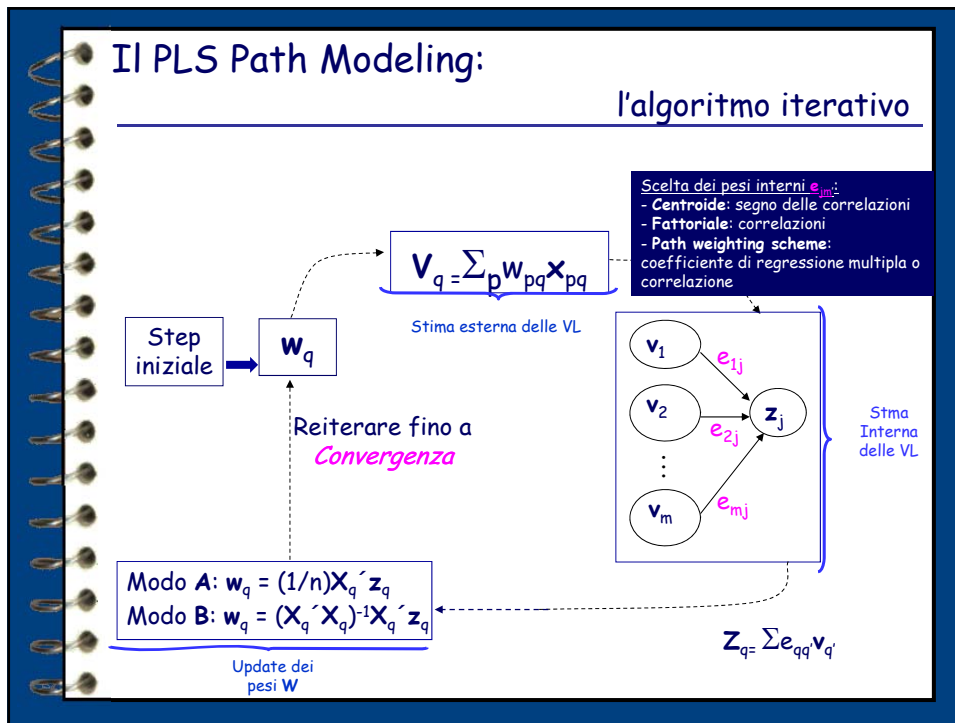
Lo scopo è quello di **identificare delle variabili latenti** che siano allo stesso tempo il **più rappresentative** possibile delle **loro variabili manifeste** e che siano in grado di **spiegare al meglio le relazioni** tra blocchi così come definite dalla struttura del modello interno.

Il PLS Path Modeling o SEM-PLS

Pro e contro del PLS-PM:

.

- Il PLS-PM non necessita **nessuna ipotesi** sulla forma della **distribuzione** delle variabili manifeste (né su quella delle variabili latenti!!)
- L'algoritmo iterativo direttamente **fornisce una stima delle variabili latenti** (latent variable scores)
- **La convergenza** del PLS-PM è assicurata **empiricamente** ma non esiste una prova matematica (tranne nel caso di due soli blocchi di variabili manifeste)
- Nel PLS-PM **non è possibile identificare un'unica funzione da ottimizzare**; tuttavia, casi specifici del PLS-PM ottimizzano specifici criteri (vedi l'ACP o le CC).



Il PLS Path Modeling:

l'algoritmo iterativo

Stima Interna delle VL

Le V.L. sono degli aggregati pesati delle VL adiacenti

$$\mathbf{z}_q \propto \sum e_{qq'} \mathbf{v}_{q'}$$

Schema del Centroide (proposto da Wold!):

$$e_{qq'} = \text{sign}[\text{cor}(\mathbf{v}_q, \mathbf{v}_{q'})]$$

-> problemi con correlazioni ≈ 0 .

Nessuna differenza tra VL
esogene ed endogene!

Schema Fattoriale (PLS, Lohmoller 1989):

$$e_{qq'} = r_{qq'} = \text{cor}(\mathbf{v}_q, \mathbf{v}_{q'})$$

Nessuna differenza tra VL
esogene ed endogene!

Schema Strutturale (Path Weighting, Lohmoller 1989):

e_{mj} = coefficiente di regressione semplice/multipla di \mathbf{v}_j on \mathbf{v}_m se ξ_j è una
variabile endogena

$e_{mj} = r_{jm}$ se ξ_m è una variabile che impatta su ξ_j **Il peso e_{mj} è diverso dal peso
 e_{jm} a seconda dal ruolo assunto
dalla VL!**

Il PLS Path Modeling:

l'algoritmo iterativo

Scelto un sistema di pesi esterni arbitrario \mathbf{w}_q (es. 1,0,...,0)

I. Calcolare la stima esterna di ogni VL:

$$\mathbf{v}_q \propto \sum_p w_{pq} \mathbf{x}_{pq} = \mathbf{X}_q \mathbf{w}_q$$

II. Stabilire lo schema per il calcolo dei pesi interni e calcolare i
pesi interni $e_{qq'}$

III. Calcolare la stima interna di ogni VL seguendo le relazioni
definite nel modello strutturale:

$$\mathbf{z}_q \propto \sum e_{qq'} \mathbf{v}_{q'}$$

IV. Aggiornare il sistema di pesi esterni w_{pq} secondo lo schema A
o B (o MIMIC)

Iterare la procedura fino al raggiungimento della **convergenza** sui pesi esterni

Una volta ottenuta la convergenza:

→ Ottenere ogni variabile latente come combinazione lineare delle sue variabili
manifeste con pesi uguali ad i pesi ottenuti dall'algoritmo iterativo: $\xi_q = \mathbf{X}_q \mathbf{w}_q$

→ Calcolare i Path-coefficients come coefficienti OLS di una regressione
semplice/multipla tra VL

Esempio Disuguaglianza Economica ed instabilità politica (dati Russett, 1964)

Disuguaglianza Economica

Disuguaglianza nell'agricoltura

GINI : Inequality of land distributions
FARM : % farmers that own half of the land (> 50%)
RENT : % farmers that rent all their land

Sviluppo Industriale

GNPR : Gross national product per capita (\$ 1955)
LABO : % of labour force employed in agriculture

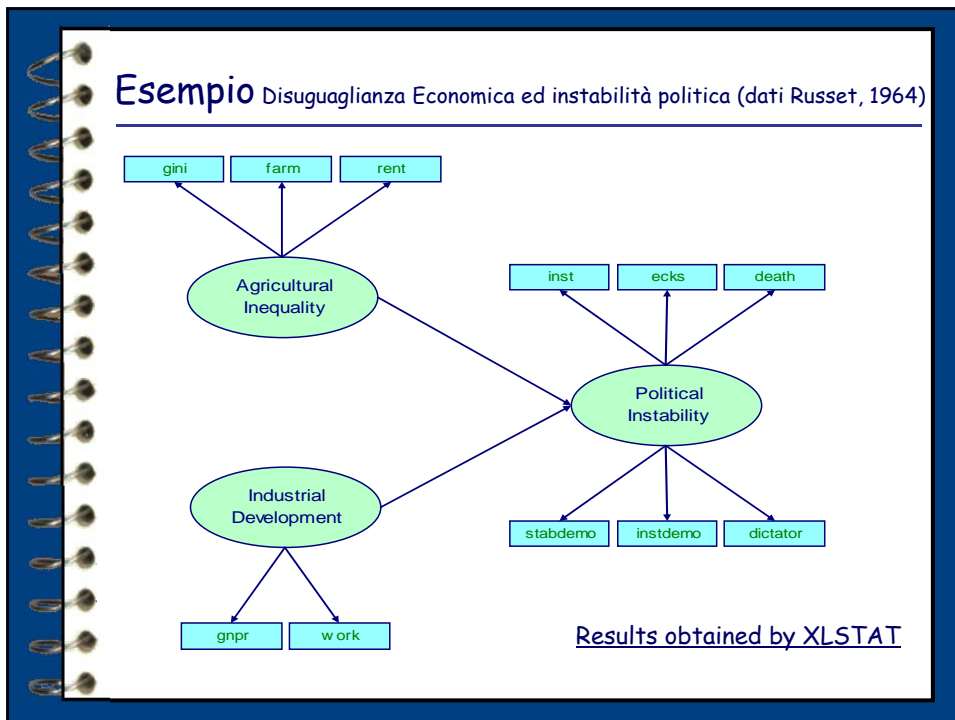
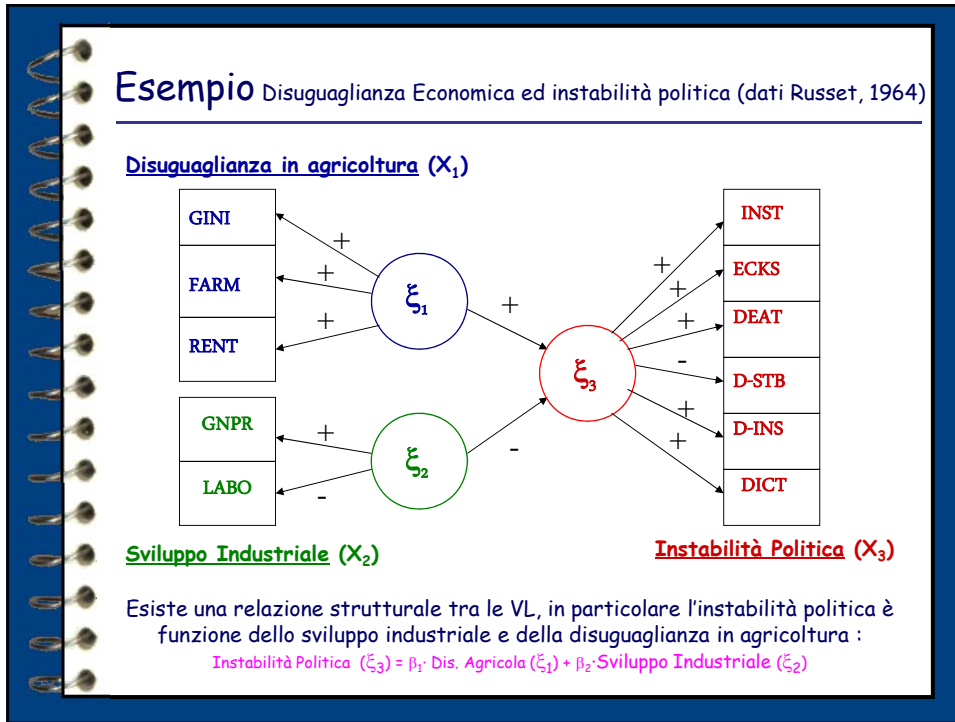
Instabilità politica

INST : Instability of executive (1945-1961)
ECKS : Nb of violent internal war incidents ('46-'61)
DEAT : Nb of people killed as a result of civil war violence ('50-'62)
DEMO :
D-STAB : Stable democracy
D-UNST : Unstable democracy
DICT : Dictatorship

Esempio Disuguaglianza Economica ed instabilità politica (dati Russett, 1964)

	Gini	Farm	Rent	Gnpr	Labo	Inst	Ecks	Deat	Demo
Argentina	86.3	98.2	32.9	374	25	13.6	57	217	2
Australia	92.9	99.6	29.5	1215	14	11.3	0	0	1
Austria	74.0	97.4	10.7	532	32	12.8	4	0	2
⋮									
France	58.3	86.1	26.0	1046	26	16.3	46	1	2
⋮									
Yugoslavia	43.7	79.8	0.0	297	67	0.0	9	0	3

La variabile DEMO è una variabile categorica con 3 modalità (e sarà dicotomizzata):
 1 = democrazia stabile
 2 = democrazia instabile
 3 = Dittatura



Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

- (1) Stima **Esterna** della VL ξ_q : \mathbf{v}_q

$$\mathbf{v}_q = \mathbf{X}_q \mathbf{w}_q$$

- (2) Stima **Interna** della VL ξ_q : \mathbf{z}_q (**Schema del Centroide**)

$$\mathbf{z}_q = \Sigma [\text{sign}(\text{cor}(\xi_q, \xi_{q'}))] \mathbf{v}_{q'}$$

Per ogni VL $\xi_{q'}$ legata alla VL ξ_q

- (3) Calcolo dei nuovi pesi esterni \mathbf{w}_q (**Modo A, schema riflessivo**)

$$\mathbf{w}_{pq} = \text{cor}(\mathbf{z}_q, \mathbf{x}_{pq})$$

Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

(1) Stima Esterna

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{X}_1 \mathbf{w}_1 \\ \mathbf{v}_2 &= \mathbf{X}_2 \mathbf{w}_2 \\ \mathbf{v}_3 &= \mathbf{X}_3 \mathbf{w}_3 \end{aligned}$$

(2) Stima Interna

$$\begin{aligned} \mathbf{z}_1 &= \text{sign}(\text{cor}(\mathbf{v}_1, \mathbf{v}_3)) \mathbf{v}_3 \\ &= (+1) \mathbf{v}_3 \end{aligned}$$

$$\begin{aligned} \mathbf{z}_2 &= \text{sign}(\text{cor}(\mathbf{v}_2, \mathbf{v}_3)) \mathbf{v}_3 \\ &= (-1) \mathbf{v}_3 \end{aligned}$$

$$\begin{aligned} \mathbf{z}_3 &= \text{sign}(\text{cor}(\mathbf{v}_3, \mathbf{v}_1)) \mathbf{v}_1 + \text{sign}(\text{cor}(\mathbf{v}_3, \mathbf{v}_2)) \mathbf{v}_2 \\ &= (+1) \mathbf{v}_1 + (-1) \mathbf{v}_2 \end{aligned}$$

(3) Calcolo dei pesi esterni \mathbf{w}_h

$$\begin{aligned} \mathbf{w}_{1j} &= \text{cor}(\mathbf{x}_{1j}, \mathbf{z}_1) \\ \mathbf{w}_{2j} &= \text{cor}(\mathbf{x}_{2j}, \mathbf{z}_2) \\ \mathbf{w}_{3j} &= \text{cor}(\mathbf{x}_{3j}, \mathbf{z}_3) \end{aligned}$$

Algoritmo

- **Iterare** la procedura fino al raggiungimento della **convergenza** (dimostrazione matematica della convergenza ottenuta solo per 2 blocchi di variabili, ma convergenza empirica sempre verificata).

Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

Outer model (Dimension 1):

Risultati - problemi nei segni
delle correlazioni VL/VM...

Weights (Dimension 1):

Latent variable	Manifest variables	Outer weight	Outer weight (Bootstrap)	Standard error	Critical ratio (CR)	Lower bound (95%)	Upper bound (95%)
Agricultural Inequality	gini	0.460	0.453	0.046	9.889	0.325	0.545
	farm	0.516	0.510	0.049	10.451	0.395	0.653
	rent	0.081	0.080	0.152	0.535	-0.275	0.347
Industrial Development	gnpr	-0.511	-0.511	0.023	-22.112	-0.557	-0.448
	labo	0.538	0.536	0.022	24.111	0.498	0.606
	inst	-0.104	-0.109	0.066	-1.579	-0.227	0.064
Political Instability	ecks	-0.270	-0.265	0.041	-6.600	-0.340	-0.151
	death	-0.302	-0.296	0.039	-7.693	-0.379	-0.218
	demostab	0.336	0.326	0.037	9.157	0.257	0.421
	demoinst	-0.037	-0.031	0.064	-0.575	-0.156	0.123
	dictatur	-0.285	-0.281	0.036	-7.974	-0.357	-0.206

Correlations (Dimension 1):

Latent variable	Manifest variables	Standardized loadings	Communalities	Redundancies
Agricultural Inequality	gini	0.977	0.955	
	farm	0.986	0.972	
	rent	0.516	0.266	
Industrial Development	gnpr	-0.950	0.903	
	labo	0.955	0.912	
	inst	-0.352	0.124	0.077
Political Instability	ecks	-0.816	0.665	0.414
	death	-0.794	0.630	0.392
	demostab	0.866	0.749	0.466
	demoinst	-0.094	0.009	0.006
	dictatur	-0.733	0.537	0.334

Loading = Coefficiente di Regressione di ogni VM sulla corrispondente VL
= Coefficiente di correlazione se le VM sono standardizzate

Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

Risultati - dopo aver invertito il segno della VL

Latent variable	Manifest variables	Outer weight
Agricultural Inequality	gini	0.460
	farm	0.516
	rent	0.081
Industrial Development	gnpr	0.511
	labo	-0.538
	inst	0.104
Political Instability	ecks	0.270
	death	0.302
	demostab	-0.336
	demoinst	0.037
	dictatur	0.285

Le VL sono ottenute al netto del segno (come ad esempio le componenti principali in un ACP!!)

Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

Latent variable scores (Dimension 1):

Scores Standardizzati delle VL

	Agricultural Inequality	Industrial Development	Political Instability
Argentina	0.953	0.238	0.751
Australia	1.265	1.371	-1.601
Austria	0.404	0.253	-0.464
Belgium	-0.848	1.530	-0.881
Bolivia	1.115	-1.584	1.503
Brasil	0.789	-0.654	0.268
Canada	-1.539	1.680	-0.972
Chile	1.239	-0.324	0.016
Colombia	0.819	-0.443	0.810
Costa Rica	0.939	-0.484	0.302
Cuba	0.734	-0.182	1.694
Denmark	-1.996	0.821	-1.528
Dominican Republic	0.720	-0.737	0.542
Ecuador	0.976	-0.690	0.956
Egypt	0.464	-1.086	0.865
Salvador	0.824	-0.718	0.426
Finland	-1.020	0.304	-0.262
France	-0.910	0.804	0.492
Guatemala	1.006	-0.959	1.099
...

Per ogni osservazione nel modello è possibile ottenere uno score per ogni VL

Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

Inner model (Dimension 1):

Risultati - Modello Strutturale

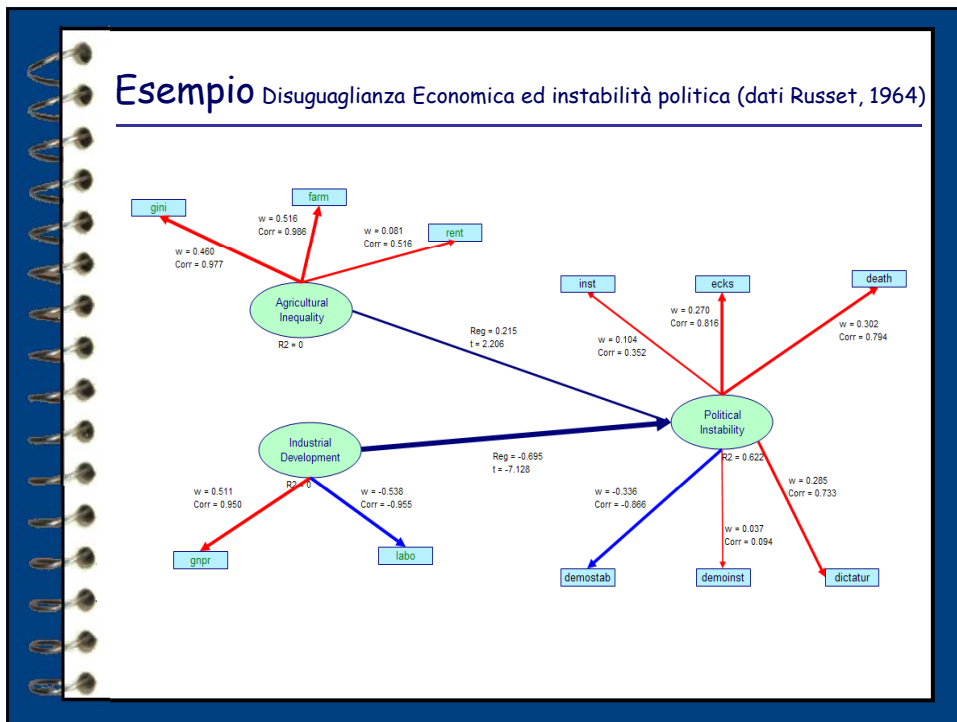
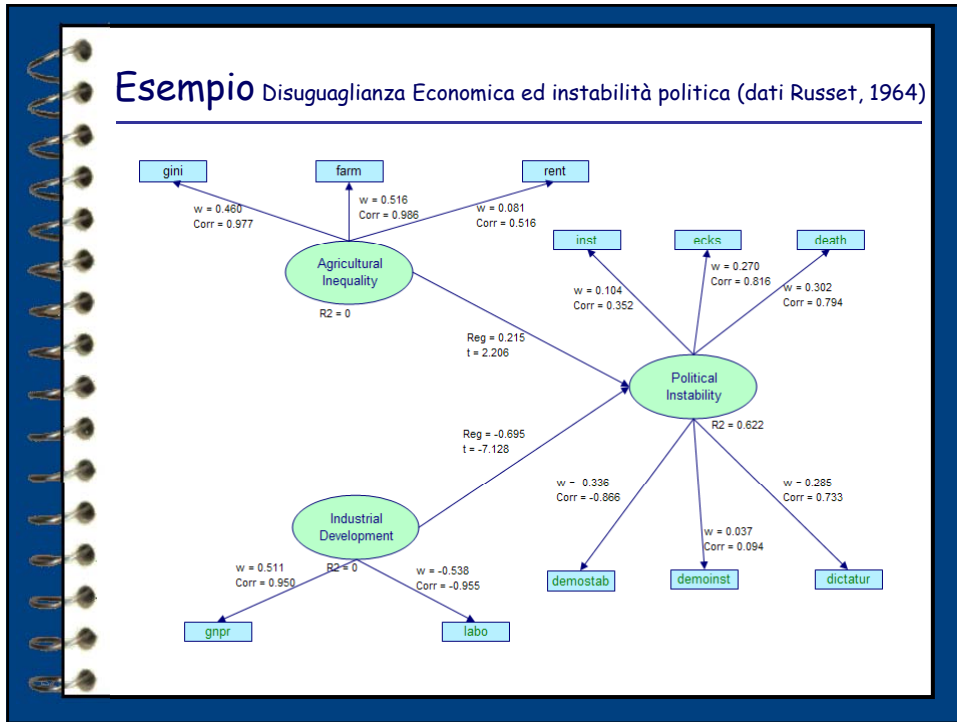
R² (Political Instability / 1):

R ²	R ² (Bootstrap)	Std. deviation	Lower bound (95%)	Upper bound (95%)
0.622	0.632	0.063	0.532	0.762

Path coefficients (Political Instability / 1):

Latent variable	Value	Standard error	t	Pr > t
Intercept	0.000	0.092	0.000	0.000
Agricultural Inequality	0.215	0.097	2.206	0.033
Industrial Development	-0.695	0.097	-7.128	0.000

Value(Bootstrap)	Standard error(Bootstrap)	Lower bound (95%)	Upper bound (95%)
0.000	0.000	0.000	0.000
0.251	0.104	0.039	0.487
-0.673	0.077	-0.846	-0.466



Esempio Disuguaglianza Economica ed instabilità politica (dati Russet, 1964)

Test sull' R^2 :

$$\frac{[R^2/p]}{[(1-R^2)/(N-p-1)]}$$

si distribuisce come una $F_{(N-p-1)}$

R^2

0.622

Interpretare il modello interno

Qual'è l'importanza relativa di una specifica variabile esogena nella predizione di una variabile endogena?

I segni dei path coefficients (effetti diretti) sono coerenti con quelli delle correlazioni?

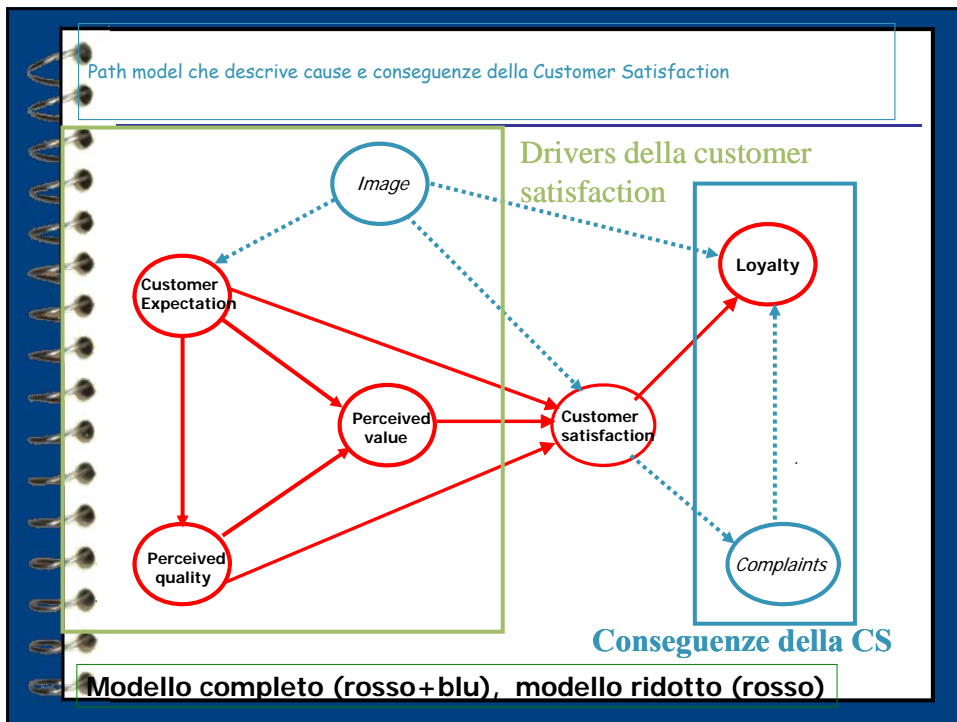
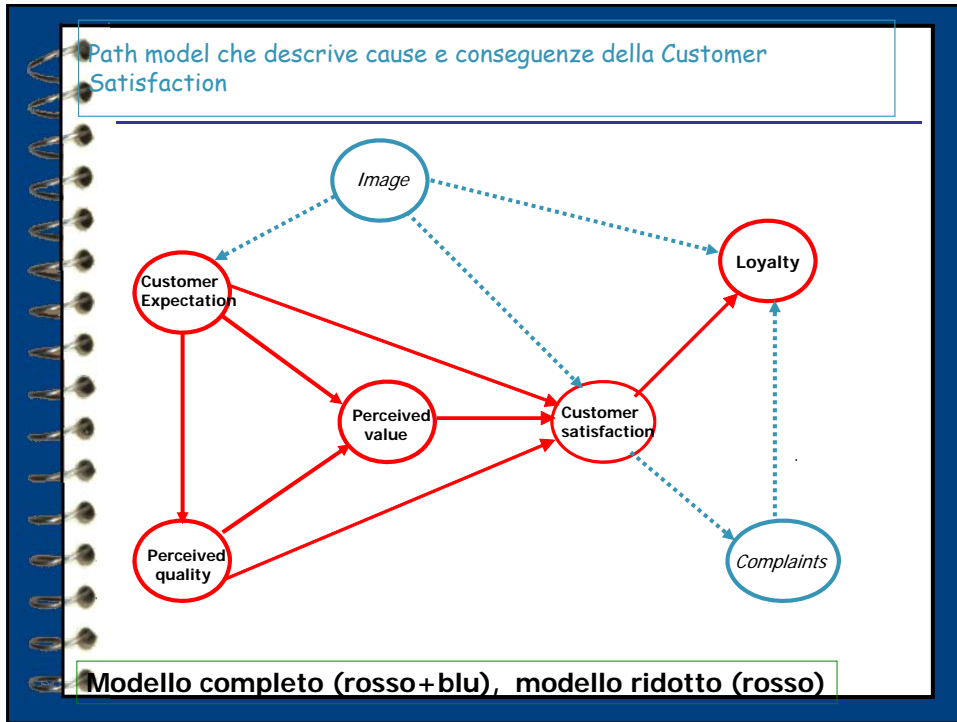
Impatto e contributo di ogni variabile latente esogena sull'instabilità politica:

	Agricultural Inequality	Industrial Development
Correlation	0.432	-0.762
Path coefficient	0.215	-0.695
Contribution to R^2 (%)	14.917	85.083
Cumulative %	14.917	100.000

L'European Customer Satisfaction Index (ECSI)

ECSI è un indicatore economico che misura la Customer Satisfaction.

Adattamento dello Swedish Customer Satisfaction Barometer e dell'American Customer Satisfaction Index (ACSI) proposto da Claes Fornell.



Determinanti della Customer Satisfaction

1. Qualità percepita

Si considerano 2 tipi di qualità percepita:

- qualità del prodotto e del servizio offerto (hardware) e
 - qualità dei servizi collaterali al prodotto/servizio offerto (software/humanware)
- (Fornell et al. 1996).

La percezione della qualità del prodotto è la valutazione del consumatore/utente dell'utilizzo di un determinato bene e quindi della sua intrinseca qualità (legata ai suoi specifici attributi).

La percezione della qualità dei servizi collaterali è la valutazione della sua esperienza legata all'utilizzo e quindi agli elementi interattivi associati al consumo/fruizione (comportamento del personale, atmosfera dell'ambiente di acquisto, ecc)

=> Hardware e Software/Humanware dovrebbero avere un effetto diretto e positivo sulla CS

Determinanti della Customer Satisfaction

2. Valore

Il valore è stato inserito nel modello in coerenza con le teorie economiche (Howard, Sheth, 1969) che gli riconoscono un ruolo essenziale nell'influenzare la CS.

Il valore è il livello percepito della qualità del prodotto/servizio relativo al prezzo pagato (rapporto qualità percepita/ prezzo)

=> Il valore dovrebbe avere un effetto diretto e positivo sulla soddisfazione (Anderson, Sullivan, 1993; Fornell, 1992) ed essere positivamente influenzato dalla qualità percepita.

Determinanti della Customer Satisfaction

3. Immagine

Si riferisce al marchio e al tipo di impulsi che il consumatore riceve dall'impresa. Tale variabile è stata introdotta nel modello NCSB (Norwegian Customer Satisfaction Barometer) (Andreassen, Lindestad, 1998). Recenti studi dimostrano che essa è un importante componente del modello di CS (Martensen, 2000).

=> Dovrebbe avere un effetto positivo sulla soddisfazione e sulla fedeltà e sulle attese.

Nota: Secondo Johnson (2001) l'immagine viene modellata per influire sulle percezioni della qualità (tale link non è però inserito nel modello ECSI).

Determinanti della Customer Satisfaction

4. Aspettative

Le aspettative si riferiscono al livello di qualità che i consumatori/utenti si aspettano di ricevere ed è il risultato di una precedente esperienza che essi hanno avuto con l'acquisto di prodotti o servizi da un'impresa.

Conseguenze della Customer Satisfaction

1. Lamentele/Reclami

Si riferisce all'intensità delle lamentele (riscontri negativi) oppure, nel caso in cui le lamentele non siano direttamente espresse, all'esigenza di lamentarsi (comunicare il dissenso) e al modo in cui l'impresa reagisce.

=> Un aumento della Customer Satisfaction dovrebbe far decrescere la loro incidenza

Conseguenze della Customer Satisfaction

2. Fedeltà

La fiducia del consumatore (fidelizzazione) assicura entrate economiche e minimizza la possibilità di defezione (abbandono) se la qualità diminuisce. Inoltre agevola il passaparola positivo riducendo costi per l'acquisizione di nuovi clienti (Anderson, Fornell, 2000).

E' misurata dall'intenzione di riacquistare, dalla tollerabilità/sensibilità agli aumenti di prezzo e dall'intenzione di consigliare ad altri il prodotto/servizio.

=>Una buona immagine e un'alta soddisfazione dovrebbero aumentare la fiducia.

=>Se la relazione tra lamentele e fiducia è positiva l'impresa riesce a trasformare i consumatori che si lamentano in consumatori fiduciosi. Viceversa l'impresa non ha saputo rimediare adeguatamente al dissenso.

Misurazione della CS per l'industria della telefonia mobile :

Esempi di variabili latenti e manifeste

La soddisfazione

In coerenza con la teoria si assumo 3 quesiti per misurare la CS.

- 1) **Soddisfazione totale:** "considerando globalmente la sua esperienza con l'impresa X, quanto si ritiene soddisfatto su una scala che va da assolutamente insoddisfatto ad assolutamente soddisfatto?"
- 2) **Soddisfazione rispetto alle attese:** "A che livello l'impresa X ha risposto alle sue aspettative, su una scala che va da 'molto meno di quanto mi aspettassi' a 'molto più di quanto mi aspettassi'?"
- 3) **Soddisfazione rispetto all'ideale:** 'Immagini un'impresa perfetta in ogni suo aspetto. L'impresa X quanto si avvicina a questo suo ideale, su una scala che va da 'molto lontana' a 'molto vicina'?"

Il questionario

Image (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 (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 Value (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"?

Dr. Laura Trincherà, Università degli Studi di Napoli " Federico II "

Il questionario (segue)

Customer Satisfaction (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 ?

Customer Complaints (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 you

Dr. Laura Trincherà, Università degli Studi di Napoli " Federico II "

I. Studio del modello completo usando l'approccio di Fornell

- Le variabili latenti V sono trasformate da una scala " 1-10 " ad una " 0-100 " :

$$x = \frac{V - 1}{9} \times 100$$

- Opzioni di stima: Mode A + Schema Centroide
- Ciascuna VL è stimata come media pesata delle sue variabili manifeste.
- Il PLS Path modelling è usato per stimare i pesi.
- Path coefficients e livelli significativi sono calcolati mediante la regressione multipla. Fornell calcola i livelli significativi mediante cross-validation (Jack-knife).

Risultati : i weights

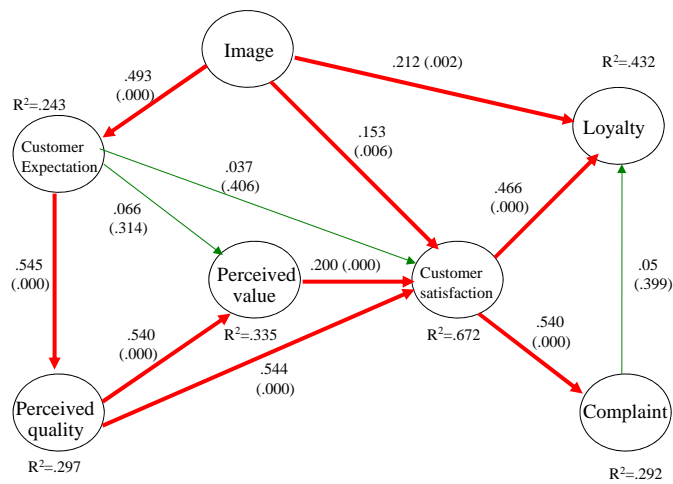
Outer Model

Variable	Weight
IMAGE outward	
IMAG1	.0145
IMAG2	.0126
IMAG3	.0136
IMAG4	.0176
IMAG5	.0144
CUS_EXP outward	
CUEX1	.0231
CUEX2	.0224
CUEX3	.0253
PER_QUAL outward	
PERQ1	.0098
PERQ2	.0085
PERQ3	.0118
PERQ4	.0094
PERQ5	.0084
PERQ6	.0095
PERQ7	.0129

Outer Model

Variable	Weight
PER_VAL outward	
PERV1	-.0239
PERV2	-.0247
ECSI outward	
CUSA1	.0158
CUSA2	.0231
CUSA3	.0264
CUS_COMP outward	
CUSCO	.0397
CUS_LOY outward	
CUSL1	-.0185
CUSL2	-.0061
CUSL3	-.0225

ECSI Path model per un "provider di Telefonia mobile" (regressione sulle variabili standardizzate)



Calcolo delle variabili latenti

Esempio : Customer Satisfaction Index

$$CSI = \frac{0.0158 \times C_sat1 + 0.0231 \times C_sat2 + 0.0264 \times C_sat3}{0.0158 + 0.0231 + 0.0264}$$

Mean and standard deviation of the latent variables

	N	Minimum	Maximum	Mean	Std. Deviation
IMAGE	250	26.49	100.00	72.6878	13.7660
CUSTOMER EXPECTATION	250	25.85	100.00	72.3198	14.1259
PERCEIVED QUALITY	250	23.95	100.00	74.5765	14.2573
PERCEIVED VALUE	250	.00	100.00	61.5887	20.5987
CUSTOMER SATISFACTION	250	23.68	100.00	71.2876	15.3417
COMPLAINT	250	.00	100.00	67.4704	25.2684
LOYALTY	250	1.29	100.00	69.1757	21.2668

Le equazioni di regressione

Linear Model

```

=====
Block          Mult. RSq
=====
IMAGE          .0000
CUS_EXP        .2431
PER_QUAL       .2971
PER_VAL        .3351
ECSI           .6717
CUS_COMP       .2916
CUS_LOY        .4318
=====
    
```

Path coefficients (Standardized coefficients)

```

=====
          IMAGE  CUS_EXP  PER_QUAL  PER_VAL  ECSI  CUS_COMP  CUS_LOY
=====
IMAGE      .000      .000      .000      .000      .000      .000      .000
CUS_EXP    .493      .000      .000      .000      .000      .000      .000
PER_QUAL   .000      .545      .000      .000      .000      .000      .000
PER_VAL    .000     -.066     -.540     .000      .000      .000      .000
ECSI       .153      .037      .544     -.200      .000      .000      .000
CUS_COMP   .000      .000      .000      .000      .540      .000      .000
CUS_LOY    -.212      .000      .000      .000     -.466     -.050      .000
=====
    
```

Correlazioni tra immagini e variabili latenti

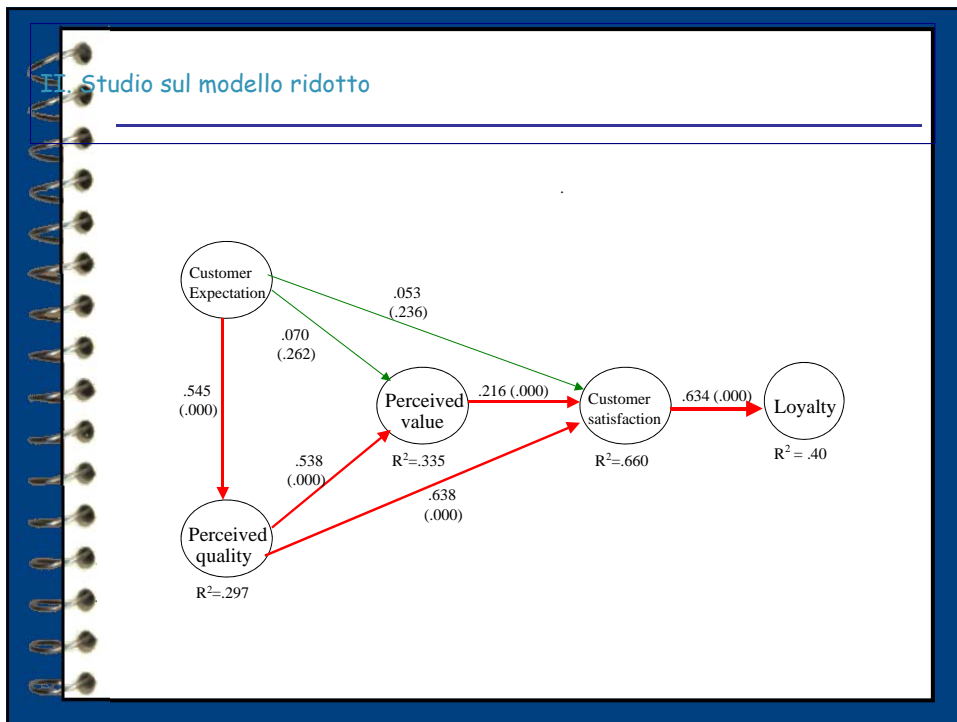
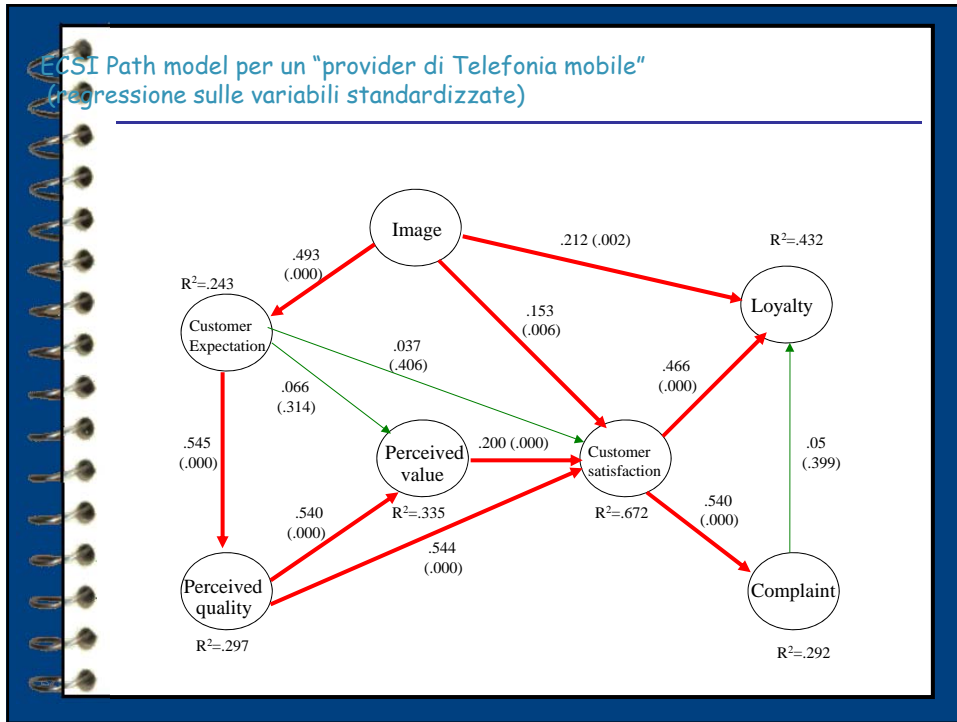
Correlations of latent variables

	IMAGE	CUS_EXP	PER_QUAL	PER_VAL	ECSI	CUS_COMP	CUS_LOY
IMAGE	1.000						
CUS_EXP	.493	1.000					
PER_QUAL	.731	.545	1.000				
PER_VAL	-.508	-.360	-.576	1.000			
ECSI	.671	.481	.791	-.604	1.000		
CUS_COMP	.469	.250	.537	-.348	.540	1.000	
CUS_LOY	-.548	-.366	-.524	.517	-.635	-.401	1.000

Correlazioni tra variabili manifeste e latenti

	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

Correlazioni in valore assoluto minore di 0.5 non sono riportate



I nuovi weights PLS

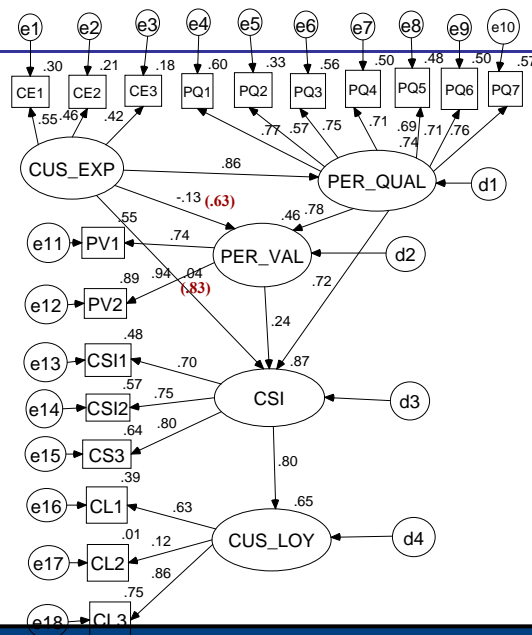
	Weight	Relative Weight
CE1	.0237	.336
CE2	.0206	.292
CE3	.0262	.372
PQ1	.0098	.139
PQ2	.0085	.121
PQ3	.0118	.168
PQ4	.0094	.134
PQ5	.0084	.119
PQ6	.0095	.135
PQ7	.0129	.183

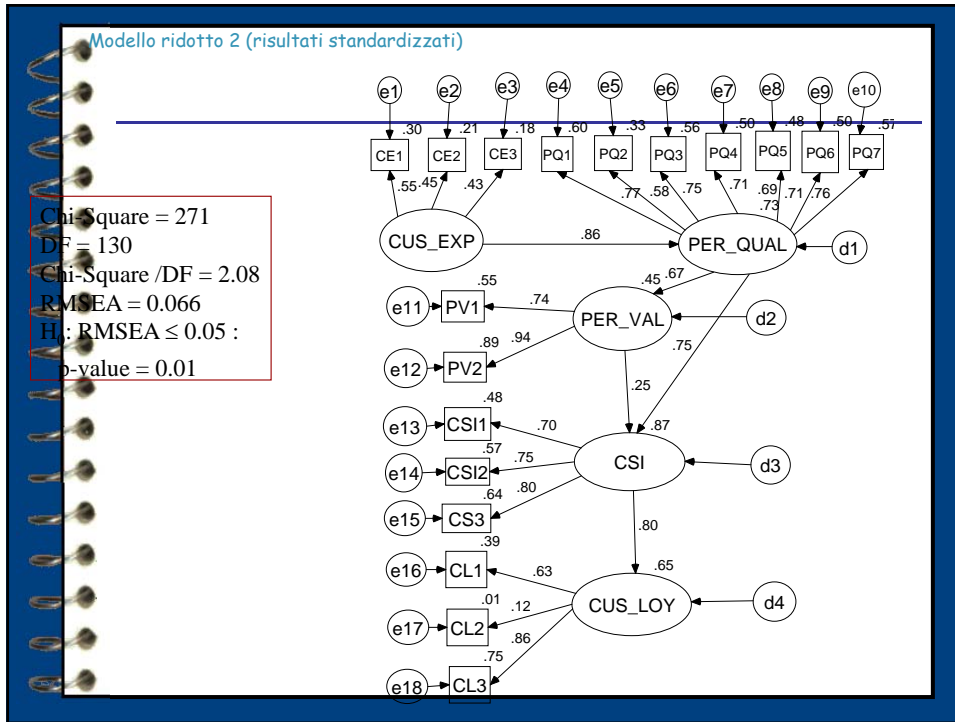
	Weight	Relative Weight
PV1	-.0239	.492
PV2	-.0247	.508
CS1	.0157	.241
CS2	.0240	.368
CS3	.0256	.392
CL1	-.0188	.405
CL2	-.0050	.108
CL3	-.0226	.487

Per ciascuna variabile i pesi (weight) relativi sommano a 1.

III. Studio del modello ridotto con LISREL : Model 1

Chi-Square = 271
 DF = 128
 Chi-Square / DF = 2.12





Stima delle VL

Ciascuna VL è stimata come media pesata delle sue manifeste.

$$E(\text{score delle attese}) = \frac{\text{cov}(CE1, CUS_EXP) * CE1 + \text{cov}(CE2, CUS_EXP) * CE2 + \text{cov}(CE3, CUS_EXP) * CE3}{\text{cov}(CE1, CUS_EXP) + \text{cov}(CE2, CUS_EXP) + \text{cov}(CE3, CUS_EXP)}$$

Il pesi delle covarianze sono interessanti

$$\text{Cov}(CE1, CUS_EXP) = \text{Cor}(CE1, CUS_EXP) * \text{STD}(CE1) * \text{STD}(CUS_EXP)$$

Per essere importante una variabile manifesta dovrebbe essere ben correlata con la sua variabile latente e avere un'alta varianza.

Stima delle variabili latenti.

Le covarianze sono calcolate usando il modello causale. Sono output standard dei programmi che applicano il LISREL

Esempio :

$$\begin{aligned} \text{Cov}(CE1, CUS_EXP) &= \text{Cov}(\lambda_{11}CUS_EXP + e_1, CUS_EXP) \\ &= \lambda_{11} \text{Var}(CUS_EXP) \end{aligned}$$

in cui λ_{11} e $\text{Var}(CUS_EXP)$ sono stimate con il LISREL .

Formula semplificata per la stima delle variabili latenti

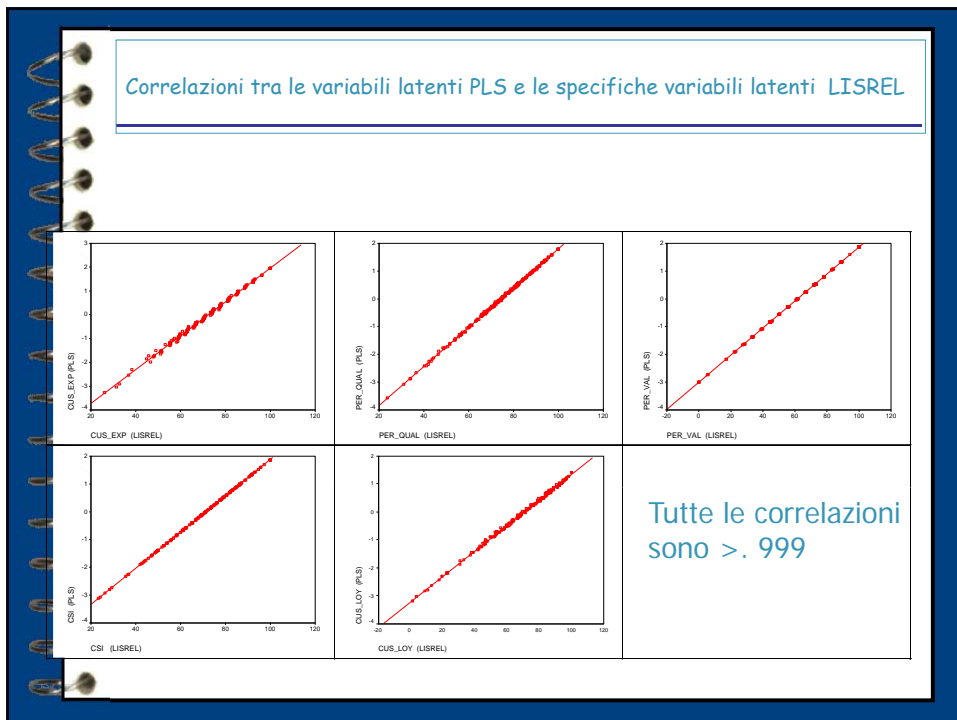
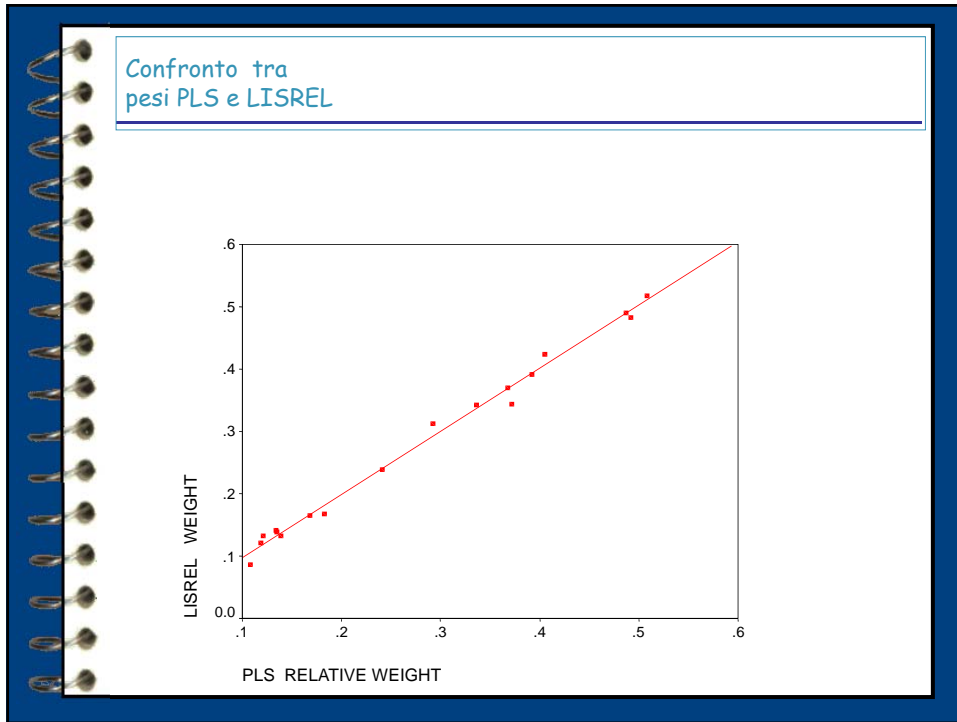
Da $\text{Cov}(CE_j, CUS_EXP) = \lambda_{1j} \text{Var}(CUS_EXP)$

e

$$Y_1 = \frac{\text{cov}(CE1, CUS_EXP) * CE1 + \text{cov}(CE2, CUS_EXP) * CE2 + \text{cov}(CE3, CUS_EXP) * CE3}{\text{cov}(CE1, CUS_EXP) + \text{cov}(CE2, CUS_EXP) + \text{cov}(CE3, CUS_EXP)}$$

si deduce la formula semplificata

$$Y_1 = \frac{\lambda_{11}CE1 + \lambda_{12}CE2 + \lambda_{13}CE3}{\lambda_{11} + \lambda_{12} + \lambda_{13}}$$



Conclusioni

- LISREL non funziona su tutto il modello.
- LISREL dà risultati migliori per il modello strutturale perchè le variabili latenti sono space-free
- PLS dà migliori risultati per il modello outer perchè ciascuna variabile latente è vincolata ad essere nello spazio definito dalle sue VM.
- Se ciascuna variabile latente LISREL è stimata come media pesata delle sue manifeste, LISREL e PLS danno le stesse stime per le variabili latenti (almeno in questo esempio).

Problemi di stima di ECSI con LISREL

- LISREL non converge con il modello completo (modello ridotto = senza "Reclami").
- Per l'identificazione del modello, è necessario aggiungere il vincolo della varianza unitaria dell'errore legato alle singole variabili manifeste della "Loyalty".
- Per migliorare la scarsa bontà dell'adattamento, LISREL propone strane modifiche:
 - stime delle covarianze degli errori di variabili non collegate;
 - aggiunge legami di dipendenza non confermati dalla teoria (i.e. "Imagine" e "Perceived Quality", "Imagine" e "Perceived Value");
 - aggiunge legami tra variabili latenti e variabili manifeste non originariamente inserite per misurare tali variabili (i.e. "Loyalty" con la prima variabile manifesta di "Perceived Value").
- Soluzioni inaccettabili possono essere ottenute (i.e. Varianze negative della "Customer Satisfaction").

Il PLS Path Modeling:

gli indici di bontà del modello

Poichè il PLS-PM è basato su un sistema di regressioni interdipendenti, e poichè ogni SEM è composto da due sottomodelli (il modello di misura ed il modello strutturale), **diversi indici** devono essere presi in considerazione per valutare la bontà di un modello PLS:

L'indice di **COMUNALITA'**

Misura la bontà del modello di misura

L'**R²** di ogni relazione strutturale

Misura la capacità predittiva del modello

interno

L'indice di **RIDONDANZA**

Misura la bontà del modello interno

Il **GoF (Goodness of Fit index)**

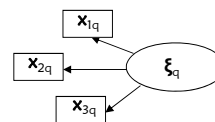
È un indice di bontà globale del modello

La caratteristica di tutti questi indici è che tutti **misurano la quantità di variabilità spiegata** dalle singole parti del modello, in altre parole sono tutti indici che misurano la capacità predittiva del modello!

Il PLS Path Modeling:

L'indice di Comunalità (1/2)

Per ogni blocco q ...



...l'indice di comunalità è la media delle correlazioni al quadrato tra ogni variabile manifesta del blocco e la variabile latente:

$$com_q = \frac{1}{p_q} \sum_{p=1}^{p_q} cor^2(x_{pq}, \xi_q)$$

N.B. Per ogni variabile manifesta x_{pq} la comunalità è il quadrato della correlazione:

$$com_{pq} = cor^2(x_{pq}, \xi_q)$$

Il PLS Path Modeling:

L'indice di Comunalità (2/2)

La bontà del modello di misura (**reliability of latent variables**) è misurata attraverso la quantità di varianza che la VL riesce a catturare dai propri indicatori (**comunalità media**) rispetto alla quantità di variabilità totale degli indicatori (i.e. delle VM).

dove λ_{pq} è il loading associato al generico indicatore del blocco q e $\text{var}(\varepsilon_{pq}) = 1 - \lambda_{pq}^2$.

$$AVE_q = \frac{\sum_{p=1}^{P_q} \lambda_{pq}^2 \times \text{var}(\xi_q)}{\left[\sum_{p=1}^{P_q} \lambda_{pq}^2 \times \text{var}(\xi_q) \right] + \sum_{p=1}^{P_q} \text{var}(\varepsilon_{pq})}$$

Nel caso di VM **standardizzate** →

$$Com_q = \frac{\sum_{p=1}^{P_q} \lambda_{pq}^2}{P_q}$$

e →

$$Com = \frac{\sum_{q=1}^Q (P_q \times Com_q)}{P}$$

Il PLS Path Modeling:

Derivazione dell'AVE (Average Variance Extracted)

Sapendo, che in uno schema formativo:

$$x_{pq} = \lambda_{pq} \xi_q + \varepsilon_{pq}$$

$$\text{e che } \sum_{p=1}^{P_q} \text{Var}(x_{pq}) = \sum_{p=1}^{P_q} \lambda_{pq}^2 \text{Var}(\xi_q) + \sum_{p=1}^{P_q} \text{Var}(\varepsilon_{pq})$$

Ed essendo le VL standardizzate ($\text{Var}(\xi_q) = 1$), si ottiene:

$$AVE_q = \frac{\sum_{p=1}^{P_q} \lambda_{pq}^2}{\sum_{p=1}^{P_q} \text{Var}(x_{pq})}$$

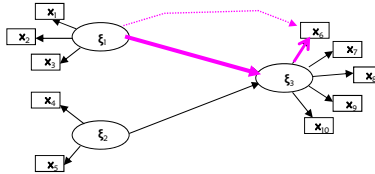
nel caso di VM **standardizzate** :

$$Com_q = \frac{\sum_{p=1}^{P_q} \lambda_{pq}^2}{P_q}$$

Il PLS Path Modeling:

L'indice di Ridondanza

Per ogni blocco **endogeno** j ...



...l'indice di ridondanza è la quantità di variabilità delle VM associate al j -esimo blocco spiegata dalle VL esogene. Può quindi essere ottenuta come variabilità media del blocco di VM endogene spiegata dalla relazione strutturale e dal modello di misura.

Il PLS Path Modeling:

L'indice di Ridondanza

La Ridondanza è la **media aritmetica delle varianze** delle VM del blocco endogeno, **spiegata dalle VL esogene**.

$$Red_{pj}^2 = \frac{Var[\lambda_{pj} \lambda_{pm} \xi_j]}{Var[x_{pj}]}$$

In modo più immediato la Ridondanza può essere ottenuta come il prodotto tra l'indice di comunalità associato al j -esimo blocco endogeno ed il valore dell' R^2 della relazione strutturale:

$$Ridondanza_j = Comunalità_j \times R^2(\xi_j, \xi_m \text{ che spiegano } \xi_j)$$

Il PLS Path Modeling:

Il Goodness of Fit Index (GoF) (1/2)

Le necessità che hanno portato alla costruzione del GoF:

Non è ben chiaro quale criterio sia ottimizzato dall'algoritmo iterativo del PLS-PM...questo non permetteva di identificare direttamente la bontà del modello. Non avendo un obiettivo è difficile stabilire quanto ci si è avvicinati!!

Un indice globale doveva essere, quindi, in grado di:

- essere un compromesso tra la performance del modello interno e quella del modello esterno;
- essere calcolabile anche per i modelli di tipo LISREL, per permettere un confronto diretto tra i due approcci;
- avere un limite superiore, in modo da sapere quanto bene stesse facendo uno specifico modello rispetto al suo massimo!

Il PLS Path Modeling:

Il Goodness of Fit Index (GoF) (2/2)

$$GoF = \sqrt{\frac{1}{\sum_{q=1}^Q P_q} \sum_{q=1}^Q \sum_{p=1}^{P_q} Cor^2(\mathbf{x}_{pq}, \xi_q)} \times \sqrt{\frac{1}{J} \sum_{j=1}^J R^2(\xi_j, \xi_m \text{ explaining } \xi_j)}$$

Validazione del modello esterno

Un indice che misura direttamente la bontà del modello interno è la comunaltà media, cioè la media aritmetica delle correlazioni al quadrato tra ogni VM e la corrispondente variabile latente, cioè come comunaltà media!

Validazione del modello interno

Un indice per misurare la bontà del modello interno è la media aritmetica dei valori degli R^2 di tutte le regressioni che compongono il modello strutturale.

Il PLS Path Modeling:

GoF normalizzato

Da una PCA si sa che $\rightarrow \sum_{p=1}^{P_q} \text{cor}^2(\mathbf{x}_{pq}, \xi_q) \leq \lambda_q$

Il più grande autovettore della matrice $\mathbf{X}_q' \mathbf{X}_q$

...e da una AC si sa che:

$\rightarrow \sum_{j=1}^J R^2(\xi_j, \xi_m \text{ explaining } \xi_j) \leq \rho_q$

La prima c of the CA of the \mathbf{X}_q

Quindi riportando ogni termine del GoF al suo massimo si ottiene il GoF normalizzato come:

$$\text{GoF} = \sqrt{\frac{1}{\sum_{q=1}^Q \rho_q} \sum_{q=1}^Q \frac{\sum_{p=1}^{P_q} \text{Cor}^2(\mathbf{x}_{pq}, \xi_q)}{\lambda_q} \times \frac{1}{J} \sum_{j=1}^J \frac{R^2(\xi_j, \xi_m \text{ explaining } \xi_j)}{\rho_j}}$$

Il PLS Path Modeling:

Validità discriminante delle VL

Le diverse **variabili latenti** devono essere correlate tra loro, ma devono anche misurare concetti diversi! In altre parole bisogna poter discriminare tra i concetti spiegati da ogni variabile latente.

Bisogna testare quindi che la **correlazione** tra VL sia **significativamente minore di 1**:

$$H_0 : \text{cor}(\xi_q, \xi_{q'}) = 1$$

Regola di decisione:

L'ipotesi nulla è rigettata se:

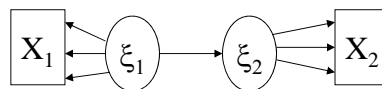
1) L'**intervallo di confidenza** al **95%** delle correlazioni tra le singole VL non contiene 1 (NB. Si tratta di intervalli di confidenza ottenuti con metodi di ricampionamento quali il bootstrap o lo jackknife);

2) $(\text{AVE}_q \text{ and } \text{AVE}_{q'}) > \text{cor}^2(\xi_q, \xi_{q'})$ in altre parole se la parte di variabilità condivisa tra una VL ed il suo blocco di indicatori è maggiore della parte di variabilità che la stessa VL condivide con altre VL! Ogni VL deve essere più correlata al suo blocco di indicatori che alle altre VL che rappresentano altri blocchi di indicatori.

Casi particolari del PLS-PM

- SEM-PLS con un solo blocco
 - Analisi in Componenti Principali PCA (Hotelling, 1933)
 - Regressione Multipla
- SEM-PLS con due blocchi
 - Analisi Correlazioni Canoniche (ACC) (Hotelling, 1936)
 - Analisi Interbatterie di Tucker (Tucker, 1958)
 - Analisi della Ridondanza (AR) (Van de Wollenberg, 1977) -
 - Analisi in Componenti Principali con Variabili Strumentali (PCAIV) (Rao, 1964)
 - Analisi in Componenti Principali in un Sottospazio di Riferimento (PCAR) (D'Ambra & Lauro, 1982)
 - Regressione PLS (Wold)
- SEM-PLS "multiblock"
 - Analisi delle Correlazione Canoniche Generalizzata di Horst (Horst, 1965)
 - Analisi delle Correlazione Canoniche Generalizzata di Carroll (Carroll, 1968)
 - Analisi Fattoriale Multipla (Escofier & Pages, 1994)

Casi particolari per un PLS-PM con 2 blocchi



A = RIFLESSIVO
B = FORMATIVO

Metodi di stima per i pesi esterni w_q

$\xi_1 = X_1 w_1$	$\xi_2 = X_2 w_2$	Metodo	Deflazione (Residui)
A	A	Regressione PLS Regressione di X_2 su X_1	Solo di X_1
B	A	Analisi della Ridondanza di X_2 rispetto a X_1	Solo di X_1
A	A	Analisi Interbatterie di Tucker	Per entrambe X_1 e X_2
B	B	Analisi Correlazioni Canoniche	Per entrambe X_1 e X_2

Alcuni casi particolari: ACP

L'ACP mira ad individuare delle variabili latenti (le Componenti Principali) $\xi_q = X_q a_q$ all'interno della matrice X in grado di spiegare "al meglio" le variabili stesse, cioè: $\max\{\text{var}(Xa_q)\}$

$$R_{11} a_q = \lambda_q a_q$$

↖ Autovalori
↘ Autovettori

Alcuni casi particolari : ACC

Cerca all'interno dei due gruppi di variabili delle componenti $\xi_q = X_q b_q$ e delle componenti $\xi_j = Y_j c_j$ che siano **massimamente correlate** tra loro \Rightarrow

$$\max\{\text{corr}(\xi_q, \xi_j)\}$$

$$R_{11}^{-1} R_{12} R_{22}^{-1} R_{21} b_q = r_q^2 b_q$$

$$R_{22}^{-1} R_{21} R_{11}^{-1} R_{12} c_j = r_j^2 c_j$$

Alcuni casi particolari : Tucker

Cerca all'interno dei due gruppi di variabili delle componenti $\xi_q = X_q d_q$ e delle componenti $\xi_j = Y_j g_j$ con lo scopo di spiegare al meglio le variabili originarie e contemporaneamente individuare le relazioni esistenti tra i due gruppi di variabili $\Rightarrow \max\{\text{cov}^2(\xi_q, \xi_j)\}$

$$R_{12}R_{21}d_q = \gamma_q^2 d_q$$

$$R_{21}R_{12}g_j = \gamma_j^2 g_j$$

Alcuni casi particolari: AR

Cerca all'interno della matrice X dei fattori $\xi_q = X_q m_q$ in grado di spiegare al meglio il gruppo di variabili Y $\max\{\text{corr}^2(\xi_q, Y)\}$

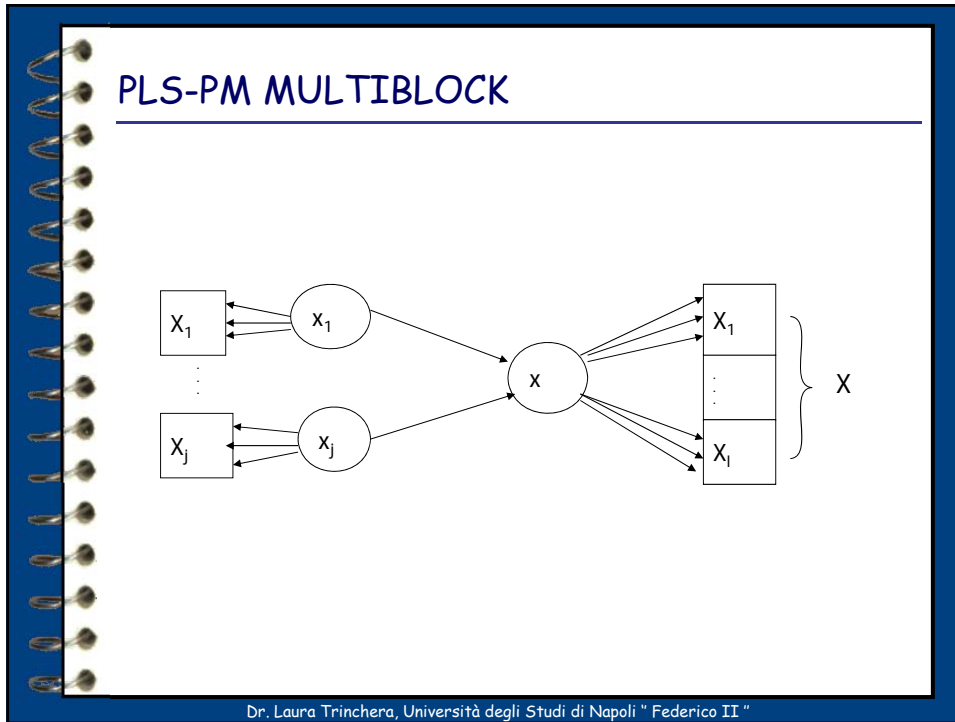
N.B. è un'analisi **asimmetrica**!! X è la matrice delle variabili INDIPENDENTI

$$R_{11}^{-1}R_{12}R_{21}m_q = \varphi_q m_q$$

(Van de Wollenberg, 1977)

$$R_{21}R_{11}^{-1}R_{12}n_q = \varphi_q n_q$$

(Lauro & D'Ambra, 1982)



Casi particolari per un PLS-PM multiblock

Internal Estimation Schemes

<i>External Estimation Modes</i>	Centroid Sign(correlation)	Factorial Correlation	Path Weighting Correlation or Regression
A PLS First Component	Area for Future Research - PLS version of GCCA by Horst - Tenenhaus (2004)	Area for Future Research PLS version of GCCA by Carroll	- Lohmöller's Split PCA - Multiple Factorial Analysis by Escofier and Pagès - Horst's Maximum Variance Algorithm - Multiple Co-Inertia Analysis (ACOM) by Chesel and Hanafi
B OLS All PLS Components	- Generalised Canonical Correlation Analysis (Horst's SUMCOR criterion) - Mathes (1993) & Hanafi (2004)	Generalised Canonical Correlation Analysis (Carroll)	Area for Future Research
PLS k components	Multiple and Partial Non Symmetrical Analysis By substituting the matrix Y of the dependent variables to the super-block X and taking into account the role of a group Z of control variables		

Deflazione: Dipende dai vincoli di ortogonalità

Dr. Laura Trincherà, Università degli Studi di Napoli " Federico II "

Riferimenti Bibliografici di approfondimento per il PLS-PM (1/3)

- M.P. Bayol, A. de la Foye, C. Tellier, M. Tenenhaus :
Use of PLS Path Modeling to Estimate the European Consumer Satisfaction Index (ECSEI) Model, *Statistica Applicata - Italian Journal of Applied Statistics*, (12), 3, 361-375, 2000
- F.L. Bookstein :
Data Analysis by Partial Least Squares, in: C. Fornell (ed.), *A Second Generation of Multivariate Analysis*, Praeger, New York, 348-366, 1982
- W.W. Chin :
The Partial Least Squares Approach to Structural Equation Modeling, in: G.A. Marcoulides (ed.), *Modern Methods for Business Research*, Lawrence Erlbaum Associates, New Jersey, 295-336, 1998
- T. Dijkstra :
Some comments on maximum likelihood and Partial Least Squares Methods, *Journal of Econometrics*, 22, 67-90, 1983
- R.F. Falk & N.B. Miller :
A primer for soft modeling, The University of Akron Press, 1992
- C. Fornell :
A National Customer Satisfaction Barometer: The Swedish Experience, *Journal of Marketing*, (56), 6-21, 1992
- C. Fornell & F.L. Bookstein :
A Comparative Analysis of Two Structural Equation Models: LISREL and PLS Applied to Market Data, in: C. Fornell (ed.), *A Second Generation of Multivariate Analysis*, Praeger, New York, 289-324, 1982
- C. Fornell & F.L. Bookstein :
Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory, *Journal of Marketing Research*, 19, 440-452, 1982
- C. Fornell & J. Cha :
Partial Least Squares, in: R.P. Bagozzi (ed.), *Advanced Methods of Marketing Research*, Blackwell Business, 52-78, 1994

Dr. Laura Trinchera, Università degli Studi di Napoli " Federico II "

Riferimenti Bibliografici di approfondimento per il PLS-PM (2/3)

- R. Grewal, J.A. Cote & H. Baumgartner :
Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing, *Marketing Science*, (23), 4, 519-529, 2004
- J. Hulland :
Use of Partial Least Squares (PLS) in Strategic Management Research: A Review of Four Recent Studies, *Strategic Management Journal*, 20, 195-204, 1999
- H.S. Jaggal :
Multicollinearity in Structural Equation Models with Unobservable Variables, *Journal of Marketing Research*, (19), 431-439, 1982
- C.B. Jarvis, S.B. MacKenzie & P.M. Podsakoff :
A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research, *Journal of Consumer Research*, (30), 199-218, 2003
- J.B. Lohmöller :
LVPLS 1.8 Program Manual: Latent variable path analysis with partial least squares estimation, Universitaet zu Koeln, Zentralarchiv fuer Empirische Sozialforschung, Köln, 1987
- J.B. Lohmöller :
Latent variable path modeling with partial least squares, Physica-Verlag, Heidelberg, 1989
- J. Roos, G.S. Yip & J.K. Johansson :
Using Partial Least Squares and LISREL to research international strategies, in: M. Gherman et al. (eds.), *Statistical Methods for Strategic Management*, Kluwer Academic Publishers, The Netherlands, 111-131, 1997
- M. Tenenhaus :
La Régression PLS, Editions Technip, Paris, 1998
- M. Tenenhaus :
L'approche PLS, *Revue de Statistique Appliquée*, 47 (2), 5-40, 1999

Dr. Laura Trinchera, Università degli Studi di Napoli " Federico II "

Riferimenti Bibliografici di approfondimento per il PLS-PM (3/3)

- M. Tenenhaus, V. Esposito Vinzi :
PLS regression, PLS path modeling and generalized Procrustean analysis: a combined approach for multiblock analysis, *Journal of Chemometrics*, 19, 145-153, John Wiley & Sons, 2005
- M. Tenenhaus, V. Esposito Vinzi, Y.M. Chatelin & C. Lauro :
PLS Path Modelling, *Computational Statistics and Data Analysis*, 48 (1), 159-205, North-Holland, The Netherlands, 2005
- M. Tenenhaus, M. Hanafi :
A bridge between PLS path modelling and multi-block data analysis, in: *Handbook of Partial Least Squares (PLS): Concepts, methods and applications*, (V. Esposito Vinzi, J. Henseler, w. Chin, H. Wang, Eds), Volume II in the series of the *Handbooks of Computational Statistics*, Springer, 2007
- M. Tenenhaus & P.L. Gonzalez :
Comparaison entre les approches PLS et LISREL en modélisation d'équations structurelles: Application à la mesure de la satisfaction clientèle, 8ème Congrès de la Société Francophone de Classification, Guadeloupe, 2001
- H. Wold :
The Fix-Point Approach to Interdependent Systems: Review and Current Outlook, in: H. Wold (ed.), *The Fix-Point Approach to Interdependent Systems*, North-Holland, Amsterdam, 1981
- H. Wold :
Soft modeling. The basic design and some extensions, in: Vol.II of Jöreskog-Wold (eds.), *Systems under indirect observation*, North-Holland, Amsterdam, 1982
- H. Wold :
Quantitative Systems Analysis: the Pedigree and Broad Scope of PLS (Partial Least Squares) Soft Modeling, in: H. Martens and H. Russwurm Jr. (eds.) *Food Research and Data Analysis*, Applied Science Publisher Ltd., 1983
- H. Wold :
Partial Least Squares, in: Vol.6 of S. Kotz & N.L. Johnson (eds.), *Encyclopedia of Statistical Sciences*, John Wiley & Sons, New York, 581-591, 1985

Dr. Laura Trincherà, Università degli Studi di Napoli " Federico II "