

UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE STATISTICHE
CORSO DI LAUREA MAGISTRALE IN
SCIENZE STATISTICHE



**Modelli a equazioni strutturali:
un confronto fra i metodi
CB-SEM e PLS-SEM**

Relatore Prof. Adriano Paggiaro
Dipartimento di Scienze Statistiche

Laureanda Alessia Parin
Matricola 2062091

Anno Accademico 2023/2024

A me e a tutti gli anni dedicati allo studio

Indice

Introduzione	1
1 Modelli a Equazioni Strutturali	3
1.1 Analisi Confermativa e Analisi Esplorativa	3
1.2 Modelli a equazioni strutturali	4
1.2.1 Modello di Misura Riflessivo	6
1.2.2 Modello di Misura Formativo	7
1.2.3 Scegliere tra i due modelli	9
1.2.4 Modello Strutturale	9
1.3 Metodi di stima dei Modelli a Equazioni Strutturali	10
1.3.1 CB-SEM	10
1.3.2 PLS-SEM	12
1.3.3 Vantaggi e svantaggi dei due metodi	14
1.3.4 Confronto e campi di applicazione	15
2 Confronto tra CB-SEM e PLS-SEM per Modelli Riflessivi	17
2.1 Introduzione	17
2.2 Modello di regressione semplice	18
2.3 Modello di regressione multivariata	25
2.4 Modello di regressione multipla	30
2.5 Modello ricorsivo	34
2.6 Sintesi	38
2.7 Modello di regressione semplice: approfondimenti	39
2.7.1 Costrutti con un numero diverso di indicatori	39
2.7.2 Standard error e Verifica d'ipotesi	42
3 Applicazioni a dati reali	47
3.1 Introduzione	47
3.2 Modello di regressione semplice	48
3.3 Modello di regressione multivariata	50
3.4 Modello di regressione multipla	51
3.5 Modello ricorsivo	54
3.6 Modello ricorsivo: estensione a più costrutti	55
3.7 Sintesi	58

Conclusione	58
A Grafici delle simulazioni	61
A.1 Grafici del Modello di regressione semplice	61
A.2 Grafici del Modello di regressione multivariata	63
A.3 Grafici del Modello di regressione multipla	66
A.4 Grafici del Modello ricorsivo	70
Bibliografia	75

Introduzione

Capita molto spesso che l'interesse di ricerca sia un concetto astratto non direttamente osservabile. Di conseguenza, si rende necessaria la presenza di variabili manifeste che lo misurino. Risulta, quindi, fondamentale darne una definizione chiara, al fine di poter individuare al meglio tali variabili.

Come primo passo, è necessario verificare l'esistenza di teorie già formulate in ricerche precedenti. Se così fosse, bisogna trarre vantaggio da ciò e sfruttare i risultati già presenti per approfondirli o eventualmente per confutarli.

I Modelli a Equazioni Strutturali servono per rappresentare legami tra concetti latenti non direttamente osservabili e le variabili utilizzate per misurarli, gli indicatori. Si possono anche ipotizzare delle relazioni tra più concetti latenti. I modelli a equazioni strutturali si dividono quindi in due parti: il modello di misura e il modello strutturale. Il modello di misura, come dice il nome stesso, rappresenta il modo con cui i concetti latenti vengono misurati. Se il concetto latente è causa comune di più indicatori si dice che il costrutto è riflessivo e quindi gli indicatori saranno manifestazioni del concetto latente. Se il concetto latente è invece causato dagli indicatori, allora si dice che il modello di misura è formativo. La parte del modello che riguarda le relazioni tra i vari concetti latenti si chiama modello strutturale.

Esistono, tra gli altri, due metodi per la stima dei modelli a equazioni strutturali: il metodo Covariance Based e il Partial Least Squares. Il metodo Covariance Based si pone come obiettivo quello di trovare il modello la cui matrice di covarianza stimata si avvicini il più possibile alla covarianza osservata nei dati. Il metodo Partial Least Squares invece utilizza una serie di regressioni per stimare i parametri.

Il primo metodo è per sua natura più adatto a modelli di misura riflessivi. Esso, infatti, utilizzando solo la varianza comune, è coerente con il fatto che nei modelli di misura riflessivi la stima dei parametri non deve essere influenzata dalla parte strutturale

del modello. Il metodo Partial Least Squares, utilizzando invece tutta la varianza e applicando delle regressioni, è più adatto a modelli di misura formativi.

Tuttavia, nel corso degli anni, i ricercatori hanno provato a confrontare i due modelli anche in casi non propriamente coerenti. È quindi essenziale capire bene i principi e gli assunti alla base dei due metodi per poter applicare il più adatto al caso di studio.

Questa tesi ha l'obiettivo di dare qualche informazione in più per poter distinguere i due metodi e cercare di capire quando i loro risultati siano confrontabili. Il Capitolo 1 fornisce una breve introduzione ai Modelli a Equazioni strutturali e ai metodi di stima Covariance-Based e Partial Least Squares. Il Capitolo 2 propone delle simulazioni da quattro tipi di modelli: il Modello di Regressione Semplice, il Modello di Regressione Multivariata, il Modello di Regressione e il Modello Ricorsivo. L'obiettivo è capire se la diversa forma e forza della parte strutturale del modello influisca sulle stime dei parametri. Il Capitolo 3 conclude con delle applicazioni di suddetti modelli a dati reali, derivanti da questionari somministrati ai clienti di una catena di ristoranti. L'ultimo Capitolo riassume le principali conclusioni della tesi.

Capitolo 1

Modelli a Equazioni Strutturali

1.1 Analisi Confermativa e Analisi Esplorativa

Nel corso degli anni la ricerca ha avuto un grande sviluppo. Ciò che la rende importante e affascinante è l'interazione che esiste tra i suoi diversi campi di applicazione. Ogni nuovo risultato può confermare, confutare o ampliare la propria e altrui conoscenza. Se esiste già molta teoria alla base di un fenomeno, allora la nuova ricerca avrà lo scopo di confermare oppure confutare l'attuale teoria. Al contrario, se l'interesse della ricerca non è ancora stato approfondito in passato, essa assumerà un connotato esplorativo. La differenza tra questi due tipi di ricerca non è sempre ben definita. Lo stesso ricercatore può essere in difficoltà nel capire se stia facendo una ricerca confermativa e quindi stia partendo da qualcosa che già esiste e lo stia confermando o confutando, oppure ricerca esplorativa, ossia stia costruendo una prima base teorica di un certo fenomeno. Vista la grande quantità di studi e ricerche già presente in tutti gli ambiti di studio, l'analisi confermativa è probabilmente quella più comune. Essa parte dalla teoria presente per fare delle ipotesi a priori e poi testarle. Al contrario l'analisi esplorativa prevede che il ricercatore non abbia teoria o evidenze a supporto del modello proposto (Afthanorhan et al., 2020) e che quindi debba cercare di estrapolare delle informazioni dai dati, senza sapere a priori quali ipotesi fare e nemmeno se otterrà dei risultati interessanti. Questo tipo di analisi è spesso utilizzata per scoprire nuove relazioni tra le variabili che si hanno a disposizione. L'analisi confermativa si propone di stimare dei parametri, dando quindi risposte definitive alle ipotesi. Al contrario l'analisi esplorativa è orientata alla previsione e consente al ricercatore di formulare ipotesi, suscitando in esso domande rilevanti. Il primo è quindi un approccio theory-driven, il ricercatore si fa guidare dalla teoria e sulla

base di essa verifica delle ipotesi. Il secondo approccio invece è sicuramente data-driven, ossia il ricercatore si lascia condurre dai dati e da cosa essi propongono.

I due tipi di analisi sono applicati in molti ambiti. Ad esempio l'analisi confermativa è applicata alla Customer Satisfaction, alla Customer Loyalty e allo studio delle Emozioni. L'analisi esplorativa invece si applica in campi del tipo Brand Equity, Decision-Making Perspective e Tecnologia. Detto quindi che non sempre risulta facile distinguere tra i due tipi di analisi, diventa essenziale essere informati su ciò che esiste come base teorica di un fenomeno, così da poter meglio definire il tipo di ricerca che si sta svolgendo, per poter poi scegliere i metodi corretti. Infatti è importante sapere che la ricerca confermativa e quella esplorativa sono differenti e hanno le loro proprie assunzioni che necessitano di essere soddisfatte (Afthanorhan et al., 2020). Dopo aver definito il tipo di analisi che si vuole intraprendere, si passa alla rappresentazione del fenomeno di interesse tramite un modello. I Modelli a Equazioni Strutturali consentono appunto di stimare relazioni tra molte variabili, che siano esse osservabili o latenti. Il prossimo paragrafo fornisce una breve introduzione ai Modelli a Equazioni Strutturali o semplicemente SEM (Structural Equation Model).

1.2 Modelli a equazioni strutturali

I ricercatori hanno sempre a che fare con variabili concettuali e modelli teorici. Si pensi ad esempio al Benessere di una popolazione. Il Benessere è un concetto astratto, non direttamente misurabile. Si può però cercare di individuare il modo con cui il Benessere si manifesta. Ad esempio si potrebbe chiedere a una persona se è felice del proprio lavoro o del proprio percorso di studi oppure se è soddisfatta di come utilizza il proprio tempo libero o ancora se il proprio reddito è soddisfacente. Ovviamente tutti questi sono semplicemente degli esempi e possono essere tutti dei fattori di benessere, ma fanno anche capire l'importanza di dare una definizione chiara al concetto astratto che si vuole misurare e al modello teorico che ne consegue. Un modello teorico riflette quindi un insieme di relazioni che formalizzano la teoria e contiene le cosiddette variabili concettuali, ossia dei concetti astratti che i ricercatori cercano di misurare (Hair et al., 2016). Nel seguito le variabili concettuali verranno anche chiamate costrutti. La definizione del costrutto può cambiare sia qualitativamente che quantitativamente in base alla definizione dello studio. Si pensi ad esempio alla soddisfazione di un cliente per un certo servizio. Questa soddisfazione può essere suddivisa a sua volta in vari aspetti di interesse, quali ad esempio la velocità del servizio e le capacità dello staff che ha svolto tale servizio (Hair et al., 2020). La formulazione del costrutto definisce quindi

come una variabile concettuale, in genere latente, ossia non osservabile, si relaziona con le variabili osservate. L'insieme delle variabili osservate e la relazione che esse hanno con le variabili latenti è il cosiddetto modello di misura. Le relazioni che invece possono esistere tra i diversi costrutti fanno parte del cosiddetto modello strutturale. Il modello di misura e il modello strutturale formano assieme il Modello a Equazioni Strutturali, o semplicemente SEM. I modelli a equazioni strutturali consentono di stimare relazioni tra molte variabili, le quali possono essere dipendenti o indipendenti, osservate o latenti. Come detto in precedenza, i SEM sono utilizzati quando si vuole rappresentare un concetto astratto attraverso una variabile concettuale, in genere non osservabile. Nella Figura 1.1 si tratta delle variabili Y , indicate con i cerchi. Queste variabili latenti possono essere misurate solo attraverso delle variabili osservabili, dette "indicatori", che nella Figura 1.1 sono rappresentate dalle x , contenute nei rettangoli. Quindi si vuole quantificare qualcosa di non osservabile, utilizzando ciò che si osserva.

Esistono di conseguenza due parti del modello: la parte di misura, che riguarda il modo con cui i costrutti latenti sono appunto misurati dagli indicatori, e la parte strutturale, che riguarda le relazioni tra i costrutti latenti. Per la parte di misura, l'idea è che gli indicatori possano essere causati dal costrutto latente, come nel caso di x_7 , x_8 e x_9 in Figura 1.1, con la freccia che parte da Y_3 e va agli indicatori, oppure che causino il costrutto latente, come ad esempio x_1 , x_2 e x_3 , che sono origine delle frecce dirette verso Y_1 . Nel primo caso, che prevede anche un errore di misura per gli indicatori, si definisce Y_3 un costrutto "riflessivo". Nel secondo caso per Y_1 si parla di costrutto "formativo". La parte strutturale invece ipotizza delle relazioni tra le variabili latenti. In particolare se una variabile latente è indipendente dagli altri costrutti, allora si definisce "esogena". Essa è solo causa di altre variabili latenti, non è quindi causata da esse. Nella Figura 1.1 si tratta di Y_1 e Y_2 . Al contrario, una variabile latente come Y_3 o Y_4 , dipendente da altri costrutti, è definita "endogena". Riprendendo l'esempio sulla soddisfazione per il servizio, di cui al paragrafo precedente, il modello si dice riflessivo se le variabili osservate, soddisfazione per la velocità e soddisfazione per lo staff, sono manifestazioni della latente, ossia il nesso di causalità va dal concetto astratto alle variabili osservate. Se invece si ipotizza che siano le variabili osservate a formare il concetto latente, ossia il nesso di causalità ha senso contrario rispetto al modello riflessivo, allora il modello si dice formativo. Nel prossimo paragrafo si spiega la differenza tra i due tipi di modelli di misura di cui appena parlato, il modello riflessivo e il modello formativo.

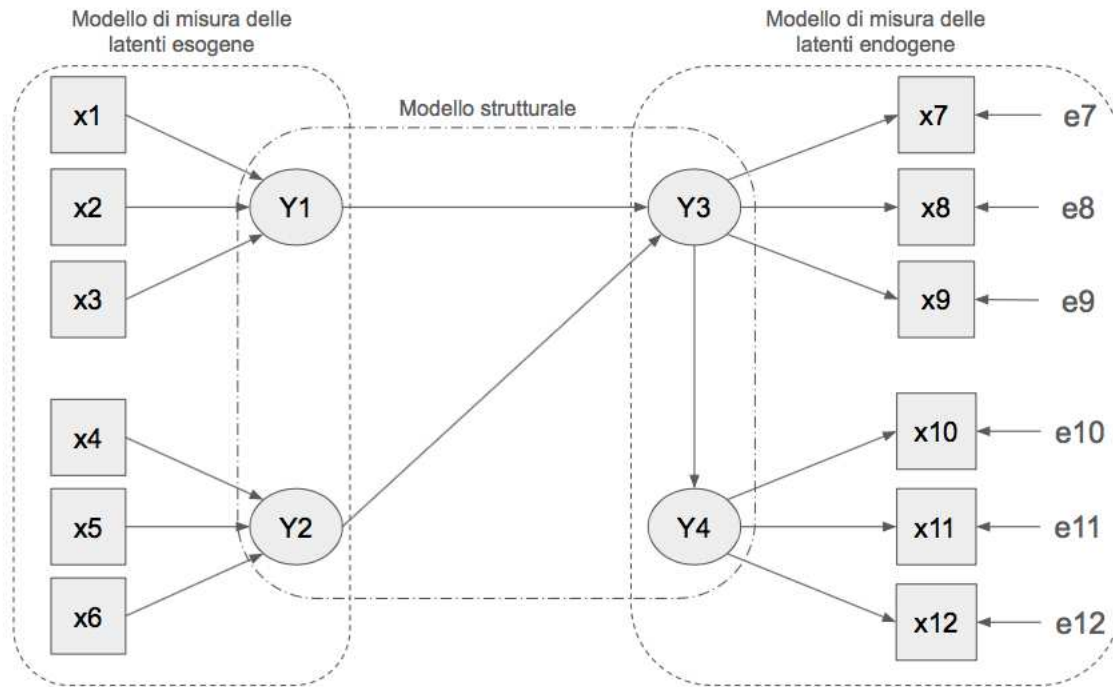


FIGURA 1.1: Modello a equazioni strutturali

1.2.1 Modello di Misura Riflessivo

Il primo modello di misura di cui si parla è quello riflessivo. Come vediamo in Figura 1.2, il nesso di causalità va dal costrutto Y alle variabili osservate x1, x2, x3 e x4, che nel seguito chiameremo semplicemente indicatori.

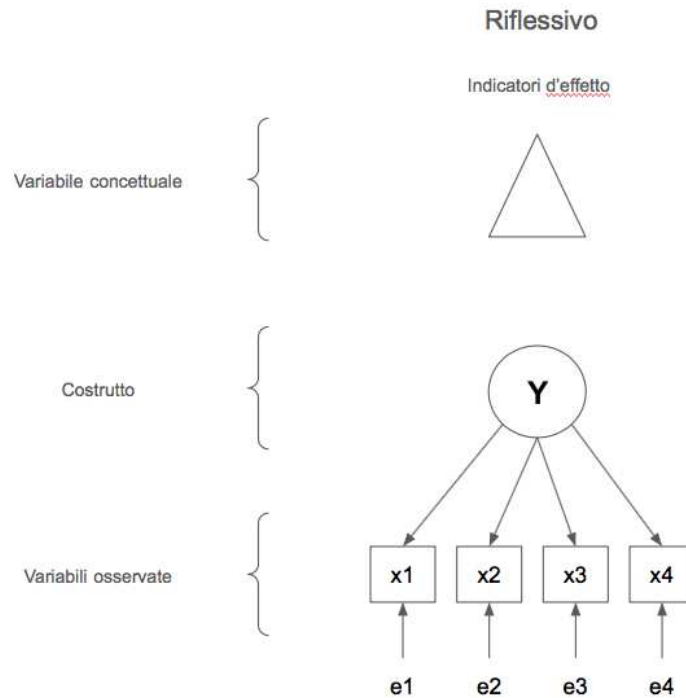


FIGURA 1.2: Modello riflessivo

Essi sono quindi considerati manifestazioni di un costrutto latente e le relative equazioni di misura possono essere rappresentate nel seguente modo

$$x_i = l_i \cdot Y + e_i \quad (1.1)$$

dove l_i , con $i = 1, \dots, k$, sono i coefficienti di regressione che quantificano la relazione tra i k indicatori e il costrutto Y ed e_i è l'errore di misura associato all' i -esimo indicatore. Si presuppone quindi che indicatori associati a uno stesso costrutto abbiano correlazione alta. Al contrario, indicatori associati a costrutti diversi dovrebbero avere correlazione più bassa. Questo vuol dire che, se un costrutto è affidabile, allora ciascuno dei suoi indicatori potrebbe ipoteticamente essere tralasciato, senza cambiare il significato del costrutto stesso (Hair et al., 2016). Si ipotizza quindi che l'eventuale correlazione tra indicatori di costrutti diversi sia attribuibile esclusivamente alla relazione tra i costrutti stessi e non ad altri tipi di relazioni.

1.2.2 Modello di Misura Formativo

Nel modello formativo invece i costrutti sono pensati come combinazioni lineari dei propri indicatori e quindi possono essere tra loro poco correlati. Come mostra la figura 1.3

esistono due tipi di indicatori formativi: gli indicatori causali e gli indicatori compositi.

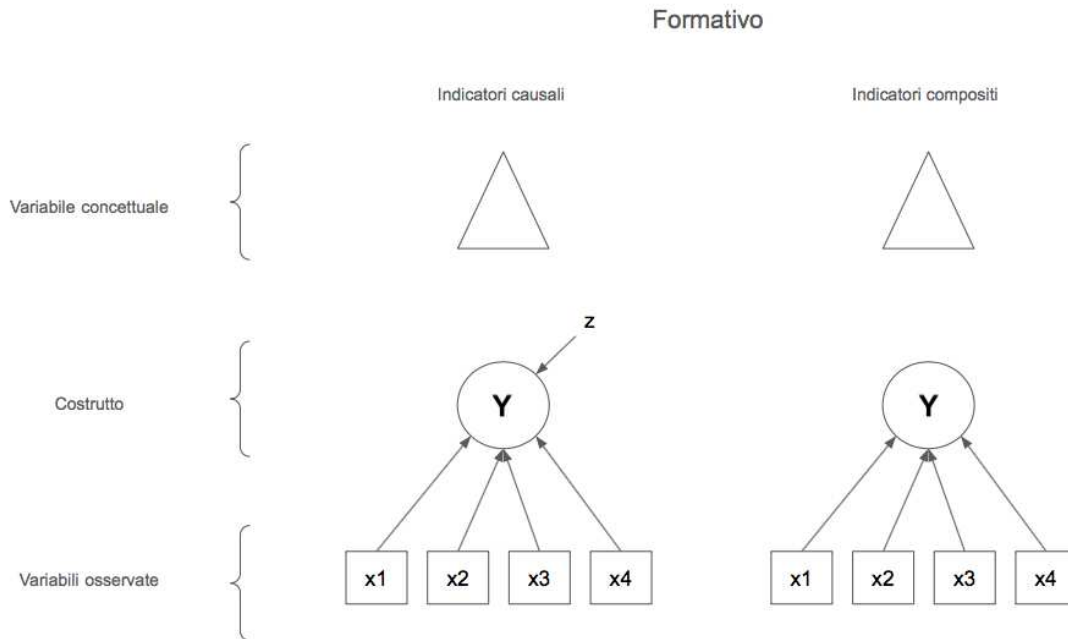


FIGURA 1.3: Modello formativo a indicatori causali e a indicatori compositi.

Nel primo caso si suppone che gli indicatori non possano spiegare in modo esaustivo il costrutto. Esiste quindi una qualche forma di errore, chiamato z nella Formula 1.2, che incide sul costrutto stesso e che non dipende dagli indicatori.

$$Y = \sum_{i=1}^k w_i \cdot x_i + z \quad (1.2)$$

dove w_i con $i = 1, \dots, k$ sono i coefficienti di regressione che quantificano la relazione tra i k indicatori e il costrutto Y e z è l'errore di misura associato al costrutto stesso.

Nel secondo caso invece si ipotizza che gli indicatori spieghino in modo esaustivo il costrutto, escludendo quindi che quest'ultimo possa avere anche un errore (Afthanorhan et al., 2020). Come mostra la Formula 1.3, la componente di errore z non è presente nel costrutto Y .

$$Y = \sum_{i=1}^k w_i \cdot x_i \quad (1.3)$$

dove w_i con $i = 1, \dots, k$ sono i coefficienti di regressione che quantificano la relazione tra i k indicatori e il costrutto Y .

Tra i due metodi è evidente come il primo sia il più realistico, in quanto è impensabile che il ricercatore riesca a cogliere tutti gli aspetti che influenzano un costrutto. Entrambi i metodi, inoltre, presumono che gli indicatori stessi non abbiano a loro volta un errore di misura, il che molto spesso risulta essere anche in questo caso poco realistico.

1.2.3 Scegliere tra i due modelli

La scelta tra i due modelli non è assolutamente facile e dipende da come il fenomeno viene concettualizzato. Ritorna quindi in campo quanto detto precedentemente sull'importanza di conoscere bene la teoria alla base di un fenomeno per poter scegliere il metodo più adatto per la ricerca. Oltre a ciò, è fondamentale avere le idee chiare su come formalizzare la domanda, pensando ad esempio se ciò che noi osserviamo, cioè gli indicatori, causi o sia causato da ciò che non osserviamo ma che vogliamo misurare, ovvero i costrutti.

I modelli riflessivi e formativi sono applicati a moltissime discipline, tra le quali marketing e psicologia. In quest'ultimo ambito, in particolare, molto spesso si abusa del modello riflessivo, anche quando il vero modello è composito e questo porta a errori estremi (Rhemtulla et al., 2020). Questo vuol dire che per ipotizzare un modello riflessivo è necessario avere delle giustificazioni fondate. Come dicono Rhemtulla et al. (2020) "in alcuni contesti, l'assunzione che tutti gli items misurino esattamente la stessa cosa, è plausibile". Nel modello riflessivo gli items sono quindi interscambiabili: tenere uno o l'altro non fa quindi differenza. Tuttavia quando si misurano costrutti psicologici "quali l'intelligenza, la depressione, l'estroversione, la prontezza scolastica, o la creatività, il set di variabili osservate è spesso non visto come manifestazione interscambiabile del costrutto".

1.2.4 Modello Strutturale

Nei Modelli a Equazioni Strutturali, oltre alla parte di misura, esiste anche una parte strutturale che consiste nella presenza di relazioni tra costrutti diversi. Si ipotizzi ad esempio una relazione strutturale tra il costrutto Y e il costrutto Z, come mostrato in Figura 1.4.



FIGURA 1.4: Esempio di relazione strutturale tra due costrutti Y e Z

Il costrutto latente Y è esogeno in quanto causa l'altro costrutto latente Z che quindi risulta essere endogeno. L'equazione di riferimento è illustrata nella Formula 1.4

$$Z = \beta \cdot Y + e_Z \quad (1.4)$$

dove β è il coefficiente della relazione causale da Y a Z ed e_Z è l'errore associato. Ovviamente a loro volta Y e Z possono essere costrutti riflessivi o formativi. Nella sua forma più generale il modello prevede relazioni causali sia fra Z e Y che fra le endogene Z :

$$Z = B \cdot Z + \Gamma \cdot Y + \zeta. \quad (1.5)$$

1.3 Metodi di stima dei Modelli a Equazioni Strutturali

Per stimare i Modelli a Equazioni Strutturali si hanno a disposizione svariati metodi di stima, tra cui i metodi Covariance-Based e il Partial Least Squares. Entrambi i metodi possono in principio stimare sia modelli formativi che riflessivi, con conseguenti vantaggi e svantaggi. In questo paragrafo si vuole illustrare brevemente i due metodi e capire quando i loro risultati possono essere confrontati.

1.3.1 CB-SEM

Il metodo Covariance-Based, nel seguito chiamato semplicemente CB-SEM o CB, si basa appunto sulla matrice di covarianza. L'obiettivo è quello di trovare il modello la cui matrice di covarianza stimata si avvicini maggiormente alla matrice di covarianza osservata nei dati.

Prendendo come riferimento un generico modello di misura con dei costrutti riflessivi endogeni z ed esogeni y , la generica matrice di covarianza è

$$\Sigma = \begin{bmatrix} \Sigma_{yy} & \\ \Sigma_{zy} & \Sigma_{zz} \end{bmatrix} \quad (1.6)$$

dove Σ_{yy} è la matrice di covarianza degli indicatori delle variabili latenti esogene, Σ_{zy} la matrice di covarianza tra indicatori endogeni ed esogeni e Σ_{zz} la matrice di covarianza di quelli endogeni.

Il modello di misura per le variabili esogene si scrive

$$x_y = \Lambda_{Y \rightarrow x} \cdot Y + \delta \quad (1.7)$$

dove $\Lambda_{Y \rightarrow x}$ la matrice dei coefficienti del modello di misura per le variabili esogene e δ sono gli errori di misura. Da ciò si ottiene

$$\Sigma_{x_y x_y} = \Lambda_{Y \rightarrow x} \Phi \Lambda_{Y \rightarrow x}^T + \Theta_\delta \quad (1.8)$$

con $\Phi = E(YY^T)$ e $\Theta_\delta = E(\delta\delta^T)$.

Per le variabili endogene, il modello di misura è

$$x_z = \Lambda_{Z \rightarrow x} \cdot Z + \epsilon \quad (1.9)$$

dove $\Lambda_{Z \rightarrow x}$ la matrice dei coefficienti del modello di misura per le variabili endogene e ϵ sono gli errori di misura. Da ciò si ottiene

$$\Sigma_{x_z x_z} = \Lambda_{Z \rightarrow x} E(ZZ^T) \Lambda_{Z \rightarrow x}^T + \Theta_\epsilon \quad (1.10)$$

con $\Theta_\epsilon = E(\epsilon\epsilon^T)$. Sapendo che $Z = (I - B)^{-1}(\Gamma \cdot Y + \zeta)$, dove B è la matrice dei coefficienti strutturali, allora si ottiene che

$$\Sigma_{x_z x_z} = \Lambda_{Z \rightarrow x} [(I - B)^{-1}(\Gamma \cdot \Phi \cdot \Gamma^T + \Psi)(I - B)^{-T}] \Lambda_{Z \rightarrow x}^T + \Theta_\epsilon \quad (1.11)$$

.

Per quanto riguarda la matrice di covarianza tra endogene ed esogene, si ha

$$\Sigma_{x_y x_z} = \Lambda_{Y \rightarrow x} \Phi \Gamma^T (I - B)^{-T} \Lambda_{Z \rightarrow x}^T \quad (1.12)$$

Il metodo CB-SEM utilizza la matrice di covarianza osservata S , che è definita

$$S = \begin{bmatrix} S_{yy} & \\ S_{zy} & S_{zz} \end{bmatrix} \quad (1.13)$$

e risolve il sistema di equazioni $S = \Sigma$, se il modello è esattamente identificato. Se invece è sovraidentificato, la stima avviene con la Massima Verosimiglianza sotto assunzione di multinormalità. I parametri del modello saranno stimati in modo da massimizzare la probabilità che S derivi da Σ , o equivalentemente minimizzare un'opportuna distanza fra S e Σ . Lo stimatore di massima verosimiglianza è asintoticamente corretto e consistente. Inoltre, all'aumentare della numerosità campionaria la distribuzione dello stimatore di Massima Verosimiglianza approssima una distribuzione normale.

Questo metodo divide la varianza degli indicatori in varianza comune, ossia la varianza condivisa dagli indicatori nella parte di misura, e varianza unica che include sia la varianza specifica sia la varianza dell'errore. Solo la varianza comune è di fatto utilizzata per la stima del modello e per questo motivo esso è più adatto alla stima di modelli riflessivi, in quanto essi prevedono che tutta la covarianza tra gli indicatori di un costrutto possa essere spiegata esclusivamente dal costrutto stesso e per questo sia sufficiente per la stima dei parametri. L'ipotesi quindi è che la varianza di un insieme di indicatori sia perfettamente spiegata dall'esistenza di un'unica variabile latente (Hair et al., 2016). Questo metodo infatti ha come obiettivo quello di stimare la matrice di covarianza degli indicatori avvicinandosi più possibile alla covarianza osservata nei dati. Con questo metodo si potrebbero anche stimare modelli formativi, ma come dicono Bollen et al. (2009) e Diamantopoulos et al. (2011), la stima può avvenire solo dopo l'imposizione di vincoli che consentano l'identificazione del modello. Come dicono infine Hair et al. (2012), questi vincoli spesso contraddicono le considerazioni teoriche e sorge la domanda se sia il modello a guidare la teoria o viceversa.

1.3.2 PLS-SEM

Al contrario del CB-SEM, il metodo PLS utilizza tutta la varianza per la stima dei parametri. L'idea è che il costrutto sia una combinazione lineare degli indicatori e che sia quindi più adatto a modelli formativi. Tuttavia, questo metodo consente anche la stima di modelli riflessivi fornendo però delle stime che potrebbero essere distorte. A correzione di questa distorsione è stato introdotto il Consistent Partial Least Squares

SEM (PLSc-SEM), il quale appone una correzione alle stime al fine di renderle non distorte. Venturini et al. (2019) illustra l'algoritmo di stima di questo metodo.

L'algoritmo prevede un primo passo in cui vengono inizializzati i punteggi delle variabili latenti come somma pesata degli indicatori associati. Il generico costrutto Y è quindi definito nella seguente Formula 1.14:

$$\hat{Y} = \sum_{i=1}^k w_i \cdot x_i \quad (1.14)$$

dove w_i con $i = 1, \dots, k$ sono i coefficienti di regressione, chiamati anche *outer weights*, che quantificano la relazione tra i k indicatori e il costrutto Y . Alla prima iterazione $w_i = 1$ per ogni $i = 1, \dots, k$.

Ora che si ha a disposizione una prima stima dei punteggi dei costrutti latenti si passa al calcolo degli *inner weights*, ossia i coefficienti del modello strutturale. In particolare ogni variabile latente è ora ristimata come somma pesata di tutti gli altri costrutti latenti. Lo schema più utilizzato per la stima degli *inner weights* è il *path scheme*. Esso prevede che il peso del generico costrutto Z nell'equazione del costrutto Y corrisponda:

- al coefficiente di correlazione se Y causa Z ;
- al coefficiente di regressione se Y è causata da Z ;
- a 0 se non esiste relazione causale tra Y e Z .

A questo punto si ristima il generico costrutto latente Y nel seguente modo

$$\hat{Y} = \sum_{j=1}^p \beta_j \cdot Z_j \quad (1.15)$$

dove i β_j , per $j = 1, \dots, p$ con p il numero di costrutti presenti nel modello oltre a Y , sono i coefficienti del modello strutturale, chiamati anche *inner weights*, che quantificano la relazione tra il costrutto Y e tutti gli altri costrutti presenti nel modello. Ovviamente la Formula 1.15 vale in modo analogo per tutti gli altri costrutti.

Il terzo passo del metodo PLS consiste nella stima degli *outer weights*. Esistono due metodi per la stima degli *outer weights*. Il primo metodo, spesso chiamato anche *Metodo A*, prevede la regressione di ciascun indicatore sul proprio costrutto di riferimento. Se il costrutto è standardizzato, allora il peso corrisponde al coefficiente di correlazione lineare, come esplicita la Formula 1.16.

$$w_i = \text{cor}(Y, x_i) \quad (1.16)$$

con $i = 1, \dots, k$.

Il secondo metodo, chiamato anche *Metodo B*, prevede che gli *outer weights* siano stimati regredendo il costrutto Y sui suoi indicatori. Il j -esimo peso sarà quindi

$$w_i = \frac{\text{cor}(Y, x_i)}{\text{var}(x_i)} \quad (1.17)$$

Si può intuire che il Metodo A, utilizzando pesi di correlazione o regressione degli indicatori sui costrutti, sia indicato per modelli riflessivi, mentre il Metodo B, che regredisce i costrutti sugli indicatori, sia adatto a modelli formativi.

Il quarto passo dell'algoritmo consiste nella stima dei punteggi dei costrutti latenti come somma pesata dei propri indicatori. Il procedimento è analogo alla Formula 1.14, ma stavolta i pesi w_i con $i = 1, \dots, k$ sono quelli calcolati al passo 3.

Questa procedura viene ripetuta fino a convergenza. Quando l'algoritmo raggiunge la convergenza e si ha quindi la stima finale dei punteggi dei costrutti latenti, si procede con la stima finale dei parametri del modello di misura. L'ultimo passo sarà la stima dei parametri strutturali.

1.3.3 Vantaggi e svantaggi dei due metodi

Vista la natura dei due metodi, si può dire che il metodo CB sia adatto a modelli riflessivi e spesso è associato a un'analisi di tipo confermativo. Il PLS invece è adatto a modelli formativi, spesso associati ad analisi esplorative e quindi è orientato a un'ottica previsiva. Quest'ultimo risulta utile quando non esiste molta teoria alla base del fenomeno di interesse oppure se il modello è mispecificato, ossia quando i ricercatori non sono sicuri degli effetti causali tra esogene ed endogene (Afthanorhan et al., 2020). Entrambi i metodi sono quindi indicati per determinati modelli e specifici tipi di analisi. Tuttavia spesso sono utilizzati anche per modelli non esattamente indicati in quanto nella realtà non è sempre facile avere a disposizione dati che diano indicazioni chiare sul tipo di modello. Ovviamente i risultati ottenuti saranno delle approssimazioni. Nella Tabella 1.1 si elencano alcuni vantaggi e svantaggi dei due metodi.

TABELLA 1.1: Vantaggi e svantaggi dei metodi PLS-SEM e CB-SEM.

	PLS	CB
Distribuzione dei dati	Nessun vincolo	Distribuzione Normale
Dimensione del campione	Nessun vincolo	Dimensioni abbastanza elevate
Indici di bontà	Non presenti	Presenti
Identificazione	Nessun vincolo	Necessaria l'identificazione del modello
Inferenza	Solo tramite simulazione	Possibile fare inferenza

Come mostra la Tabella 1.1, il PLS-SEM è un modello non parametrico, in quanto non prevede che i dati abbiano una specifica distribuzione. Al contrario, il CB-SEM è un modello parametrico; esso infatti assume che i dati abbiano una distribuzione normale. Tuttavia è ampiamente dimostrato da studi di simulazione che il CBSEM è robusto anche in assenza di normalità (Bollen, 1989). È inoltre dimostrato che il PLS si comporti meglio del CBSEM in presenza di campioni piccoli e non soffra di problemi di identificazione, al contrario del CBSEM che invece la richiede. Tuttavia PLS non consente di fare inferenza, rendendo così necessario ricorrere a delle simulazioni. Inoltre non presenta indici di bontà, al contrario del CB-SEM. Si nota come non sia facile e forse nemmeno corretto dire che un metodo sia migliore dell'altro. È quindi buona norma capire bene quali sono le esigenze del ricercatore e di conseguenza scegliere il metodo più adatto alle proprie necessità, seguendo le linee guida date dalla teoria proposta finora a riguardo. Come dicono Hair et al. (2012), "le debolezze del PLS-SEM sono i punti di forza del CB-SEM e viceversa". Nella prossima sezione si cerca di dare indicazione sui modelli per i quali i risultati ottenuti con i due metodi siano confrontabili .

1.3.4 Confronto e campi di applicazione

Come detto in precedenza è utile cercare di capire se e in quali casi i due metodi CB e PLS siano confrontabili. Per quanto riguarda la stima dei modelli formativi, il PLS è sicuramente il metodo più indicato. In linea di principio anche il CB-SEM consente la stima di tali modelli, ma in questo caso si rende necessaria l'imposizione di vincoli che spesso contraddicono le considerazioni teoriche (Hair et al., 2012). Essendo quindi sconsigliato l'utilizzo dei CBSEM per modelli formativi, è da evitare anche il confronto tra i risultati dei due metodi. Ci si chiede quindi se sia possibile il confronto per modelli riflessivi. Se si ipotizza che il vero modello sia riflessivo, allora si può certamente dire che il CB-SEM è un metodo robusto. È però dimostrato che anche il PLS dà stime consistenti. Quindi forse il confronto tra i due metodi per questi tipi di modelli può avere un senso. Nel Capitolo 2 si procede con delle simulazioni da alcuni modelli riflessivi e

si analizzano le stime CB e PLS della parte di misura del modello cambiando di volta in volta la parte strutturale, sia in termini di intensità delle relazioni strutturali sia in termini di direzione di quest'ultime. L'obiettivo è quello di approfondire come e quanto il metodo PLS-SEM, che utilizza tutta la varianza presente nei dati, possa essere influenzato dal modello strutturale. Come detto in precedenza, nelle analisi del prossimo Capitolo sarà utilizzato il metodo PLSc, ossia il Consistent Partial Least Squares. Per semplicità sarà chiamato solo PLS.

Capitolo 2

Confronto tra CB-SEM e PLS-SEM per Modelli Riflessivi

2.1 Introduzione

In questo capitolo si vuole analizzare il comportamento dei due metodi Covariance-Based e Partial Least Squares nella stima di modelli riflessivi. In particolare è di interesse capire se la stima dei parametri di misura con i due metodi sia influenzata dal modello strutturale. Si analizza il comportamento delle stime dei parametri per relazioni strutturali con direzioni e intensità diverse.

Si sceglie quindi di simulare da quattro tipi di modelli riflessivi con diverse parti strutturali. Per ogni tipo di modello si fanno 5000 simulazioni facendo variare i parametri strutturali per capire se abbiano incidenza sulla stima dei parametri di misura ed eventualmente in che modo.

I modelli scelti per le simulazioni sono i seguenti:

- Modello di regressione semplice
- Modello di regressione multivariata
- Modello di regressione multipla
- Modello ricorsivo.

Per ogni modello si osservano i comportamenti dei due metodi al crescere del valore del parametro strutturale, al fine di avere un'indicazione su quando sia possibile confrontare i due metodi in quanto assumono comportamenti simili e quando invece sia meglio scegliere uno a scapito dell'altro. L'analisi si focalizza su costrutti misurati da

tre indicatori ciascuno, in quanto, in questo caso, il modello di misura di ogni costrutto è esattamente identificato. Alla fine del Capitolo si replica brevemente il modello di regressione semplice cambiando il numero di indicatori: si effettuano simulazioni da due modelli di regressione semplice, uno sovraidentificato con costrutti da quattro indicatori e uno con costrutti da due indicatori, per il quale l'identificazione necessita della correlazione fra diversi fattori.

2.2 Modello di regressione semplice

In questo paragrafo si fanno 5000 simulazioni da un modello di regressione semplice con due costrutti riflessivi. Ogni costrutto ha tre indicatori, in particolare:

- il costrutto X ha indicatori x_1 , x_2 e x_3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;
- il costrutto Y ha indicatori y_1 , y_2 e y_3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8.

La parte strutturale prevede una relazione causale da X a Y. Si vuole osservare come cambiano le stime della parte di misura al crescere del vero valore del parametro strutturale. Si sceglie quindi di simulare da sei modelli in cui cambia solo il valore del parametro strutturale da 0.2 a 0.7 come illustrato nella Tabella 2.1. Per semplicità, in questa e nelle seguenti Tabelle, la parola “Mod” indicherà il Modello.

TABELLA 2.1: Parametri strutturali nei modelli di regressione semplice simulati.

	Mod 1a	Mod 1b	Mod 1c	Mod 1d	Mod 1e	Mod 1f
$X \rightarrow Y$	0.2	0.3	0.4	0.5	0.6	0.7

La Figura 2.1 illustra il modello simulato. Il parametro strutturale non è inserito nel grafico, in quanto cambia nelle simulazioni. Il costrutto X ha tre indicatori i cui parametri di misura sono rispettivamente 0.6, 0.7 e 0.8. Lo stesso modello di misura vale per il costrutto Y e i suoi indicatori. Si vogliono analizzare anche eventuali comportamenti diversi per parametri di misura diversi. Ogni indicatore è soggetto a un errore di misura con distribuzione Normale. Anche il costrutto Y, essendo la variabile endogena della parte strutturale, è soggetto a errore.

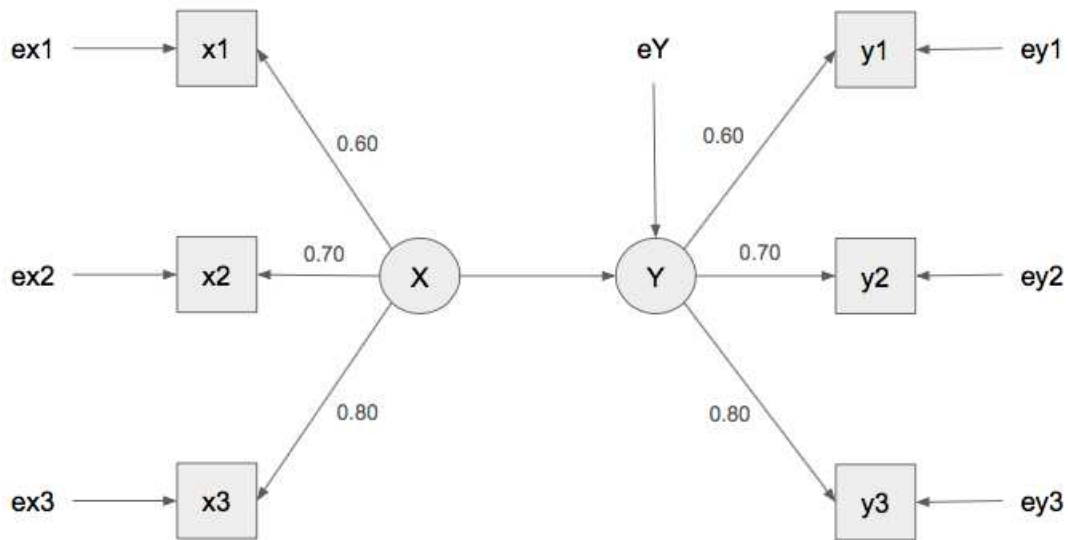


FIGURA 2.1: Modello di regressione semplice con due costrutti riflessivi X e Y, ognuno con tre indicatori.

Nel seguito si riportano le distribuzioni delle stime dei modelli CB e PLS applicati alle 5000 simulazioni. La Figura 2.2 mostra le densità delle stime ottenute applicando alle simulazioni il metodo CB (linea rossa) e il metodo PLS (linea verde) cambiando di volta in volta il vero valore del parametro strutturale. Considerando il primo grafico della Figura 2.2 si nota la sostanziale differenza tra le due curve. La curva rossa, riferita al metodo CB è centrata in 0.6, che si ricorda essere il vero valore del parametro $X \rightarrow x_1$, e ha delle code ridotte. Questo vuol dire che le stime ottenute con il metodo CB sono mediamente corrette e hanno una deviazione standard contenuta. Tutt'altro accade per le stime PLS le quali, pur essendo anch'esse centrate in 0.6, presentano una deviazione standard molto più alta. Passando invece all'ultimo grafico della Figura 2.2 si nota che le due curve sono molto più simili rispetto ai grafici precedenti. In generale si nota che per gli indicatori del costrutto X il metodo CB fornisce delle stime buone sia in termini di media che di varianza indipendentemente dal vero valore del parametro strutturale. Per quanto riguarda invece il metodo PLS, si può dire che la varianza delle stime sia molto influenzata dalla parte strutturale. In particolare, la varianza si riduce all'aumentare del vero valore del parametro strutturale. Nell'Appendice al paragrafo A.1 sono disponibili i grafici per $X \rightarrow x_2$ e $X \rightarrow x_3$, che presentano evidenze simili.

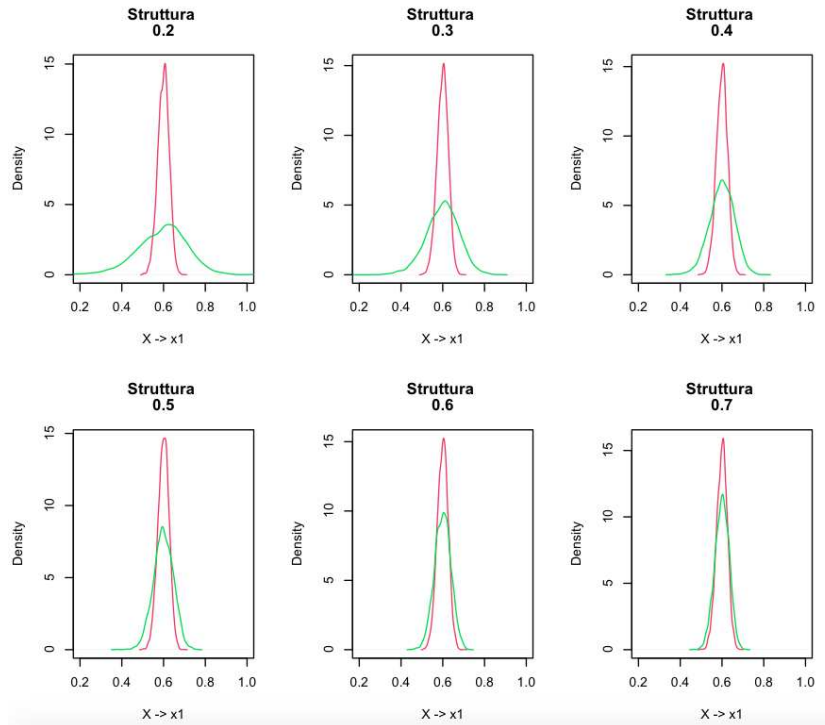


FIGURA 2.2: Distribuzione delle stime del parametro di misura $X \rightarrow x_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS.

La Figura 2.3 raffigura il comportamento delle medie e delle deviazioni standard delle stime di misura per gli indicatori di X al crescere del parametro strutturale $X \rightarrow Y$. Si nota infatti un comportamento analogo anche per le stime dei parametri associati a x_2 e x_3 . In termini di media, le stime sono molto simili tra i due metodi CB e PLS, anche se si nota una leggera differenza per valori bassi del parametro strutturale: il metodo PLS fornisce stime leggermente inferiori rispetto al metodo CB. Per quanto riguarda la deviazione standard delle stime, i due metodi ancora una volta forniscono risultati molto diversi, soprattutto per valori bassi del parametro strutturale. Il problema si attenua quando il vero valore del parametro strutturale aumenta.

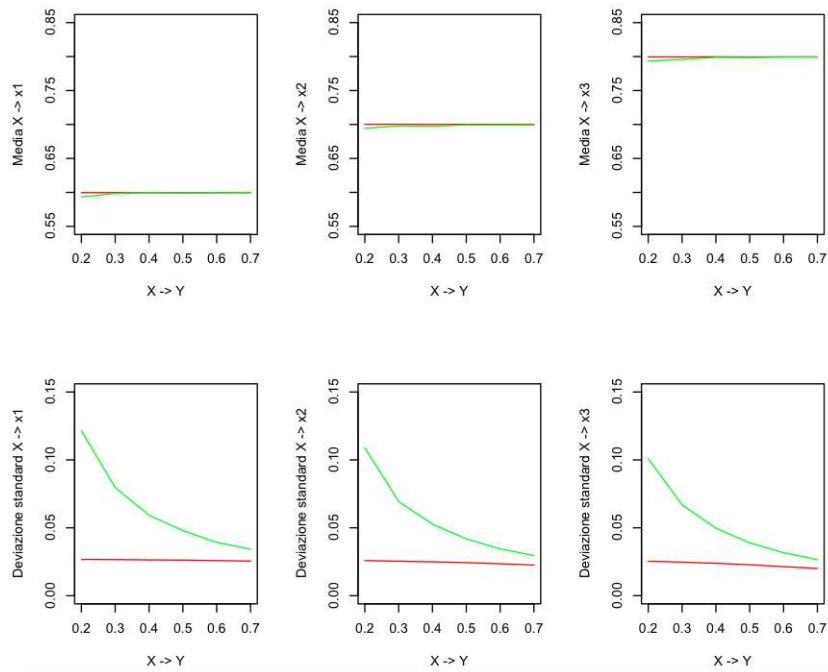


FIGURA 2.3: Media e deviazione standard delle stime dei parametri di misura del costrutto X con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

Gli stessi commenti valgono per le stime della parte di misura del costrutto Y, come si vede nella Figura 2.4. Nell'Appendice al paragrafo A.1 si trovano anche le densità delle stime di $Y \rightarrow y_1$, $Y \rightarrow y_2$ e $Y \rightarrow y_3$, per cui valgono commenti analoghi a quelli degli indicatori di X.

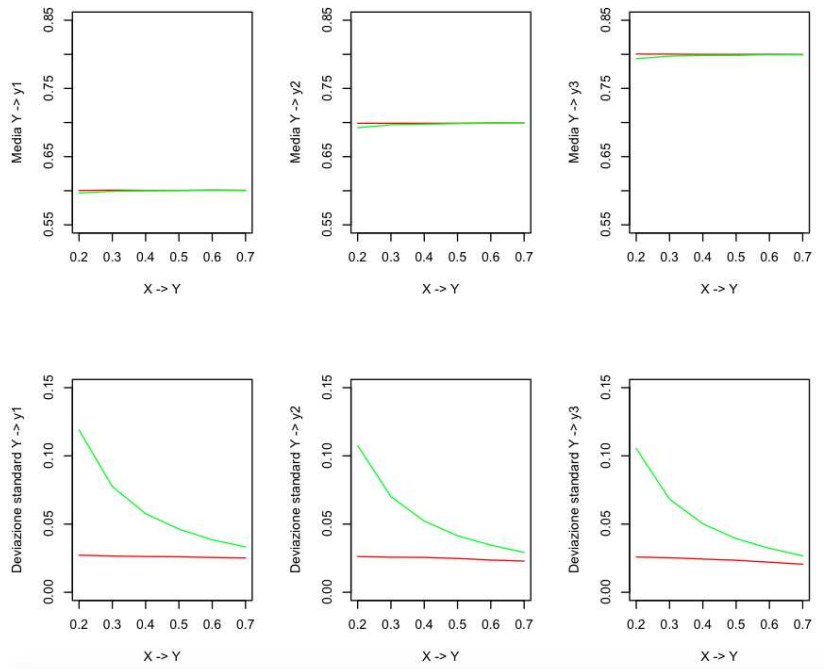


FIGURA 2.4: Media e deviazione standard delle stime dei parametri di misura del costrutto Y con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

Da questo primo modello si deduce che sia l'intensità della relazione strutturale più che la sua direzione ad influenzare le stime del metodo PLS. Infatti, si ricorda che X è una variabile esogena, mentre Y è una variabile endogena; tuttavia il comportamento del metodo PLS risulta uguale per i parametri di misura di entrambi i costrutti.

Per quanto riguarda la stima del parametro strutturale, i due metodi forniscono stime molto simili sia in termini di media che di varianza, come mostrano le Figure 2.5 e 2.6.

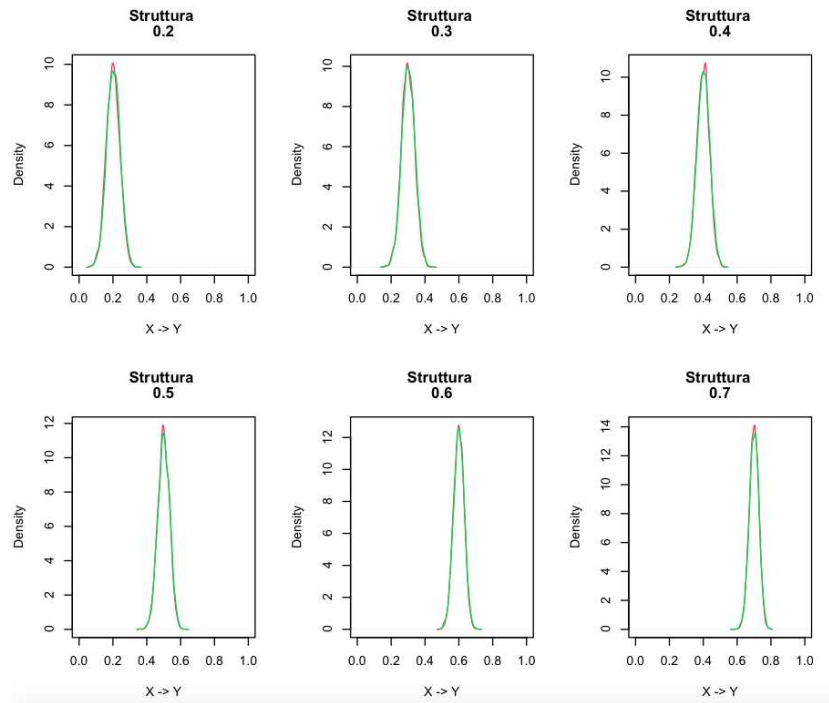


FIGURA 2.5: Distribuzione delle stime del parametro strutturale $X \rightarrow Y$. In rosso le stime con CB e in verde le stime PLS

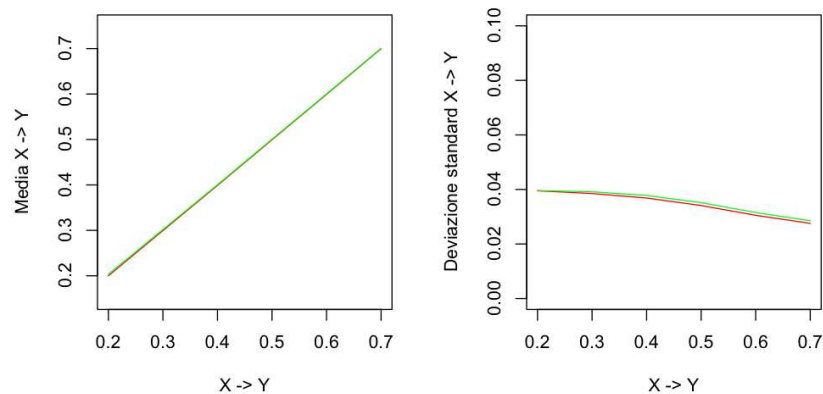


FIGURA 2.6: Media e deviazione standard delle stime del parametro strutturale $X \rightarrow Y$ con CB (linea rossa) e PLS (linea verde) al crescere del vero valore del parametro stesso.

Queste evidenze possono essere numericamente riassunte calcolando gli Errori Quadratici Medi (MSE). La Tabella 2.2 mostra che quelli ottenuti con il metodo CB, in generale, sono simili e non sono quindi molto influenzati dalla parte strutturale del modello. Al contrario, la Tabella 2.3 mostra che gli MSE per il metodo PLS diminuiscono notevolmente all'aumentare del vero valore del parametro strutturale. Inoltre, si nota che

l'MSE tende a diminuire anche per parametri di misura maggiori. Ad esempio, per il costruito X, l'MSE ottenuto nei vari modelli per $X \rightarrow x1$, il cui vero valore è pari a 0.6, supera l'MSE di $X \rightarrow x3$, che invece ha come vero valore 0.8. È infine evidente che il metodo PLS fornisca MSE nettamente superiori rispetto al CB. Si consideri, ad esempio, l'MSE di $X \rightarrow x1$ per il Modello 1a: il PLS fornisce un valore pari a 147.857, che risulta essere oltre 20 volte superiore a quello ottenuto con CB, pari a 7.145. La stessa riflessione vale per $Y \rightarrow y1$ e $Y \rightarrow y3$.

TABELLA 2.2: Errori quadratici medi delle stime CB dei parametri di misura del modello di regressione semplice. Ordine di grandezza 10^{-4} .

	Mod 1a	Mod 1b	Mod 1c	Mod 1d	Mod 1e	Mod 1f
MSE CB $X \rightarrow x1$	7.145	7.074	6.954	6.884	6.688	6.505
MSE CB $X \rightarrow x2$	6.698	6.442	6.227	5.936	5.529	5.111
MSE CB $X \rightarrow x3$	6.452	6.102	5.692	5.191	4.580	4.032
MSE CB $Y \rightarrow y1$	7.399	7.077	6.877	6.775	6.519	6.342
MSE CB $Y \rightarrow y2$	6.883	6.623	6.552	6.169	5.591	5.235
MSE CB $Y \rightarrow y3$	6.682	6.412	5.939	5.507	4.872	4.236
MSE CB $X \rightarrow Y$	15.615	14.829	13.593	11.643	9.317	7.609

TABELLA 2.3: Errori quadratici medi delle stime PLS dei parametri di misura del modello di regressione semplice. Ordine di grandezza 10^{-4} .

	Mod 1a	Mod 1b	Mod 1c	Mod 1d	Mod 1e	Mod 1f
MSE PLS $X \rightarrow x1$	147.857	63.396	35.141	22.911	15.468	11.712
MSE PLS $X \rightarrow x2$	118.927	48.271	27.795	17.527	11.873	8.744
MSE PLS $X \rightarrow x3$	102.217	44.940	24.760	15.207	10.041	7.027
MSE PLS $Y \rightarrow y1$	141.711	60.006	33.208	21.447	14.775	11.086
MSE PLS $Y \rightarrow y2$	116.163	49.307	27.311	17.215	11.947	8.556
MSE PLS $Y \rightarrow y3$	111.494	46.692	25.235	15.483	10.368	7.125
MSE PLS $X \rightarrow Y$	15.819	15.360	14.283	12.367	9.952	8.142

Da questo primo modello si nota quindi che il metodo PLS è fortemente influenzato dalla parte strutturale. Nello specifico i due metodi hanno comportamenti diversi per valori bassi del parametro strutturale, mentre si assomigliano per valori alti. In un modello riflessivo la parte di misura dovrebbe essere indipendente dalla parte strutturale. Si ipotizza quindi che i due metodi siano confrontabili solo nel caso in cui il parametro strutturale assuma un valore alto.

2.3 Modello di regressione multivariata

In questo paragrafo si procede con delle simulazioni da un modello di regressione multivariata a tre costrutti riflessivi. Ogni costrutto avrà a sua volta tre indicatori, in particolare:

- il costrutto X ha indicatori x1, x2 e x3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;
- il costrutto Y ha indicatori y1, y2 e y3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;
- il costrutto Z ha indicatori z1, z2 e z3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8.

La parte strutturale prevede una relazione causale da X a Y e da X a Z, come si vede in Figura 2.7.

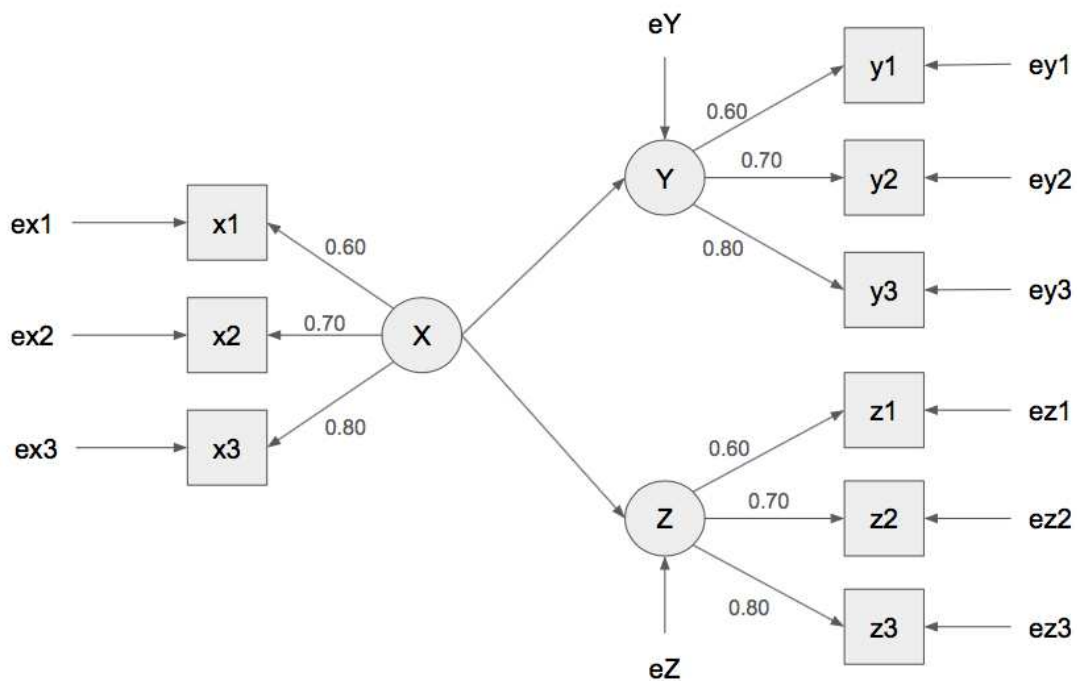


FIGURA 2.7: Modello di regressione multivariata a tre costrutti riflessivi X, Y e Z, ognuno con tre indicatori, con relazione causale da X a Y e da X a Z.

I parametri strutturali assumeranno i valori 0.2 e 0.7. Si avranno quindi in tutto quattro modelli, come mostrato nella Tabella 2.4. Il Modello 2a prevede delle relazioni strutturali

di lieve intensità. I Modelli 2b e 2c contengono una relazione strutturale forte e una debole, quindi si prevede diano risultati simili rispettivamente per i costrutti Y e Z. Nel Modello 2d invece entrambe le relazioni strutturali sono forti.

TABELLA 2.4: Parametri strutturali nei modelli di regressione multivariata simulati.

	Mod 2a	Mod 2b	Mod 2c	Mod 2d
$X \rightarrow Y$	0.2	0.2	0.7	0.7
$X \rightarrow Z$	0.2	0.7	0.2	0.7

Per ogni modello si fanno 5000 simulazioni. Tra i quattro modelli di simulazione cambia solo il valore dei parametri strutturali, non la direzione delle relazioni causali. Come si nota nella Figura 2.8 (e analogamente nell'Appendice al paragrafo A.2 per $X \rightarrow x_2$ e $X \rightarrow x_3$), il metodo CB fornisce stime molto buone sia in termini di media che di varianza. Il metodo PLS, invece, continua a restituire stime con bassa varianza solo per valori alti dei parametri strutturali. In particolare, per il modello di misura per X, le stime hanno una varianza elevata quando sia $X \rightarrow Y$ che $X \rightarrow Z$ assumono valore 0.2. La varianza si riduce molto quando entra in gioco anche solo una relazione strutturale forte, che sia $X \rightarrow Y$ oppure $X \rightarrow Z$.

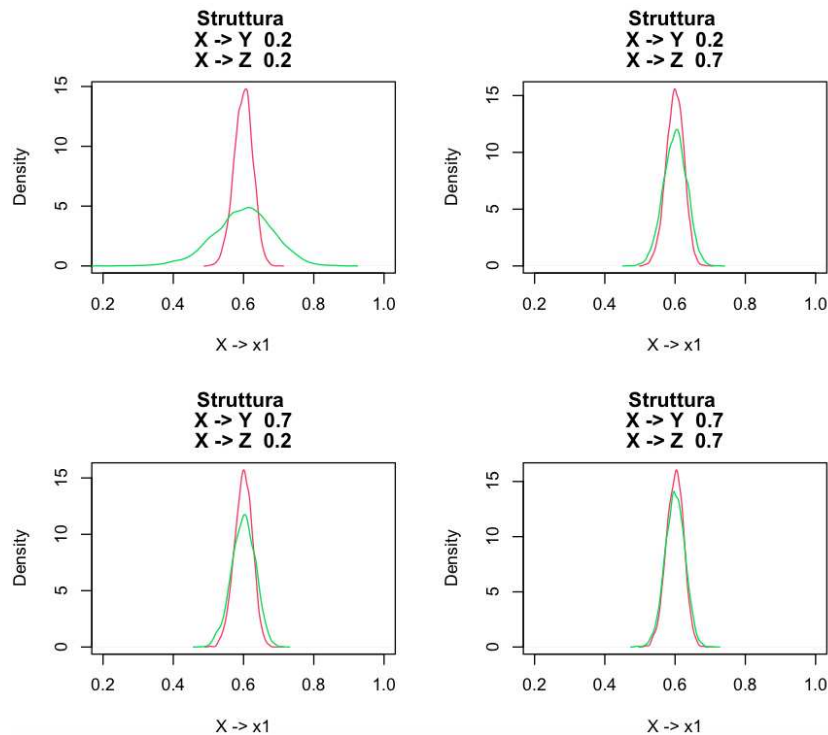


FIGURA 2.8: Distribuzione delle stime del parametro di misura $X \rightarrow x_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS.

Per quanto riguarda gli indicatori y_1 , y_2 e y_3 , il metodo CB fornisce stime con varianza bassa per entrambi i valori del parametro strutturale, mentre con il PLS le stime hanno deviazione standard alta quando il parametro strutturale $X \rightarrow Y$ vale 0.2. Quando invece vale 0.7, il metodo funziona bene, indipendentemente dal valore dell'altra relazione causale, ossia $X \rightarrow Z$ (vedi Figura 2.9). Nell'Appendice al paragrafo A.2 si trovano anche le densità delle stime di $Y \rightarrow y_2$ e $Y \rightarrow y_3$, per cui valgono commenti analoghi. Si presume quindi che lo stesso ragionamento valga anche per gli indicatori di Z , in particolare che le stime dei modelli con un valore basso per il parametro strutturale $X \rightarrow Z$, assumano un comportamento analogo.

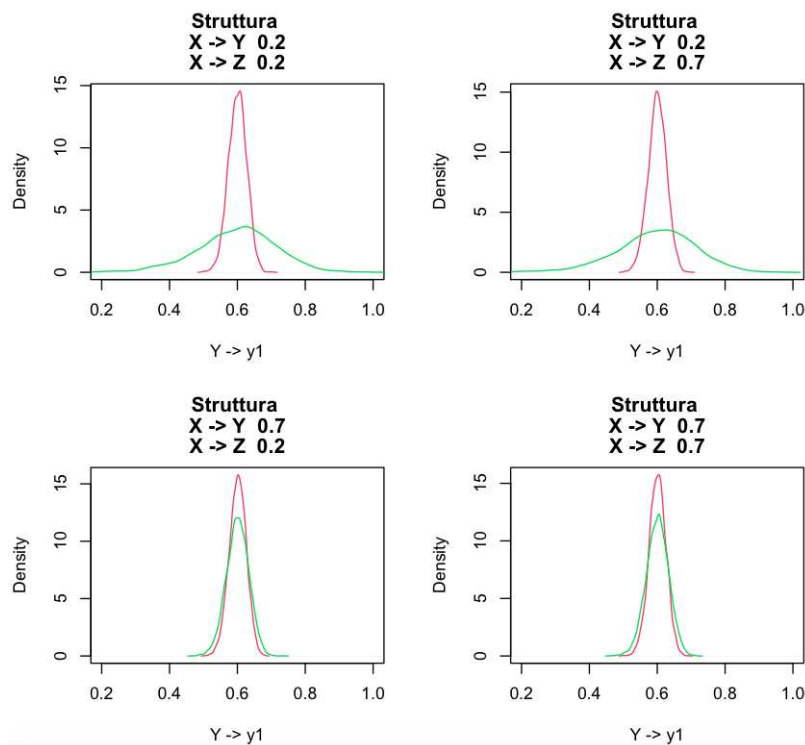


FIGURA 2.9: Distribuzione delle stime del parametro di misura $Y \rightarrow y_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS.

Effettivamente si verifica quanto appena detto. Per quanto riguarda gli indicatori z_1 , z_2 e z_3 , il metodo CB funziona bene per tutti i modelli, mentre il PLS ha problemi quando il parametro strutturale $X \rightarrow Z$ vale 0.2. Quando invece vale 0.7, il metodo funziona bene, indipendentemente dal valore di $X \rightarrow Y$ come mostra la Figura 2.10. Nell'Appendice paragrafo A.2 si trovano anche le distribuzioni delle stime di $Z \rightarrow z_2$ e $Z \rightarrow z_3$, per cui valgono commenti analoghi.

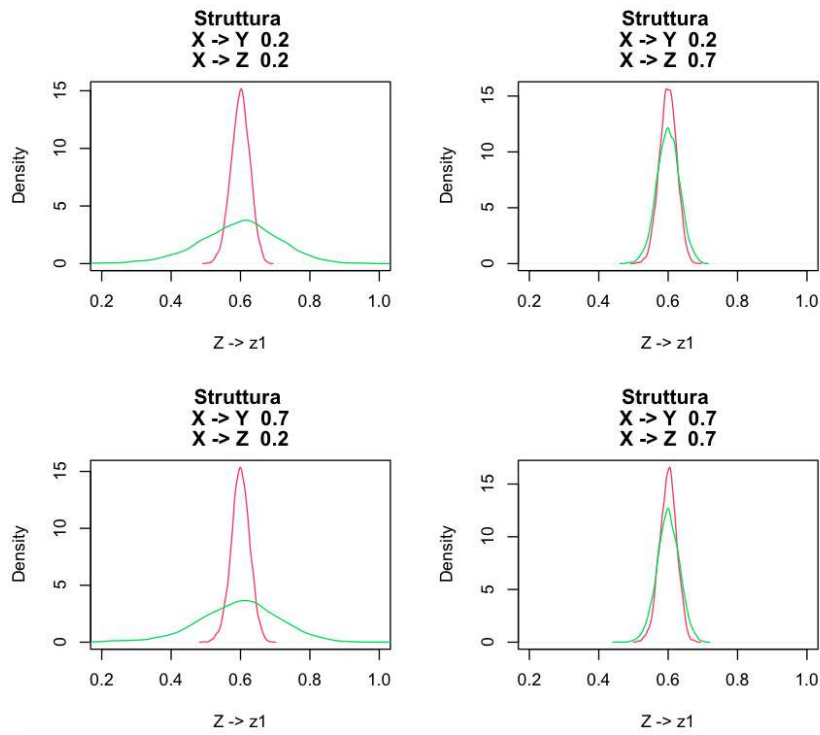


FIGURA 2.10: Distribuzione delle stime del parametro di misura $Z \rightarrow z_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Per la parte di struttura, come per il Modello di regressione semplice, i due metodi danno risultati simili, come si vede nelle Figure 2.11 e 2.12.

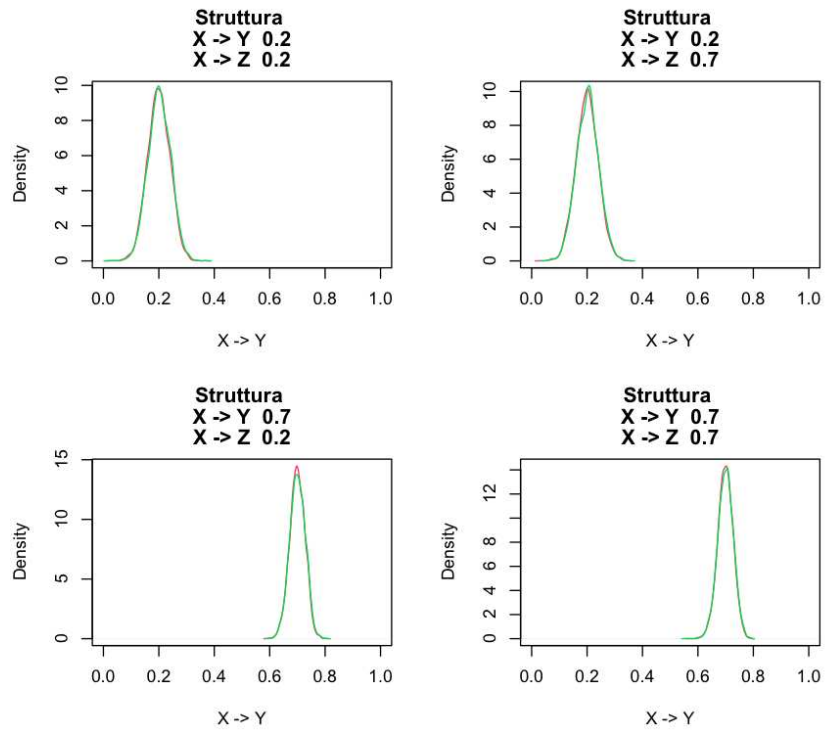


FIGURA 2.11: Distribuzione delle stime del parametro strutturale $X \rightarrow Y$. In rosso le stime con CB e in verde le stime PLS

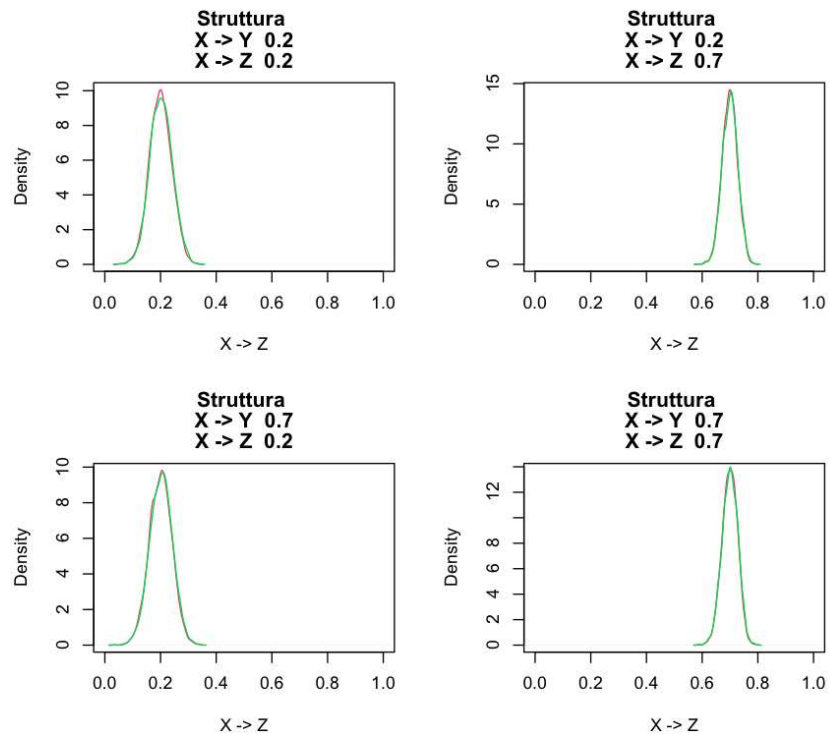


FIGURA 2.12: Distribuzione delle stime del parametro strutturale $X \rightarrow Z$. In rosso le stime con CB e in verde le stime PLS

I risultati ottenuti dal Modello di regressione semplice e dal Modello di regressione multivariata fanno intuire che il metodo PLS sia influenzato dalla parte strutturale. Questo non dovrebbe verificarsi nel caso di modelli riflessivi, per quanto spiegato nei capitoli precedenti. Tuttavia la presenza delle relazioni strutturali risulta influente in termini di intensità. Nel prossimo paragrafo si simula un Modello di regressione multipla, mantenendo gli stessi costrutti e indicatori, invertendo però la direzione delle relazioni causali. In questo modo si vuole capire se anche la direzione dei nessi causali influisca sulle stime.

2.4 Modello di regressione multipla

Si procede ora con delle simulazioni da un modello di regressione multipla a tre costrutti riflessivi. Ogni costrutto ha a sua volta tre indicatori, in particolare:

- il costrutto X ha indicatori x1, x2 e x3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;
- il costrutto Y ha indicatori y1, y2 e y3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;
- il costrutto Z ha indicatori z1, z2 e z3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8.

La parte strutturale prevede relazioni causali invertite rispetto al Modello di regressione multivariata, ossia da Y a X e da Z a X. I parametri strutturali assumono i valori 0.2 o 0.7. Si avranno quattro modelli, come mostrato nella Tabella 2.5. Si vogliono quindi analizzare eventuali cambiamenti dovuti alla direzione della causalità.

TABELLA 2.5: Parametri strutturali nei modelli di regressione multipla simulati.

	Mod 3a	Mod 3b	Mod 3c	Mod 3d
$Y \rightarrow X$	0.2	0.2	0.7	0.7
$Z \rightarrow X$	0.2	0.7	0.2	0.7

La Figura 2.13 illustra il modello di simulazione. I valori dei parametri strutturali non sono inseriti, in quanto cambiano nelle diverse simulazioni.

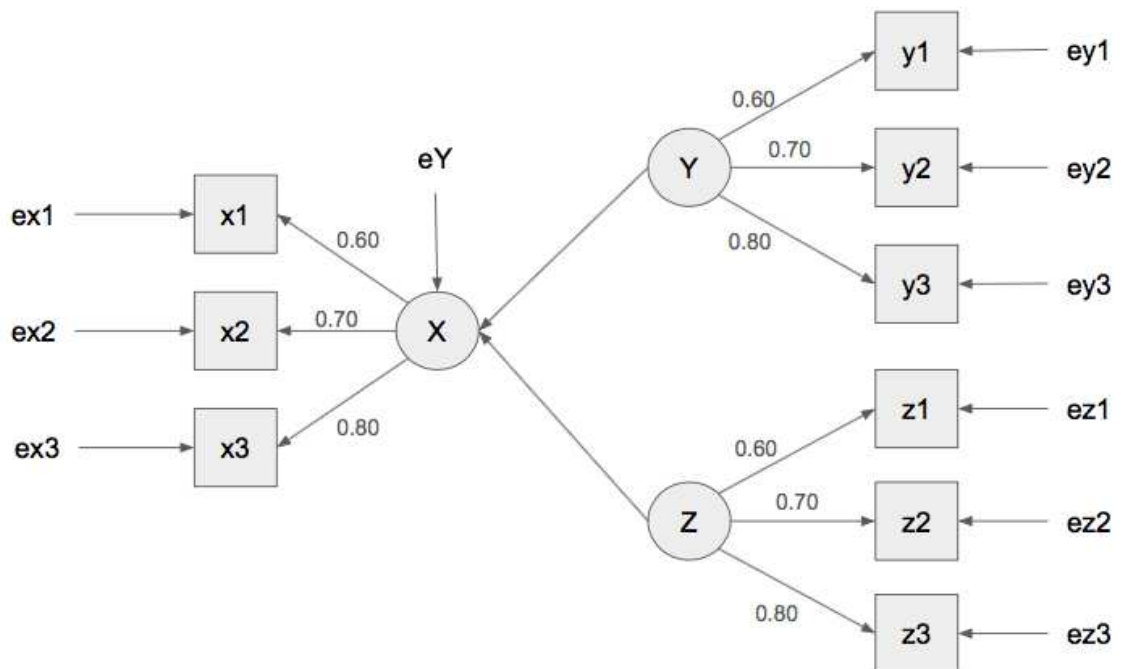


FIGURA 2.13: Modello di regressione multipla a tre costrutti riflessivi X, Y e Z, ognuno con tre indicatori, con relazione causale da Y a X e da Z a X

La Figura 2.14 conferma ciò che è stato detto finora. Il metodo PLS continua a dare problemi quando le relazioni strutturali sono entrambe deboli. Anche una sola relazione forte consente stime migliori; se sono entrambe forti, le stime si avvicinano a quelle del CB-SEM. Nell'Appendice al paragrafo A.3 sono contenuti anche i grafici relativi alle stime dei parametri di misura per x2 e x3.

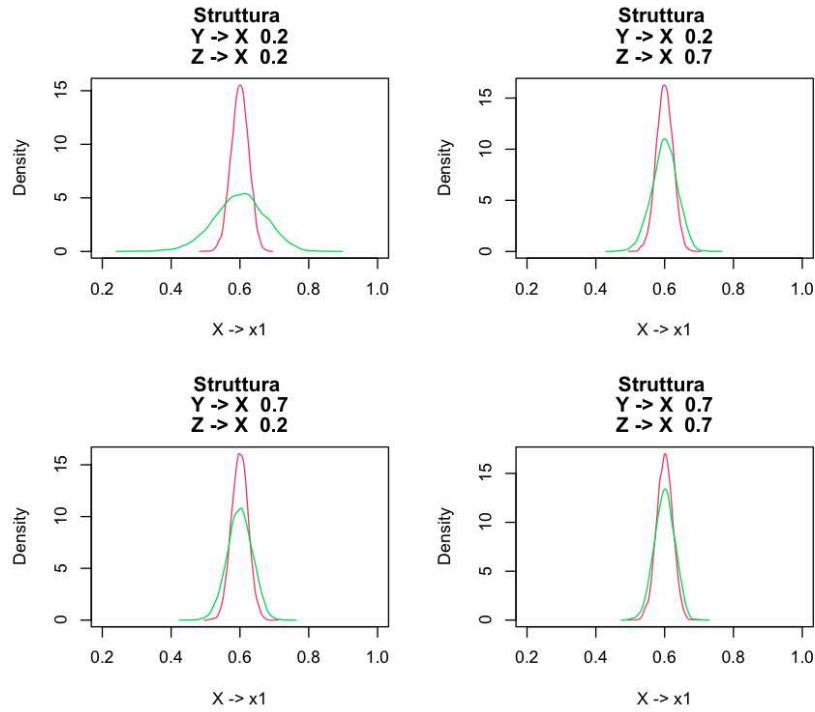


FIGURA 2.14: Distribuzione delle stime del parametro di misura $X \rightarrow x_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Anche per gli indicatori di Y il metodo CB continua a fornire risultati centrati nei veri valori dei parametri e con deviazione standard bassa, mentre il metodo PLS dimostra ancora problemi legati alla varianza elevata per valori strutturali bassi. Si osserva però che, anche con valori alti, la varianza rimane superiore a quella del CB. La seguente Figura 2.15 dimostra quanto detto. Per le stime dei parametri di misura per y_2 e y_3 si veda l'Appendice al paragrafo A.3.

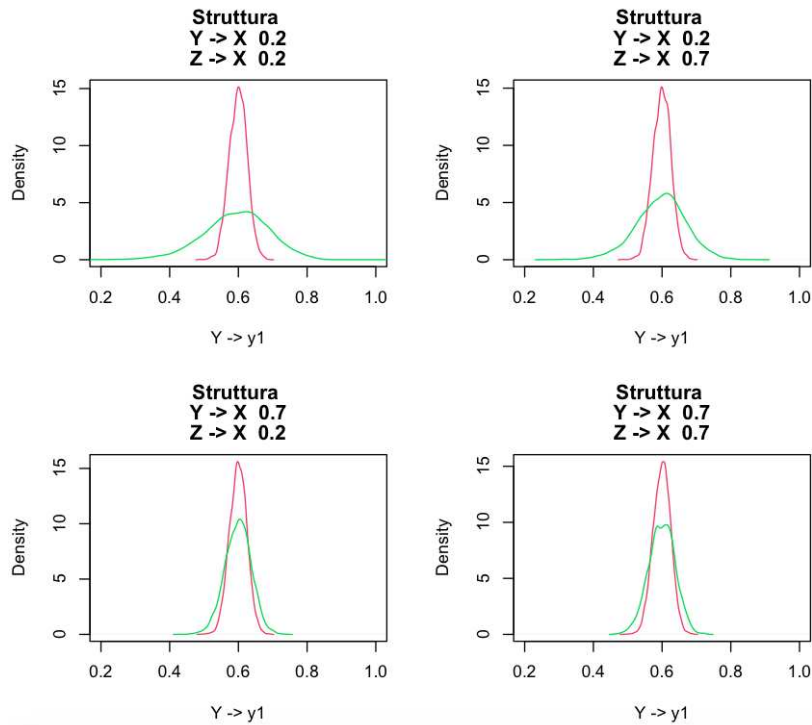


FIGURA 2.15: Distribuzione delle stime del parametro di misura $Y \rightarrow y_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

I risultati sono analoghi anche per le stime dei parametri di misura da Z a z_1 , z_2 e z_3 . Si veda la seguente Figura 2.16 per z_1 e le Figure nel paragrafo A.3 dell'Appendice per z_2 e z_3 . Sembra quindi che, dovendo "condividere" l'effetto su X, per Y e Z i modelli di misura non sfruttino completamente la forza della componente strutturale.

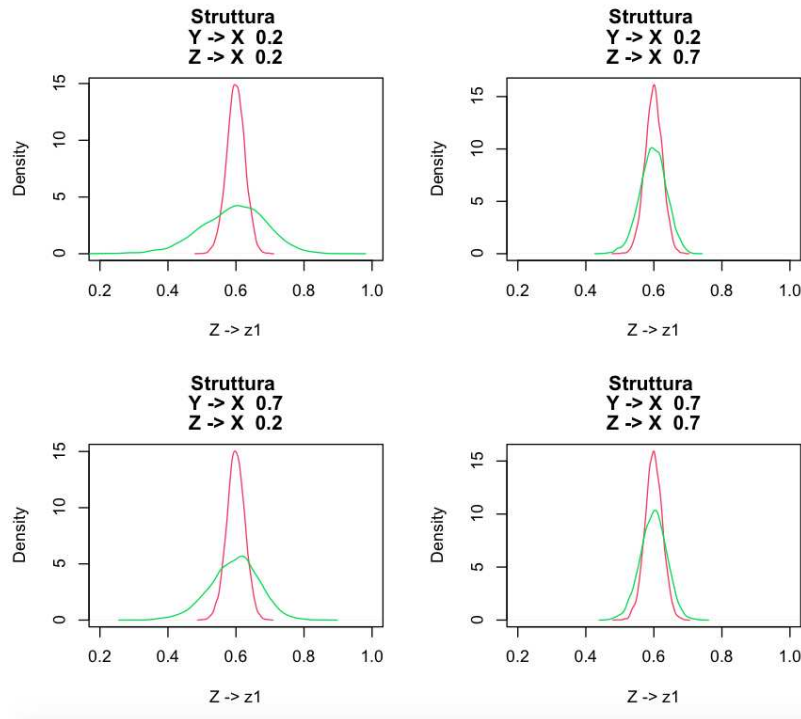


FIGURA 2.16: Distribuzione delle stime del parametro di misura $Z \rightarrow z1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Le stime della parte strutturale sono riportate nell'Appendice al paragrafo A.3 e valgono commenti analoghi a quelli dei paragrafi precedenti, ossia le stime dei parametri strutturali ottenute con CB e PLS sono tra loro molto simili.

Anche per questo modello il comportamento del metodo PLS migliora con l'aumentare dei valori strutturali. Inoltre, non è influenzato dalla direzione dei nessi causali, quanto piuttosto alla loro presenza e forza. Quindi anche questo tipo di modello porta a pensare che il confronto tra CB e PLS sia possibile solo quando le relazioni strutturali sono forti. Nel prossimo paragrafo si specifica un modello ricorsivo che, rispetto al modello di regressione multivariata, introduce una relazione causale anche tra Y e Z .

2.5 Modello ricorsivo

In questo paragrafo si procede con delle simulazioni da un modello ricorsivo a tre costrutti riflessivi. Ogni costrutto avrà tre indicatori, in particolare:

- il costrutto X ha indicatori $x1$, $x2$ e $x3$ con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;

- il costrutto Y ha indicatori y1, y2 e y3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8;
- il costrutto Z ha indicatori z1, z2 e z3 con parametri di misura pari rispettivamente a 0.6, 0.7 e 0.8.

La parte strutturale riprende il Modello di regressione multivariata e aggiunge una relazione causale da Y a Z, come mostrato in Figura 2.17.

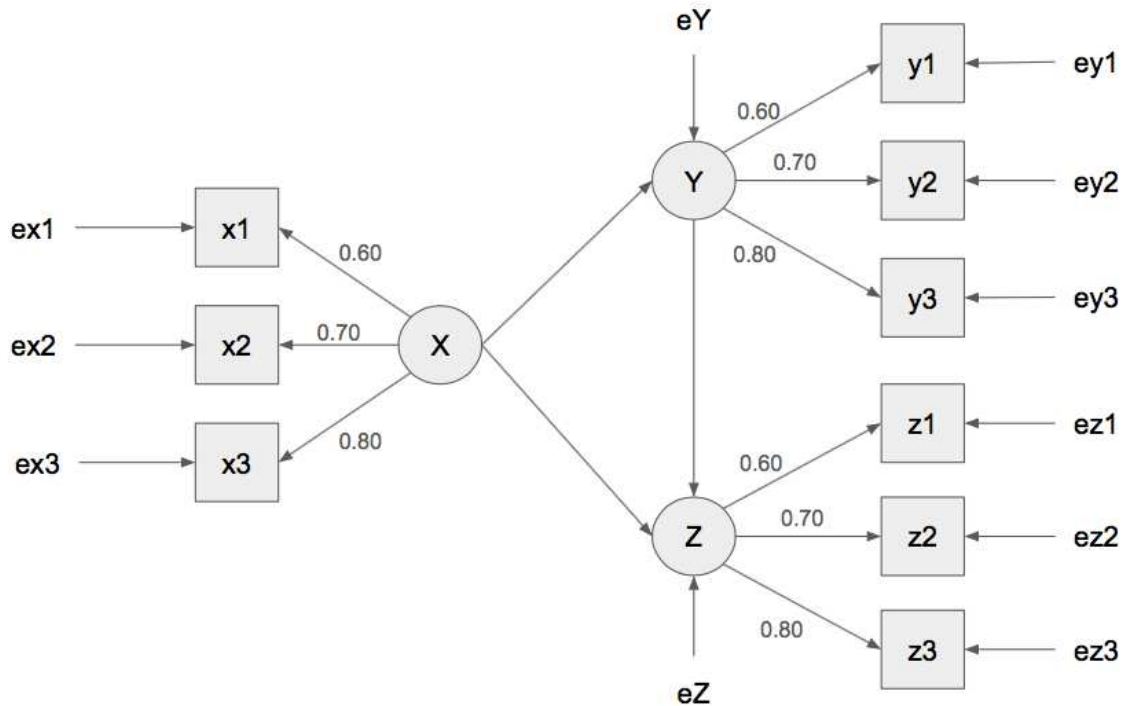


FIGURA 2.17: Modello ricorsivo con tre costrutti riflessivi X, Y e Z, ognuno con tre indicatori, con relazione causale da X a Y, da X a Z e da Y a Z.

Come per le precedenti simulazioni, i parametri di misura assumono i valori 0.2 o 0.7. Avremo quindi otto simulazioni, come mostrato nella Tabella 2.6.

TABELLA 2.6: Parametri strutturali nei modelli ricorsivi simulati.

	Mod 4a	Mod 4b	Mod 4c	Mod 4d	Mod 4e	Mod 4f	Mod 4g	Mod 4h
X → Y	0.2	0.2	0.2	0.2	0.7	0.7	0.7	0.7
X → Z	0.2	0.2	0.7	0.7	0.2	0.2	0.7	0.7
Y → Z	0.2	0.7	0.2	0.7	0.2	0.7	0.2	0.7

Anche in questo caso le stime PLS risentono della presenza di relazioni strutturali deboli e migliorano all'aumentare della forza della relazione strutturale.

La Figura 2.18 mostra come, anche per questo tipo di modello, il PLS fornisca stime peggiori rispetto al CB, in particolare quando le relazioni strutturali $X \rightarrow Y$ e $X \rightarrow Z$ sono deboli. Si nota, inoltre, che la relazione strutturale $Y \rightarrow Z$ non ha influenza sulle stime di x_1 . Come per i precedenti modelli, è sufficiente una relazione strutturale forte affinché le stime del PLS migliorino. Nell'Appendice, al paragrafo A.4, si trovano i medesimi grafici per x_2 e x_3 , per le quali valgono commenti analoghi.

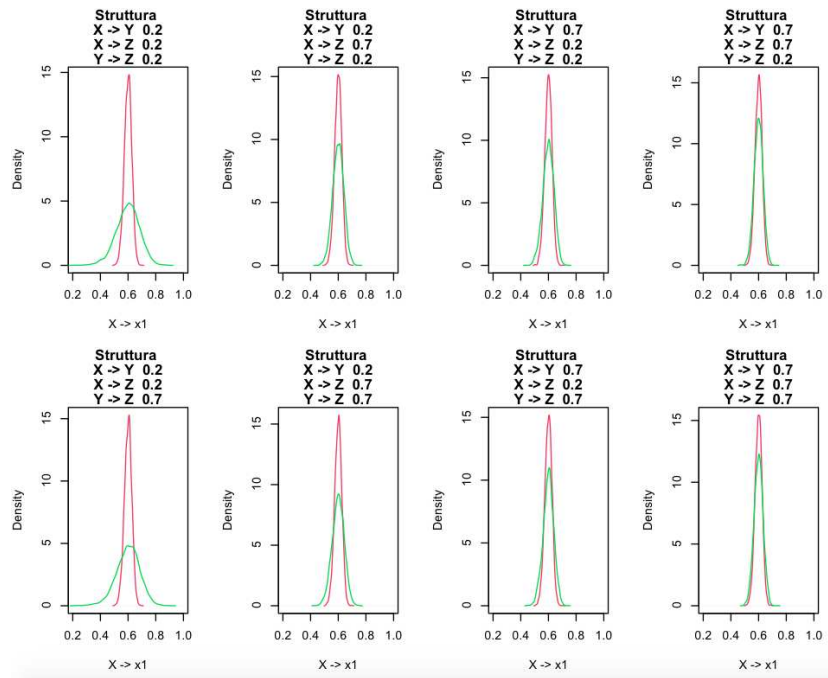


FIGURA 2.18: Distribuzione delle stime del parametro di misura $X \rightarrow x_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Le stesse riflessioni valgono anche per i parametri di misura di Y . In questo caso la relazione $X \rightarrow Z$ non ha influenza sulle stime PLS. È sufficiente che una delle relazioni strutturali, tra $X \rightarrow Y$ e $Y \rightarrow Z$, sia forte affinché le stime del PLS migliorino. Nell'Appendice, al paragrafo A.4, si trovano i medesimi grafici per y_2 e y_3 , per le quali valgono commenti analoghi.

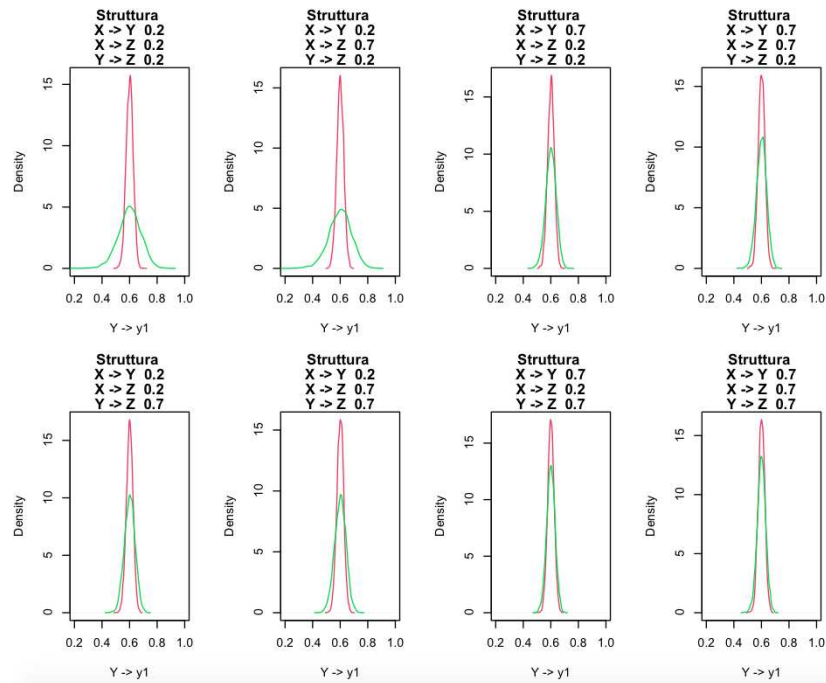


FIGURA 2.19: Distribuzione delle stime del parametro di misura $Y \rightarrow y_1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Infine, per quanto riguarda il costrutto Z, si nota che le uniche relazioni strutturali ad avere effetto sulle stime PLS sono $X \rightarrow Z$ e $Y \rightarrow Z$. È sufficiente che una di queste relazioni strutturali sia forte affinché le stime del PLS migliorino. Nell'Appendice, al paragrafo A.4, si trovano i medesimi grafici per z2 e z3, per le quali valgono commenti analoghi.

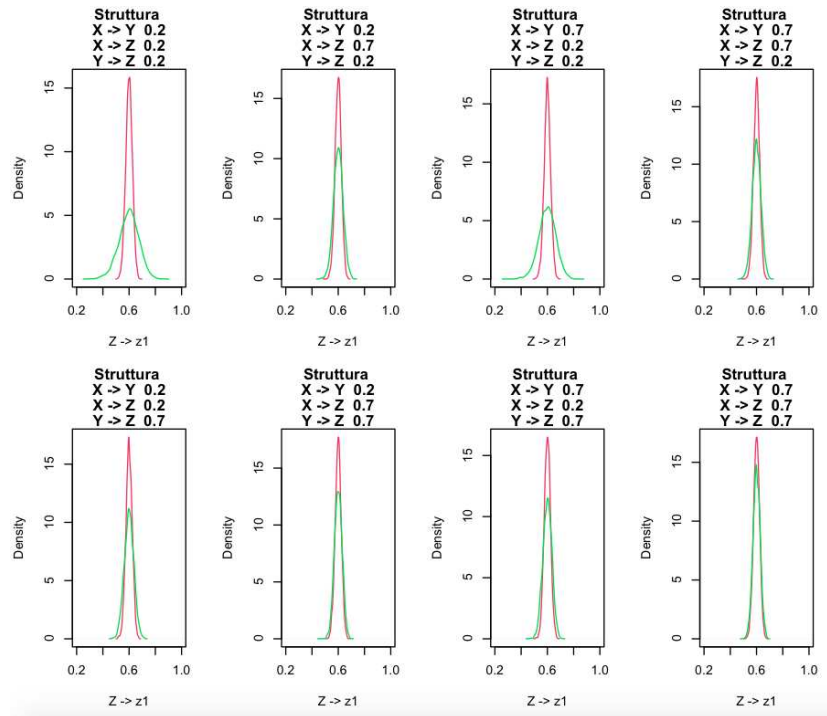


FIGURA 2.20: Distribuzione delle stime del parametro di misura $Z \rightarrow z1$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Per questo tipo di modello si conclude che il PLS sia influenzato dalle relazioni strutturali e dalla loro intensità, come risulta per tutti i modelli simulati in precedenza. Le stime migliorano quando i veri valori dei parametri strutturali aumentano.

2.6 Sintesi

Le simulazioni effettuate nei paragrafi precedenti evidenziano il differente comportamento dei due metodi Covariance Based e Partial Least Squares nella stima di alcuni tipi di modelli riflessivi. Dai risultati ottenuti si nota che il metodo CB fornisce stime mediamente centrate nel vero valore del parametro e con deviazione standard contenuta, indipendentemente dalla forma e forza delle relazioni strutturali. Il metodo PLS, invece, ha un comportamento che migliora al crescere del valore dei parametri strutturali. Questo vuol dire che la stima dei parametri di misura nel PLS dipende dalla parte strutturale, in particolare dalla sua forza più che dalla forma.

Nei modelli riflessivi, per loro natura, la parte di misura non dipende dalla parte strutturale, quindi il metodo adottato dovrebbe dare risultati simili indipendentemente dal valore del parametro strutturale. Quindi, si può concludere che il confronto dei due modelli è fattibile solo quando il vero valore dei parametri strutturali è alto. Tenendo

conto del fatto che i due modelli stimano bene i parametri strutturali, è anche possibile applicare i due modelli a dati reali e verificare se le stime della parte strutturale sono alte. Allora in quel caso i due modelli dovrebbero stimare similmente anche la parte di misura e quindi possono essere utilizzati entrambi.

2.7 Modello di regressione semplice: approfondimenti

2.7.1 Costrutti con un numero diverso di indicatori

Un possibile sviluppo consta nell'analisi del comportamento dei due metodi quando il numero di indicatori cambia. In particolare, potrebbe essere utile verificare cosa accade quando i costrutti hanno quattro o due indicatori. Come prima esplorazione, nel seguito si procede simulando da un modello di regressione semplice, prima con costrutti da quattro indicatori e poi con costrutti da due indicatori ciascuno.

Il primo è un modello di regressione lineare semplice con due costrutti X e Y, rispettivamente esogeno ed endogeno, da quattro indicatori ciascuno. Le Figure 2.21 e 2.22 confermano quanto detto nei paragrafi precedenti. Si accentua la differenza nella media delle stime dei due metodi per valori bassi del parametro strutturale. Questo potrebbe essere un ulteriore punto di approfondimento per ricerche future. Le deviazioni standard restano molto differenti tra i due metodi, in particolare molto alte per il metodo PLS quando la relazione strutturale è debole.

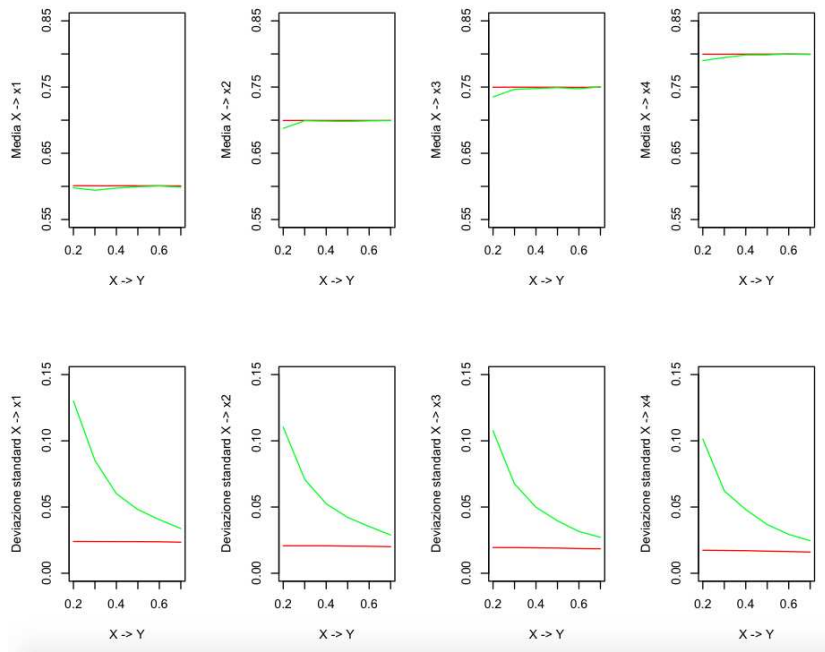


FIGURA 2.21: Media e deviazione standard delle stime dei parametri di misura del costrutto X (a 4 indicatori) con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

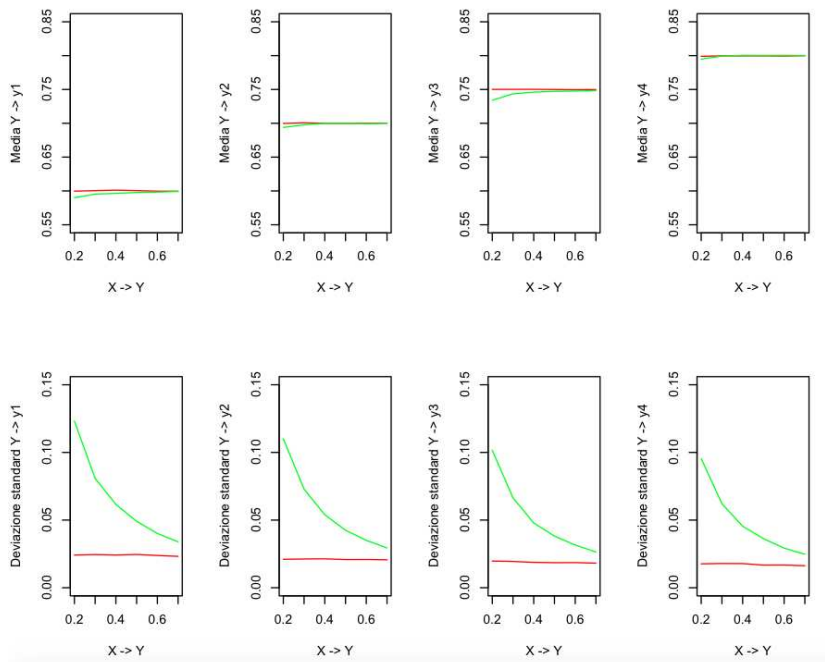


FIGURA 2.22: Media e deviazione standard delle stime dei parametri di misura del costrutto Y (a 4 indicatori) con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

Simulando invece da un modello di regressione semplice con costrutti da due indicatori, il metodo CB segnala, in alcuni casi, degli errori nella stima, che potrebbero dipendere da problemi di identificazione. Prendendo in considerazione solo le simulazioni per le quali entrambi i metodi arrivano a convergenza, si nota un peggioramento del metodo CB circa la deviazione standard delle stime. Si ha infatti che il metodo PLS non subisce variazioni rispetto a quanto detto precedentemente. Ciò che cambia è appunto la varianza delle stime ottenute con il metodo CB, la quale subisce un aumento non banale. Sarebbe interessante approfondire il tema dei costrutti con pochi indicatori in sviluppi futuri.

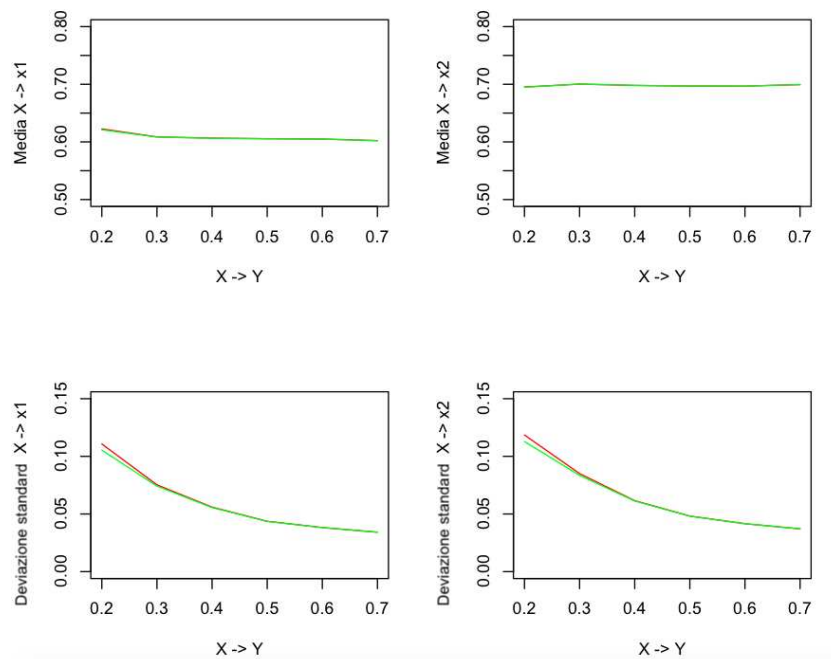


FIGURA 2.23: Media e deviazione standard delle stime dei parametri di misura del costrutto X (a 2 indicatori) con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

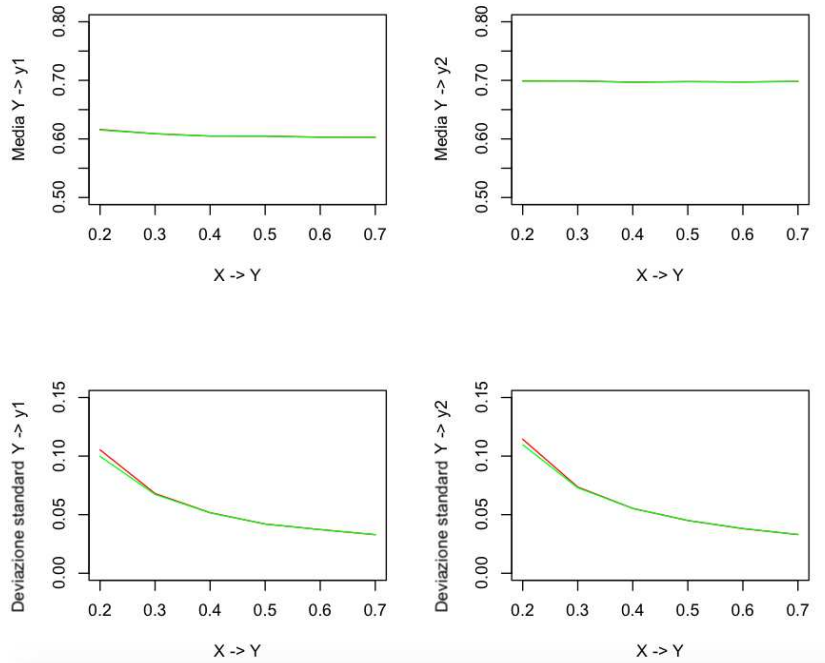


FIGURA 2.24: Media e deviazione standard delle stime dei parametri di misura del costruito Y (a 2 indicatori) con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

2.7.2 Standard error e Verifica d'ipotesi

Nelle sezioni precedenti è stato analizzato il comportamento delle stime puntuali ottenute con i due metodi CB e PLS. Si fa notare che si è parlato di deviazioni standard delle stime e non di standard error dei parametri. Sarebbe interessante approfondire questa tematica, per capire se la parte strutturale del modello influenzi effettivamente anche gli standard error.

In questo breve paragrafo si fornisce un'introduzione all'argomento, facendo delle simulazioni da un modello di regressione semplice e analizzando gli standard error medi ottenuti con i due metodi. In particolare, si fanno 1000 simulazioni e, per ognuna di esse, si estraggono gli standard error ottenuti con il metodo CB e con il metodo PLS. Si fa notare che lo standard error ottenuto con il metodo PLS necessita di simulazioni bootstrap, vista la sua natura non parametrica. Tuttavia, come si vedrà anche nel Capitolo 3 sulle applicazioni a dati reali, non persiste una differenza degna di nota tra gli standard error ottenuti con il metodo CB applicato all'insieme di dati originale e quello ottenuto tramite simulazioni bootstrap. La Figura 2.25 riporta il comportamento degli standard error per i parametri di misura relativi agli indicatori x_1 , x_2 , x_3 , y_1 , y_2 e y_3 .

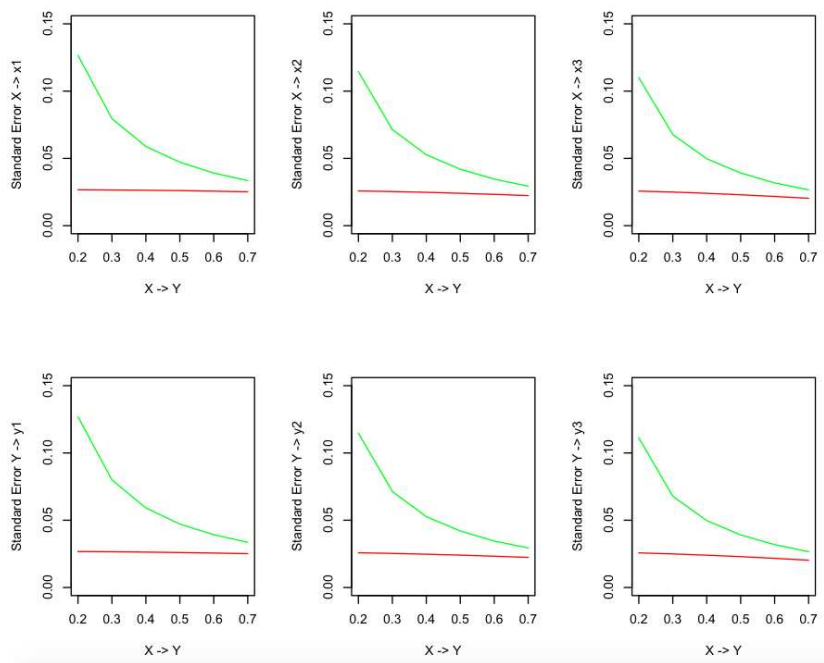


FIGURA 2.25: Media degli standard error dei parametri di misura per un modello di regressione semplice, ottenuti con CB (linea rossa) e PLS (linea verde) al crescere del parametro strutturale $X \rightarrow Y$.

Si nota che gli standard error ottenuti con il metodo CB non sono influenzati dal vero valore del parametro strutturale, come potevasi immaginare. Al contrario, quelli ottenuti con il metodo PLS sono decrescenti all'aumentare dell'intensità della relazione strutturale. I comportamenti degli standard error sono in genere molto simili alle deviazioni standard ottenute per le stime puntuali dei parametri, come si può notare nei paragrafi precedenti.

L'ultima analisi proposta vuole verificare in quante delle 1000 simulazioni i modelli non rifiutano l'ipotesi che il singolo parametro sia uguale al suo vero valore. Ad esempio, la Tabella 2.7 mostra quante simulazioni non rifiutano l'ipotesi H_0 , illustrata nel sistema d'ipotesi 2.1.

$$\begin{cases} H_0 : X \rightarrow x_1 = 0.6 \\ H_1 : X \rightarrow x_1 \neq 0.6 \end{cases} \quad (2.1)$$

In generale il metodo PLS rifiuta l'ipotesi nulla in più simulazioni rispetto al CB, anche se non c'è una differenza sostanziale. Questo potrebbe significare che, quando il metodo PLS fornisce una stima lontana dal vero valore del parametro, allora lo standard error

aumenta. Osservando la Figura 2.27, si nota che, nel caso di parametri strutturali deboli, le stime PLS risultano instabili. Oltre a ciò, si può notare un andamento decrescente degli standard error nel primo grafico, ossia quello riferito al modello Mod1a, con parametro strutturale $X \rightarrow Y$ pari a 0.2. Questo significa che, nei casi in cui il PLS fornisce una stima inferiore al vero valore del parametro, allora il suo standard error tende ad essere elevato, facendo sì che la relativa ipotesi nulla venga accettata. Al contrario, si nota che per stime PLS superiori al vero valore del parametro, lo standard error tende a diminuire, facendo così rifiutare l'ipotesi nulla. Risultati analoghi si ottengono per gli altri parametri di misura.

TABELLA 2.7: Numero di simulazioni in cui non viene rifiutata l'ipotesi che $X \rightarrow x1$ sia uguale a 0.6.

	Mod 1a	Mod 1b	Mod 1c	Mod 1d	Mod 1e	Mod 1f
CB $X \rightarrow x1$	958	962	957	949	952	949
PLS $X \rightarrow x1$	965	940	940	937	943	943

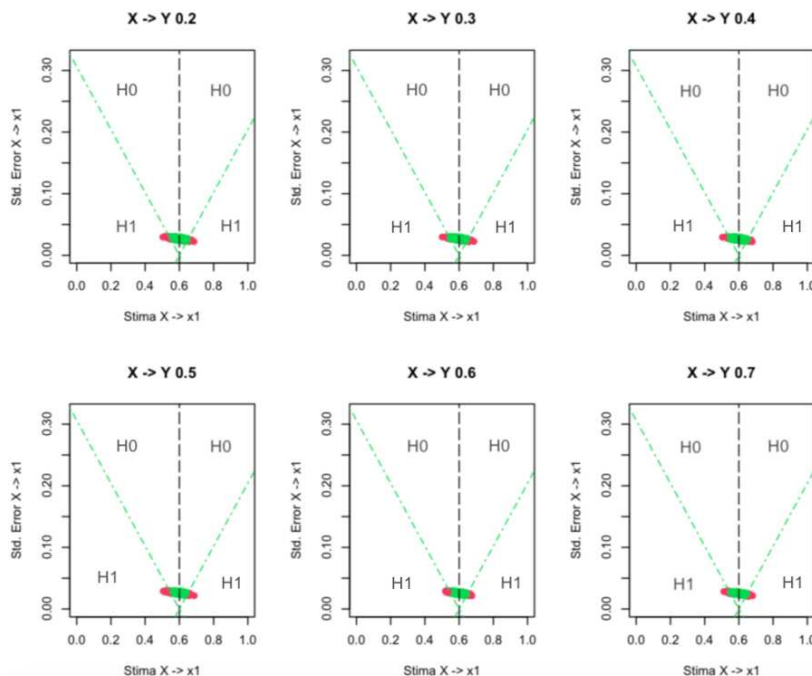


FIGURA 2.26: Grafico delle stime puntuali e dei relativi standard error per il parametro $X \rightarrow x1$ con CB. In verde le simulazioni che non rifiutano l'ipotesi nulla $H0 : X \rightarrow x1 = 0.6$. In rosso le simulazioni che rifiutano l'ipotesi nulla.

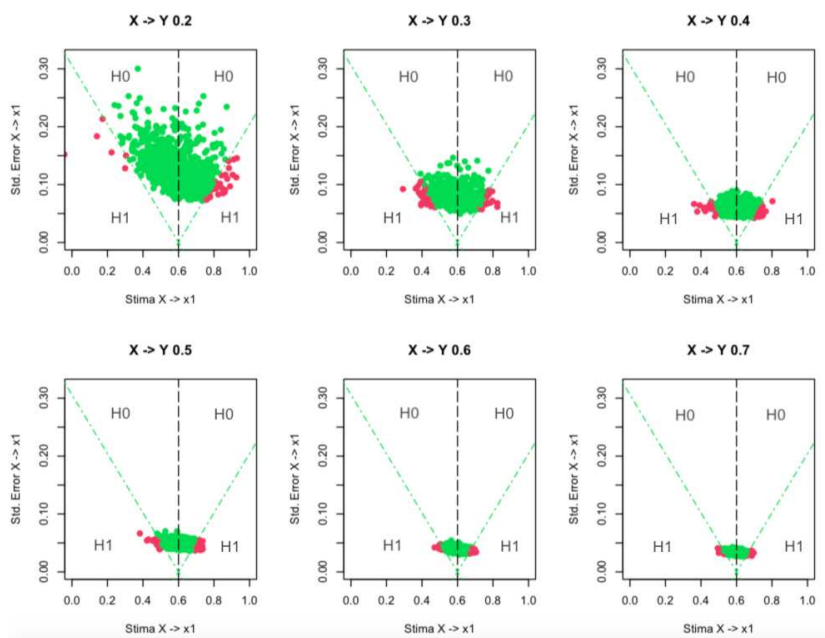


FIGURA 2.27: Grafico delle stime puntuali e dei relativi standard error per il parametro $X \rightarrow x1$ con PLS. In verde le simulazioni che non rifiutano l'ipotesi nulla. In rosso le simulazioni che rifiutano l'ipotesi nulla $H0 : X \rightarrow x1 = 0.6$.

Si prova, infine, ad analizzare il comportamento dei due metodi nella verifica d'ipotesi che $X \rightarrow x1$ sia uguale a determinati valori t . Le Tabelle 2.8 e 2.9 mostrano, rispettivamente, il numero di ipotesi non rifiutate da CB e PLS. Si nota che il metodo CB non è particolarmente influenzato dal vero valore del parametro strutturale. Oltre a ciò, esso rifiuta sempre o nella maggior parte dei casi quando si ipotizza che il parametro di $X \rightarrow x1$ sia diverso dal suo vero valore 0.6. Al contrario, il metodo PLS risulta fortemente influenzato dalla parte strutturale: si vede infatti che, al crescere del parametro strutturale, il numero di ipotesi accettate diminuisce fortemente, soprattutto quando si ipotizza che $X \rightarrow x1$ sia diverso dal suo vero valore 0.6. Infine si nota che, per valori inferiori a 0.6, il metodo accetta in meno casi rispetto alle ipotesi con valori superiori a 0.6, quando la relazione strutturale è debole, e in più casi quando la relazione strutturale è forte. Questo conferma il fatto che il PLS funzioni meglio quando le relazioni strutturali sono forti e che fornisca standard error maggiori quando sottostima il parametro e standard error minori quando sovrastima il parametro.

TABELLA 2.8: Numero di simulazioni in cui CB non rifiuta l'ipotesi che $X \rightarrow x_1$ sia uguale a un determinato valore t .

	Mod 1a	Mod 1b	Mod 1c	Mod 1d	Mod 1e	Mod 1f
$t = 0.3$	0	0	0	0	0	0
$t = 0.4$	0	0	0	0	0	0
$t = 0.5$	63	53	47	49	39	40
$t = 0.6$	958	962	957	949	952	949
$t = 0.7$	16	13	14	14	13	12
$t = 0.8$	0	0	0	0	0	0
$t = 0.9$	0	0	0	0	0	0

TABELLA 2.9: Numero di simulazioni in cui PLS non rifiuta l'ipotesi che $X \rightarrow x_1$ sia uguale a un determinato valore t .

	Mod 1a	Mod 1b	Mod 1c	Mod 1d	Mod 1e	Mod 1f
$t = 0.3$	368	60	3	1	0	0
$t = 0.4$	562	284	87	21	1	0
$t = 0.5$	829	728	561	448	300	169
$t = 0.6$	965	940	940	937	943	943
$t = 0.7$	936	778	652	407	270	159
$t = 0.8$	652	294	77	13	0	0
$t = 0.9$	248	31	1	0	0	0

Capitolo 3

Applicazioni a dati reali

3.1 Introduzione

In questo Capitolo si adattano i modelli descritti e simulati nel Capitolo 2 a dati reali. In particolare, i dati si riferiscono a dei questionari somministrati a 1090 clienti dei ristoranti di una catena. Le domande sono in scala da 1 a 7 e riguardano la soddisfazione dei clienti per alcuni aspetti della loro esperienza in questi ristoranti. Alcune domande riguardano aspetti legati al cibo, quali ad esempio la presentazione del piatto e il gusto della pietanza. Altri aspetti riguardano il servizio ricevuto dal personale, come l'efficienza e la gentilezza dello staff e ancora aspetti legati al design, come i colori e la luminosità dei locali.

Oltre a questi aspetti più concreti, sono state sottoposte alcune domande legate alle emozioni positive e negative provate, come ad esempio l'aver provato felicità oppure noia durante l'esperienza nei ristoranti. Infine, è stato chiesto un giudizio complessivo, in termini di consiglio del ristorante ad altre persone o alla possibilità di ritornare in futuro.

Ecco quindi che si potrebbe ipotizzare la presenza di sei macrogruppi legati al *Cibo*, al *Design*, al *Servizio*, alle *Emozioni Negative*, alle *Emozioni Positive* e al *Giudizio Finale*. Nelle prossime analisi non si vuole trovare il modello che meglio si adatta ai dati, ma semplicemente confrontare le differenze tra i due metodi Covariance Based e Partial Least Squares nella stima dei modelli stessi.

3.2 Modello di regressione semplice

Il primo modello che si vuole adattare ipotizza la presenza di due costrutti latenti, il *Design* e le *Emozioni Positive*. Per il *Design* gli indicatori utilizzati sono il *Design Interno*, il *Design Esterno* e i *Colori* dei locali, per semplicità D1, D2 e D3. Le *Emozioni Positive* hanno invece come indicatori il *Divertimento*, l'*Interesse* e la *Felicità*, chiamati brevemente EP1, EP2 ed EP3. Tutti gli indicatori sono soggetti a errore di misura. Si ipotizza che la soddisfazione provata per il *Design* del ristorante comporti un effetto causale verso le *Emozioni Positive*. La Figura 3.1 mostra infatti i due costrutti riflessivi con relazione causale rappresentata dalla freccia che ha punto di origine in *Design* e punto di arrivo nelle *Emozioni Positive*.

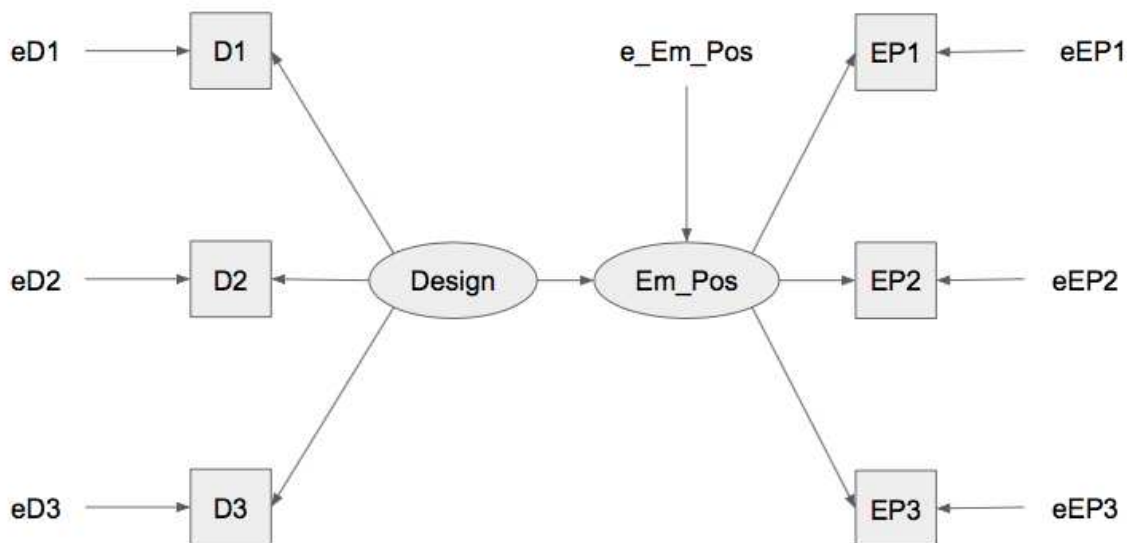


FIGURA 3.1: Modello di regressione semplice delle *Emozioni Positive* sul costrutto del *Design*.

Si fa notare che, essendo il PLS un metodo non parametrico, l'inferenza è possibile solo ricorrendo a simulazioni bootstrap. Per completezza si sceglie di calcolare lo standard error per il metodo CB anche tramite simulazioni bootstrap. La Tabella 3.1 contiene le stime e gli standard error ottenuti con il metodo parametrico CB. La Tabella 3.2 contiene le medie e gli standard error dei parametri ottenuti ricorrendo a simulazioni bootstrap sia per il metodo CB e ovviamente anche per il metodo PLS.

Come si può notare nella Tabella 3.1, tutte le stime sono elevate e anche il parametro strutturale è stimato pari a 0.617. In Tabella 3.2 le stime ottenute con i due metodi sono tra loro simili, salvo in alcuni casi come per i parametri Em_pos \rightarrow EP2 e Em_pos

TABELLA 3.1: Stime del modello di regressione semplice ottenute con il metodo Covariance-Based.

Parametro	Stima CB	S.e. CB
Design → D1	0.745	0.024
Design → D2	0.839	0.027
Design → D3	0.861	0.016
Em_pos → EP1	0.823	0.021
Em_pos → EP2	0.854	0.015
Em_pos → EP3	0.831	0.025
Design → Em_pos	0.617	0.031

TABELLA 3.2: Stime del modello di regressione semplice ottenute con i metodi CB e PLS tramite simulazioni bootstrap.

Parametro	Media CB	S.e. CB	Media PLS	S.e. PLS
Design → D1	0.745	0.023	0.732	0.028
Design → D2	0.839	0.027	0.830	0.027
Design → D3	0.861	0.017	0.883	0.024
Em_pos → EP1	0.822	0.021	0.817	0.029
Em_pos → EP2	0.853	0.014	0.930	0.024
Em_pos → EP3	0.831	0.026	0.761	0.034
Design → Em_pos	0.616	0.031	0.612	0.031

→ EP3 che presentano valori abbastanza diversi. Anche gli standard error sono diversi. Ad esempio gli standard error per il parametro Em_pos → EP2 ottenuti con CB e PLS sono rispettivamente pari a 0.015 per il CB (0.014 tramite simulazioni bootstrap) e 0.024 per il metodo PLS. Il rapporto tra questi due valori è pari a 1.6.

3.3 Modello di regressione multivariata

Il secondo modello applicato è un Modello di regressione multivariata che, rispetto al modello di regressione semplice, aggiunge una relazione causale dal *Design* alle *Emozioni Negative*. Le *Emozioni Negative* costituiscono un costrutto latente misurato dagli indicatori *Rabbia*, *Stress* e *Noia*, chiamati brevemente EN1, EN2 e EN3. La Figura 3.2 rappresenta il modello adattato ai dati.

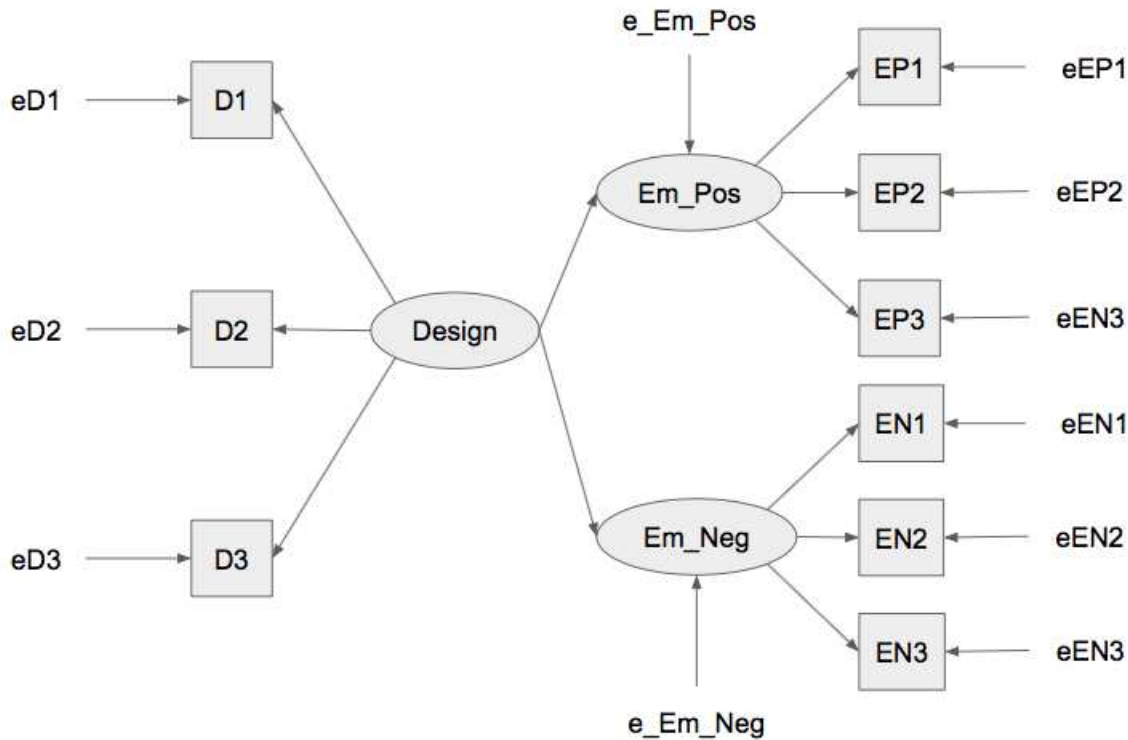


FIGURA 3.2: Modello di regressione multivariata delle *Emozioni Positive* e *Negative* sul costrutto del *Design*.

Le relazioni strutturali $\text{Design} \rightarrow \text{Em_pos}$ e $\text{Design} \rightarrow \text{Em_neg}$ sono stimate rispettivamente pari a 0.617 e -0.473, come mostra la Tabella 3.3. La Tabella 3.4 presenta medie molto simili tra i due metodi. Se si considera il parametro $\text{Em_neg} \rightarrow \text{EN3}$, lo standard error ottenuto con metodo CB è pari a 0.022 (0.022 anche tramite simulazioni bootstrap), mentre lo standard error ottenuto con metodo PLS è pari a 0.043, ossia quasi il doppio. Questo è coerente con le simulazioni del Capitolo 2, in quanto la relazione strutturale delle *Emozioni Negative* è meno forte.

TABELLA 3.3: Stime del modello di regressione multivariata ottenute con il metodo Covariance-Based.

Parametro	Stima CB	S.e. CB
Design → D1	0.747	0.024
Design → D2	0.838	0.026
Design → D3	0.859	0.016
Em_pos → EP1	0.827	0.020
Em_pos → EP2	0.851	0.014
Em_pos → EP3	0.830	0.024
Em_neg → EN1	0.816	0.027
Em_neg → EN2	0.871	0.020
Em_neg → EN3	0.808	0.022
Design → Em_pos	0.617	0.031
Design → Em_neg	-0.473	0.040

TABELLA 3.4: Stime del modello di regressione multivariata ottenute con i metodi CB e PLS tramite simulazioni bootstrap.

Parametro	Media CB	S.e. CB	Media PLS	S.e. PLS
Design → D1	0.747	0.023	0.756	0.028
Design → D2	0.839	0.027	0.824	0.028
Design → D3	0.859	0.016	0.865	0.025
Em_pos → EP1	0.826	0.020	0.816	0.030
Em_pos → EP2	0.851	0.014	0.926	0.024
Em_pos → EP3	0.830	0.025	0.761	0.034
Em_neg → EN1	0.817	0.026	0.900	0.034
Em_neg → EN2	0.872	0.021	0.833	0.039
Em_neg → EN3	0.808	0.022	0.757	0.043
Design → Em_pos	0.616	0.031	0.613	0.031
Design → Em_neg	-0.470	0.040	-0.479	0.040

3.4 Modello di regressione multipla

In questo paragrafo si adatta il Modello di regressione multipla visto nel Capitolo 2. In questo caso, si tengono i costrutti delle *Emozioni Positive* e delle *Emozioni Negative* e si aggiunge il costrutto latente del *Giudizio Positivo* finale. I suoi indicatori sono le variabili inerenti al *Ritorno* nel ristorante, alla *Raccomandazione* del ristorante ad amici e al *Parlare Bene* del ristorante ad altre persone. Le relazioni causali ipotizzate vanno dalle *Emozioni Positive* al *Giudizio Positivo* e dalle *Emozioni Negative* al *Giudizio Positivo*. La Figura 3.3 mostra il grafico del modello.

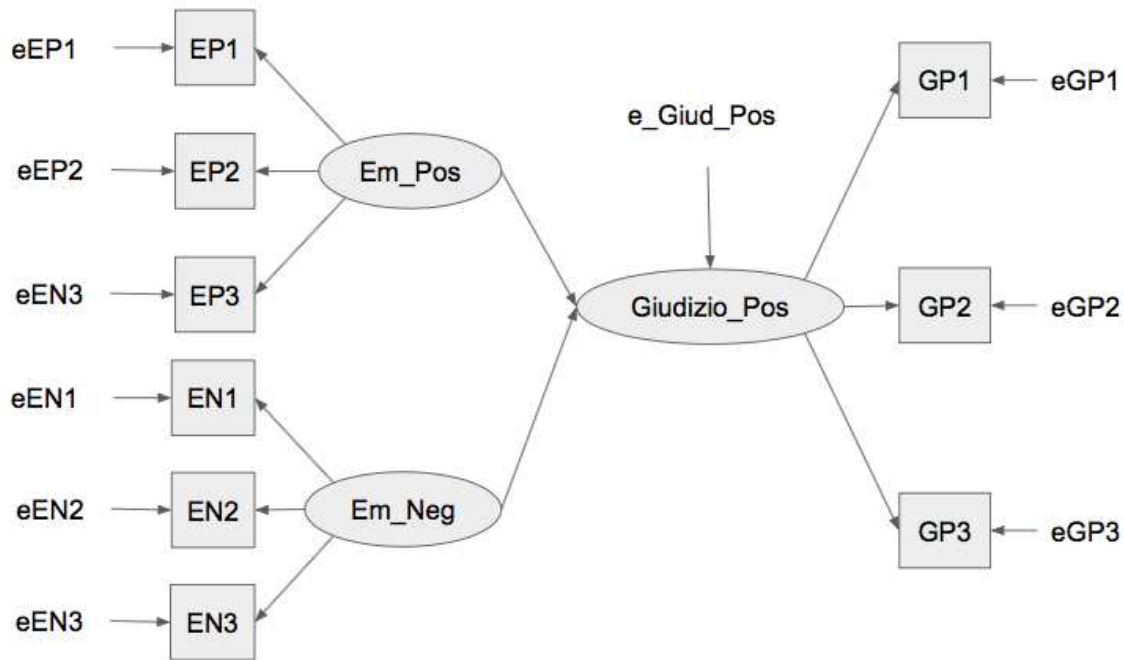


FIGURA 3.3: Modello di regressione multipla del *Giudizio Positivo* sulle *Emozioni Positive e Negative*.

Le Tabelle 3.5 e 3.6 mostrano, nel complesso, stime simili fra i due metodi. Gli effetti delle emozioni sono entrambi significativi ma le *Emozioni Positive* hanno un effetto più forte.

TABELLA 3.5: Stime del modello di regressione multipla ottenute con il metodo Covariance-Based.

Parametro	Stima CB	S.e. CB
Em_pos → EP1	0.843	0.018
Em_pos → EP2	0.840	0.016
Em_pos → EP3	0.826	0.024
Em_neg → EN1	0.815	0.027
Em_neg → EN2	0.876	0.019
Em_neg → EN3	0.804	0.023
Giudizio_pos → GP1	0.893	0.012
Giudizio_pos → GP2	0.886	0.014
Giudizio_pos → GP3	0.934	0.006
Em_pos → Giudizio_pos	0.622	0.033
Em_neg → Giudizio_pos	-0.226	0.036

Anche per questo modello si notano delle differenze negli standard error di alcuni parametri di misura. Si prenda in considerazione ad esempio il parametro Em_neg → EN1.

TABELLA 3.6: Stime del modello di regressione multipla ottenute con i metodi CB e PLS tramite simulazioni bootstrap.

Parametro	Media CB	S.e. CB	Media PLS	S.e. PLS
Em_pos → EP1	0.842	0.018	0.872	0.019
Em_pos → EP2	0.839	0.015	0.849	0.018
Em_pos → EP3	0.825	0.025	0.784	0.027
Em_neg → EN1	0.816	0.026	0.846	0.036
Em_neg → EN2	0.877	0.019	0.877	0.027
Em_neg → EN3	0.803	0.022	0.769	0.034
Giudizio_pos → GP1	0.892	0.013	0.867	0.019
Giudizio_pos → GP2	0.885	0.014	0.913	0.010
Giudizio_pos → GP3	0.934	0.006	0.930	0.011
Em_pos → Giudizio_pos	0.623	0.034	0.635	0.033
Em_neg → Giudizio_pos	-0.225	0.035	-0.210	0.034

Lo standard error ottenuto con il metodo CB è pari a 0.027 (0.026 con simulazioni bootstrap) e a 0.036 con il metodo PLS. Il rapporto è pari a 1.4. Anche in questo caso si confermano differenze maggiori dove il parametro strutturale è meno forte.

3.5 Modello ricorsivo

Infine si adatta ai dati il Modello ricorsivo. In particolare, si ipotizzano tre costrutti riflessivi, quali il *Design*, le *Emozioni Positive* e il *Giudizio Positivo*. Le relazioni strutturali vanno dal *Design* alle *Emozioni Positive*, dal *Design* al *Giudizio Positivo* e dalle *Emozioni Positive* al *Giudizio Positivo*. La Figura 3.4 mostra il grafico del modello stimato.

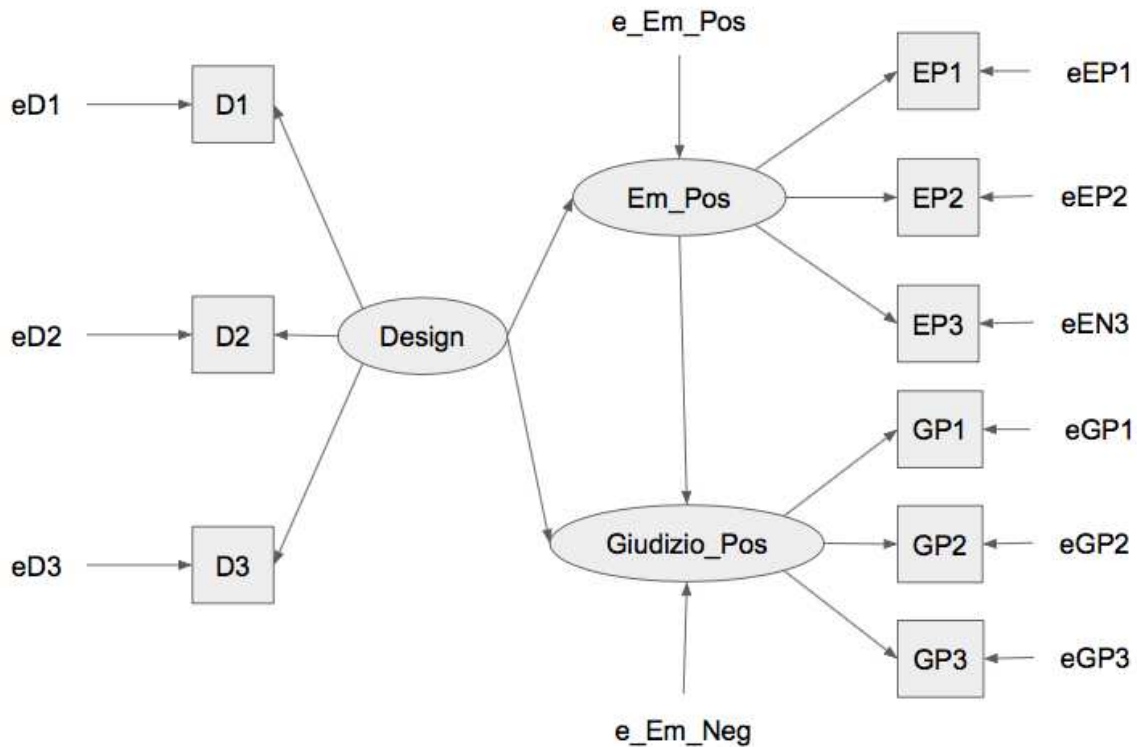


FIGURA 3.4: Modello ricorsivo.

Le Tabelle 3.7 e 3.8 riportano le stime del modello con i metodi CB e PLS: diventa più complicato analizzare il comportamento delle stime in quanto i costrutti sono tra loro tutti relazionati. Per alcuni indicatori resta una differenza non banale con i due metodi, mentre per altri i due metodi tendono a fornire risultati simili. Si nota, infine, la presenza di un parametro strutturale con stima bassa, per l'effetto diretto del *Design* sul *Giudizio*, mentre le stime sono più elevate per i parametri legati a un effetto indiretto. Con entrambi i metodi di stima si ottiene quindi una mediazione parziale delle emozioni nella relazione fra *Design* e *Giudizio*.

TABELLA 3.7: Stime del modello ricorsivo ottenute con il metodo Covariance-Based.

Parametro	Media CB	S.e. CB
Design → D1	0.743	0.023
Design → D2	0.833	0.027
Design → D3	0.867	0.016
Em_pos → EP1	0.838	0.019
Em_pos → EP2	0.847	0.014
Em_pos → EP3	0.821	0.026
Giudizio_pos → GP1	0.892	0.014
Giudizio_pos → GP2	0.895	0.013
Giudizio_pos → GP3	0.925	0.007
Design → Em_pos	0.616	0.031
Design → Giudizio_pos	0.217	0.039
Em_pos → Giudizio_pos	0.629	0.036

TABELLA 3.8: Stime del modello ricorsivo ottenute con i metodi CB e PLS tramite simulazioni bootstrap.

Parametro	Stima CB	S.e. CB	Stima PLS	S.e. PLS
Design → D1	0.744	0.024	0.734	0.029
Design → D2	0.833	0.027	0.818	0.030
Design → D3	0.867	0.016	0.893	0.021
Em_pos → EP1	0.838	0.019	0.853	0.021
Em_pos → EP2	0.847	0.015	0.878	0.017
Em_pos → EP3	0.822	0.025	0.776	0.026
Giudizio_pos → GP1	0.892	0.013	0.860	0.021
Giudizio_pos → GP2	0.896	0.013	0.953	0.009
Giudizio_pos → GP3	0.925	0.007	0.896	0.012
Design → Em_pos	0.617	0.031	0.612	0.030
Design → Giudizio_pos	0.227	0.037	0.227	0.036
Em_pos → Giudizio_pos	0.629	0.036	0.624	0.036

3.6 Modello ricorsivo: estensione a più costrutti

In quest'ultimo paragrafo si adatta ai dati un modello ricorsivo con sei costrutti riflessivi: il *Cibo*, il *Design*, il *Servizio*, le *Emozioni Positive*, le *Emozioni Negative* e il *Giudizio Positivo*. La Figura 3.5 rappresenta il modello adattato ai dati. Tutti i costrutti hanno tre indicatori, tranne il *Servizio* che ne ha quattro. L'ipotesi è che il *Cibo*, il *Design* e il *Servizio* causino direttamente le *Emozioni Positive* e le *Emozioni Negative*, le quali a loro volta causino il *Giudizio Positivo* finale. Nel modello sono inserite anche delle

relazioni dirette da *Cibo*, *Design* e *Servizio* al *Giudizio Positivo* finale.

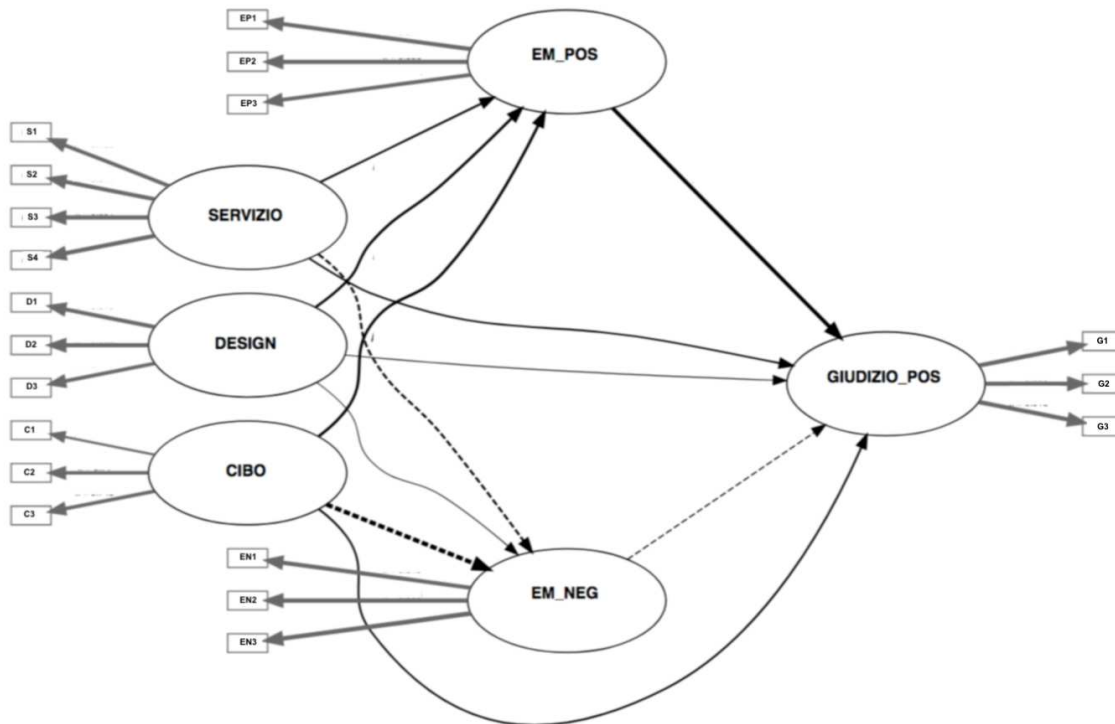


FIGURA 3.5: Modello ricorsivo con sei costrutti riflessivi.

Visto che nei precedenti modelli non si notavano sostanziali differenze tra gli standard error del metodo CB, ottenuti con e senza le simulazioni bootstrap, si riporta solo la Tabella 3.9, relativa alle stime dei parametri di misura ottenute tramite bootstrap.

Anche per questo modello, si nota che gli standard error ottenuti con il metodo PLS sono tendenzialmente maggiori rispetto a quelli ottenuti con il metodo CB. Si guardi, ad esempio, il costrutto *Servizio*: per gli indicatori S1 e S2, gli standard error ottenuti con PLS sono quasi il doppio rispetto a quelli ottenuti con CB. Infatti, il rapporto tra 0.22 e 0.12, che sono gli standard error per S1 ottenuti rispettivamente con PLS e CB, è pari a 1.8. Anche per gli indicatori D3 e EN2 persiste una differenza degna di nota tra gli standard error ottenuti con i due metodi. Guardando, invece, il costrutto *Cibo*, si nota che contiene un indicatore con coefficiente molto basso: *Cibo* → C2 presenta una stima pari a 0.390. Inoltre, guardando gli standard error ottenuti con PLS, si nota che risultano inferiori rispetto a quelli CB. Questo potrebbe essere un punto da approfondire in sviluppi futuri: sarebbe interessante analizzare il comportamento del metodo PLS, ad esempio, in presenza di indicatori che, secondo il CB, non risultano appartenere a nessun gruppo.

Ancora una volta i risultati confermano quanto detto finora, nonostante l'interpretazione di modelli articolati sia più complicata da analizzare, visto che le stime dei parametri di misura sono influenzate da molte relazioni strutturali.

TABELLA 3.9: Stime del modello ricorsivo ottenute con i metodi CB e PLS tramite simulazioni bootstrap.

Parametro	Stima CB	S.e. CB	Stima PLS	S.e. PLS
Servizio → S1	0.866	0.012	0.890	0.022
Servizio → S2	0.886	0.010	0.901	0.017
Servizio → S3	0.833	0.023	0.784	0.027
Servizio → S4	0.831	0.021	0.839	0.030
Design → D1	0.739	0.024	0.743	0.027
Design → D2	0.818	0.027	0.817	0.028
Design → D3	0.881	0.014	0.881	0.022
Cibo → C1	0.751	0.041	0.742	0.034
Cibo → C2	0.390	0.024	0.428	0.027
Cibo → C3	0.707	0.037	0.708	0.030
Em_pos → EP1	0.844	0.018	0.877	0.018
Em_pos → EP2	0.842	0.015	0.855	0.017
Em_pos → EP3	0.819	0.026	0.773	0.027
Em_neg → EN1	0.880	0.020	0.848	0.026
Em_neg → EN2	0.920	0.015	0.974	0.023
Em_neg → EN3	0.929	0.016	0.904	0.021
Giudizio_pos → GP1	0.899	0.013	0.901	0.015
Giudizio_pos → GP2	0.884	0.014	0.895	0.011
Giudizio_pos → GP3	0.928	0.007	0.914	0.009
Cibo → Em_pos	0.442	0.116	0.309	0.077
Design → Em_pos	0.214	0.093	0.278	0.064
Servizio → Em_pos	0.145	0.072	0.196	0.055
Cibo → Em_neg	-0.641	0.127	-0.557	0.072
Design → Em_neg	0.087	0.094	0.043	0.054
Servizio → Em_neg	-0.209	0.068	-0.236	0.051
Cibo → Giud_pos	0.149	0.076	0.196	0.061
Design → Giud_pos	0.050	0.043	0.046	0.040
Servizio → Giud_pos	0.193	0.040	0.172	0.036
Em_neg → Giud_pos	-0.127	0.047	-0.090	0.040
Em_pos → Giud_pos	0.444	0.040	0.459	0.036

3.7 Sintesi

Il presente Capitolo aveva come obiettivo quello di adattare alcuni tipi di modelli riflessivi a dati reali. Sicuramente, questi tipi di modelli sono un piccolo sottoinsieme di tutti quelli possibili; tuttavia mostrano come le stime ottenute con il metodo Partial Least Squares siano certamente influenzate dalla parte strutturale, cosa che non dovrebbe accadere vista la natura stessa dei modelli riflessivi. In particolare, il PLS fornisce stime meno accurate quando le relazioni strutturali sono deboli. Quando invece le relazioni strutturali sono forti, i due metodi CB e PLS danno stime simili. Per quanto riguarda le stime dei parametri strutturali, i due modelli forniscono risultati simili, indipendentemente dal vero valore dei parametri stessi. Potrebbe essere quindi un aiuto nel capire quando confrontare o meno i due metodi. Quando le stime dei parametri strutturali sono elevate, allora il confronto tra i due metodi è fattibile. Quando invece le stime dei parametri strutturali sono basse, il confronto è da evitare, o perlomeno bisogna essere consapevoli del fatto che i due metodi danno risultati diversi. Se i due metodi forniscono risultati diversi è meglio scegliere il metodo CB, che per sua natura è comunque più adatto per la stima dei modelli riflessivi.

Conclusione

I modelli riflessivi ipotizzano che il costrutto latente sia fattore comune tra i suoi indicatori, le cui correlazioni sono appunto spiegate esclusivamente dalla presenza del costrutto stesso. Un'eventuale correlazione tra indicatori di costrutti diversi è dovuta solamente alla relazione tra i costrutti stessi. Questo implica che il modello di misura di un costrutto possa, in linea di principio, essere stimato senza prendere in considerazione il modello strutturale. Di conseguenza, il metodo utilizzato per la stima dovrebbe tenere conto di ciò.

Sicuramente il metodo Covariance Based, utilizzando solo la varianza comune per la stima, risulta essere adatto ai modelli riflessivi. Il metodo Partial Least Squares, invece, utilizzando tutta la varianza, sia quella comune sia quella specifica, non risulta concettualmente adatto alla stima di questi modelli, nonostante venga spesso utilizzato per approssimarli.

Questa tesi ha l'obiettivo di confrontare i due metodi nella stima di alcuni modelli riflessivi. I risultati ottenuti confermano la bontà del metodo Covariance Based e mostrano che il comportamento del Partial Least Squares è influenzato dalla presenza del modello strutturale. In particolare, il metodo risulta più efficiente quando le relazioni strutturali presenti nel modello sono forti, dando in questo caso risultati simili a quelli ottenuti con il metodo Covariance Based.

Si può quindi concludere che in caso di relazioni forti tra i costrutti latenti del modello, per le quali entrambi i metodi forniscono stime simili, i risultati ottenuti con i due metodi sono confrontabili. Resta comunque essenziale ricordare che il metodo Covariance Based rimane il migliore nella stima di modelli riflessivi.

Appendice A

Grafici delle simulazioni

A.1 Grafici del Modello di regressione semplice

Come visto nel Capitolo 2 alla sezione riferita al Modello di regressione semplice, le stime con il metodo PLS sono influenzate dalla forza del parametro di struttura. In particolare, si nota che per parametri strutturali deboli, la varianza delle stime PLS è molto più elevata rispetto alla varianza delle stime CB. Questo problema svanisce per parametri strutturali con valori alti.

Si riportano in questa sezione i grafici riferiti alle stime CB e PLS per i parametri di misura $X \rightarrow x_2$ (Figura A.1), e $X \rightarrow x_3$ (Figura A.2).

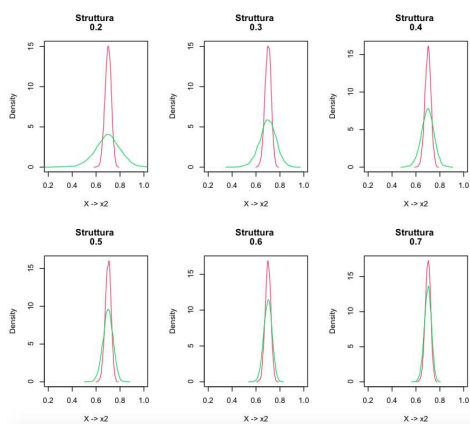


FIGURA A.1: Distribuzione delle stime del parametro di misura $X \rightarrow x_2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

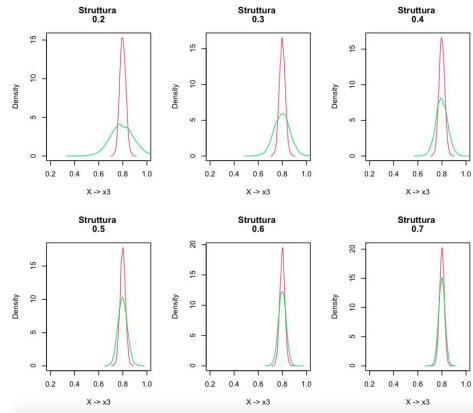


FIGURA A.2: Distribuzione delle stime del parametro di misura $X \rightarrow x3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Nel seguito i grafici riferiti alle stime CB e PLS per i parametri di misura $Y \rightarrow y2$ (Figura A.3), e $Y \rightarrow y3$ (Figura A.4).

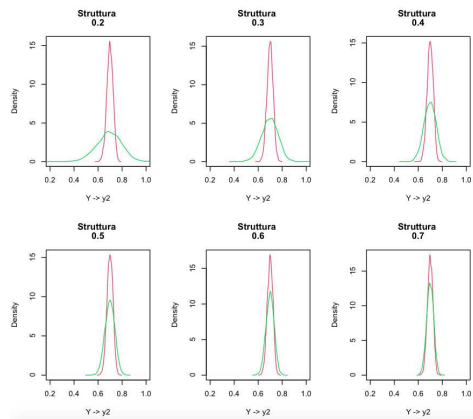


FIGURA A.3: Distribuzione delle stime del parametro di misura $Y \rightarrow y2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

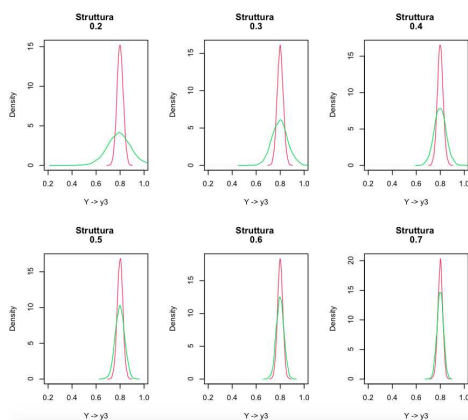


FIGURA A.4: Distribuzione delle stime del parametro di misura $Y \rightarrow y_3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

I grafici appena proposti confermano quanto descritto nel Capitolo 2. Anche per i parametri di misura riferiti agli indicatori x_2 , x_3 , y_2 e y_3 si nota che le stime ottenute con i due metodi sono centrate nei veri valori dei parametri, ma la varianza delle stime ottenuta con il metodo PLS risulta superare di molto quella ottenuta con il CB.

A.2 Grafici del Modello di regressione multivariata

Come visto nel Capitolo 2, il metodo PLS risulta fortemente influenzato dalla presenza della parte strutturale anche per il Modello di regressione multivariata.

Questa sezione integra il paragrafo sul Modello di regressione multivariata del Capitolo 2 con i grafici riferiti alle stime CB e PLS per i parametri di misura $X \rightarrow x_2$ (Figura A.5), e $X \rightarrow x_3$ (Figura A.6).

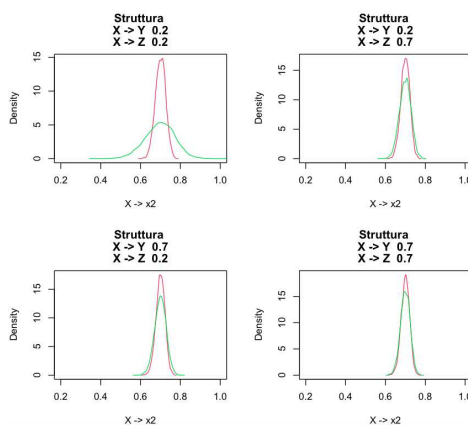


FIGURA A.5: Distribuzione delle stime del parametro di misura $X \rightarrow x_2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

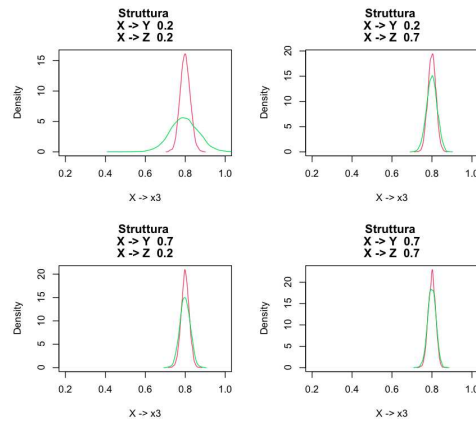


FIGURA A.6: Distribuzione delle stime del parametro di misura $X \rightarrow x_3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Nel seguito i grafici riferiti alle stime CB e PLS per i parametri di misura $Y \rightarrow y_2$ (Figura A.7), e $Y \rightarrow y_3$ (Figura A.8).

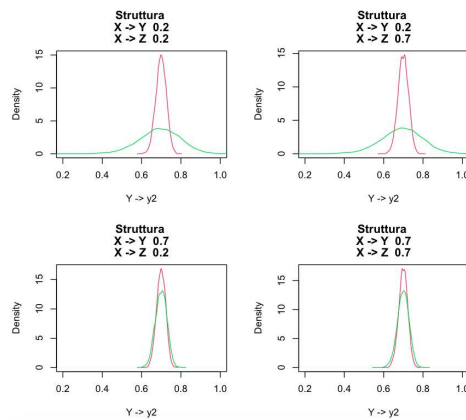


FIGURA A.7: Distribuzione delle stime del parametro di misura $Y \rightarrow y_2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

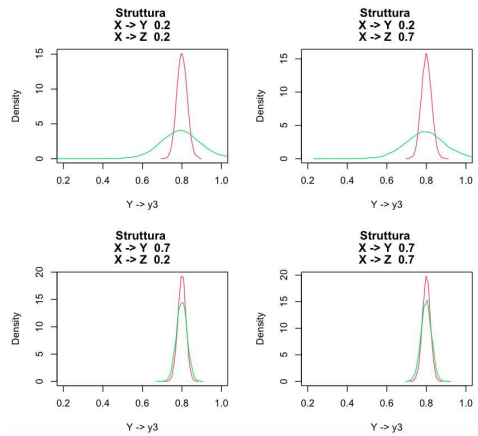


FIGURA A.8: Distribuzione delle stime del parametro di misura $Y \rightarrow y_3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Infine i grafici riferiti alle stime CB e PLS per i parametri di misura $Z \rightarrow z_2$ (Figura A.9), e $Z \rightarrow z_3$ (Figura A.10).

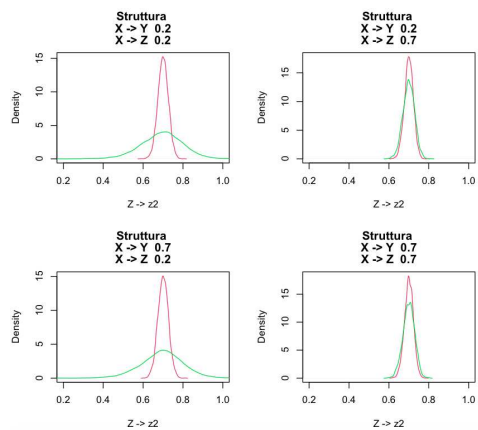


FIGURA A.9: Distribuzione delle stime del parametro di misura $Z \rightarrow z_2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

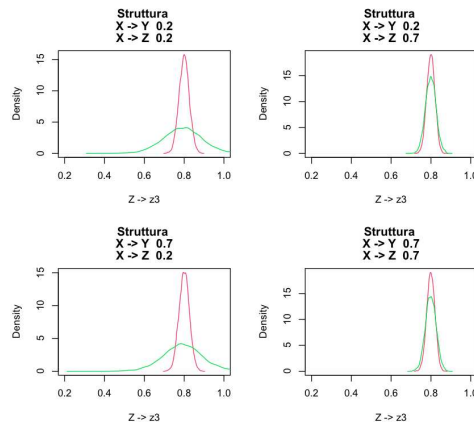


FIGURA A.10: Distribuzione delle stime del parametro di misura $Z \rightarrow z3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Le osservazioni dei paragrafi precedenti valgono anche per gli indicatori $x2$, $x3$, $y2$, $y3$, $z2$ e $z3$.

A.3 Grafici del Modello di regressione multipla

Come per i Modelli di regressione semplice e di regressione multivariata, si riportano in questa sezione dell'Appendice le distribuzioni delle stime di misura del Modello di regressione multipla per gli indicatori $x2$ e $x3$ del costrutto X. Nelle Figure A.11 e A.12 si nota ancora una volta che i due metodi assumono comportamenti diversi al crescere del parametro strutturale.

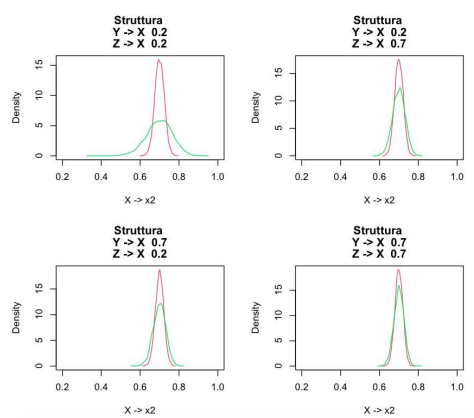


FIGURA A.11: Distribuzione delle stime del parametro di misura $X \rightarrow x2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

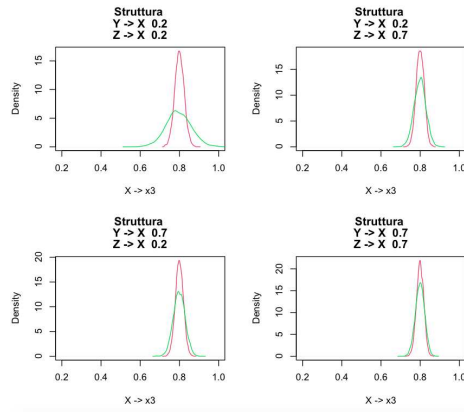


FIGURA A.12: Distribuzione delle stime del parametro di misura $X \rightarrow x3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Nel seguito i grafici riferiti alle stime CB e PLS per i parametri di misura $Y \rightarrow y2$ (Figura A.13), e $Y \rightarrow y3$ (Figura A.14).

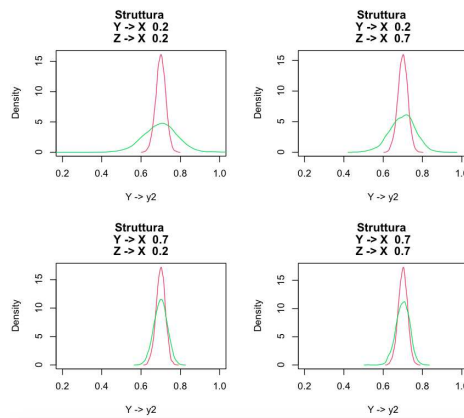


FIGURA A.13: Distribuzione delle stime del parametro di misura $Y \rightarrow y2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

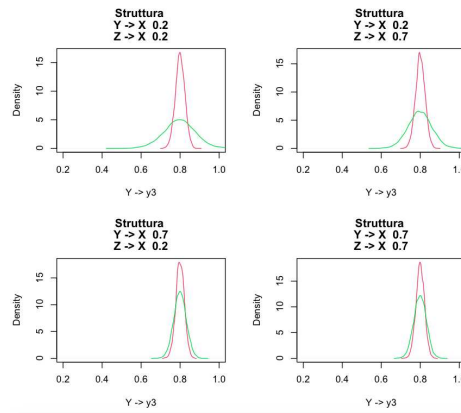


FIGURA A.14: Distribuzione delle stime del parametro di misura $Y \rightarrow y3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Infine i grafici riferiti alle stime CB e PLS per i parametri di misura $Z \rightarrow z2$ (Figura A.15), e $Z \rightarrow z3$ (Figura A.16).

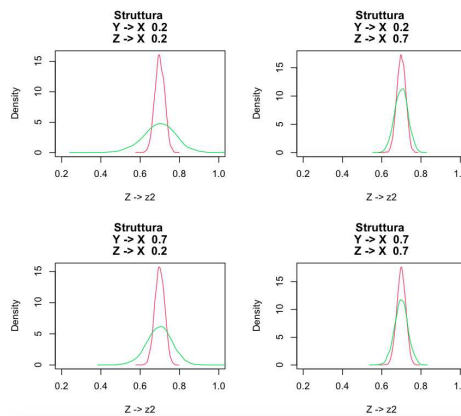


FIGURA A.15: Distribuzione delle stime del parametro di misura $Z \rightarrow z2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

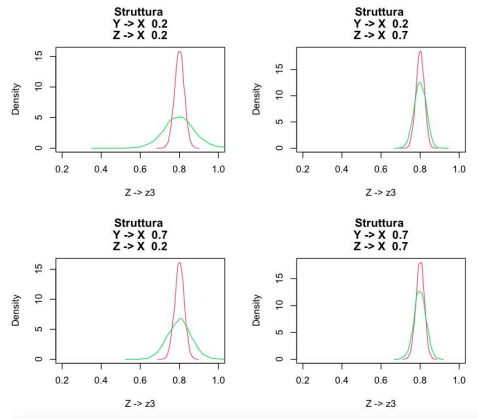


FIGURA A.16: Distribuzione delle stime del parametro di misura $Z \rightarrow z3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Nelle Figure A.17 e A.18 si riportano anche i grafici relativi alla parte strutturale del Modello di regressione multipla, notando che, anche in questo caso, i due metodi CB e PLS forniscono stime simili sia in termini di media che di varianza.

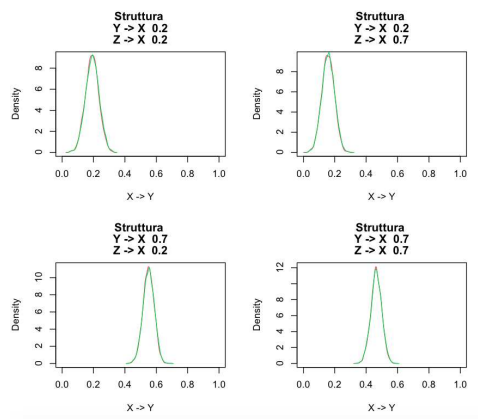


FIGURA A.17: Distribuzione delle stime del parametro strutturale $Y \rightarrow X$. In rosso le stime con CB e in verde le stime PLS

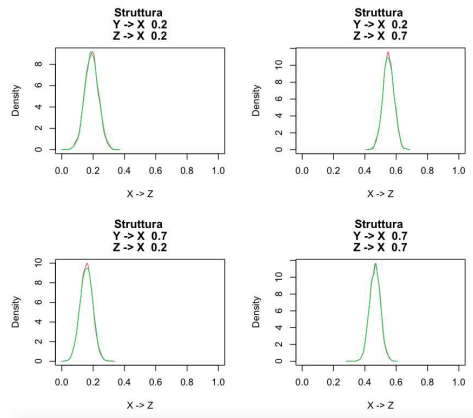


FIGURA A.18: Distribuzione delle stime del parametro strutturale $Z \rightarrow X$. In rosso le stime con CB e in verde le stime PLS

Anche per questo modello vale quanto detto, ossia che la varianza delle stime con il metodo PLS cala all'aumentare della forza nelle relazioni strutturali.

A.4 Grafici del Modello ricorsivo

In questo paragrafo si riportano i grafici del paragrafo sul Modello ricorsivo del Capitolo 2. Ancora una volta vale quanto detto nel Capitolo 2 per tutti i tipi di modelli analizzati.

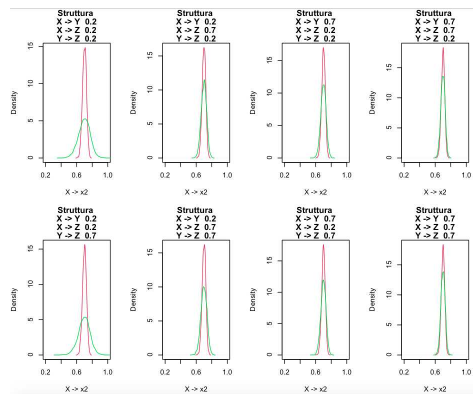


FIGURA A.19: Distribuzione delle stime del parametro di misura $X \rightarrow x2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

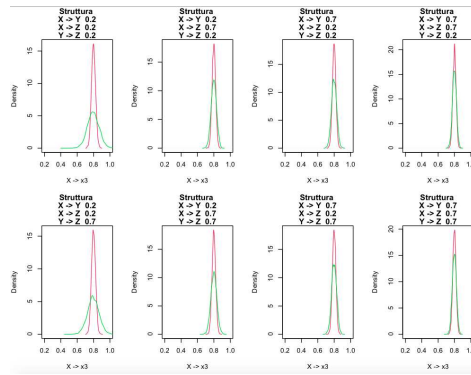


FIGURA A.20: Distribuzione delle stime del parametro di misura $X \rightarrow x_3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

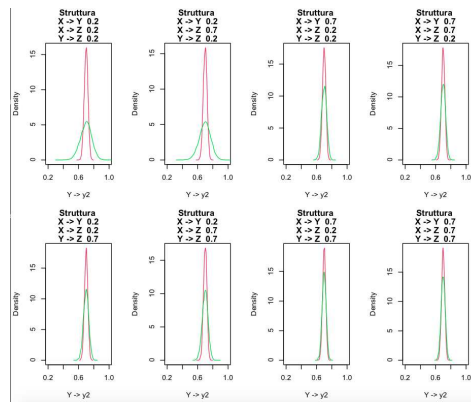


FIGURA A.21: Distribuzione delle stime del parametro di misura $Y \rightarrow y_2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

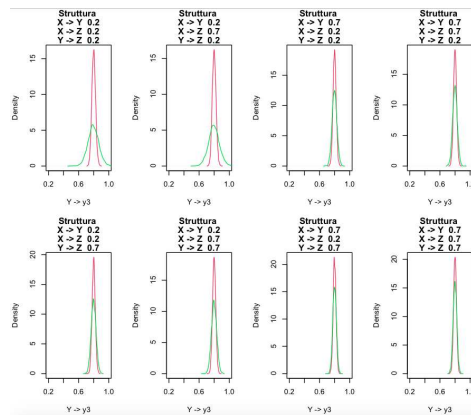


FIGURA A.22: Distribuzione delle stime del parametro di misura $Y \rightarrow y_3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

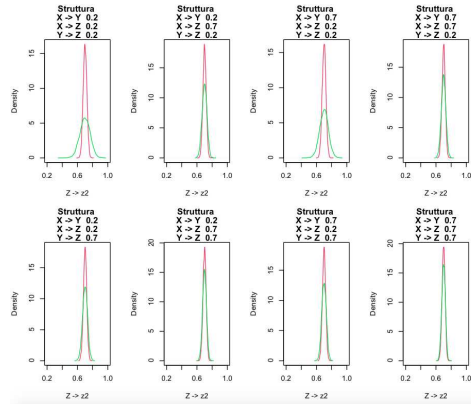


FIGURA A.23: Distribuzione delle stime del parametro di misura $Z \rightarrow z_2$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

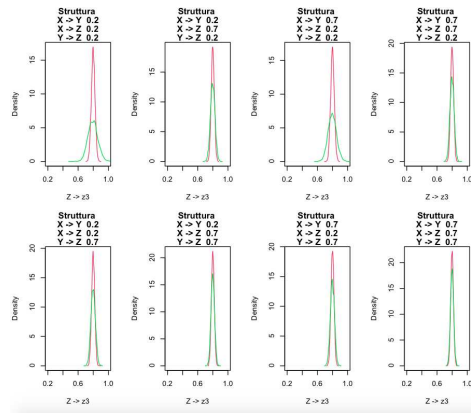


FIGURA A.24: Distribuzione delle stime del parametro di misura $Z \rightarrow z_3$ al variare del parametro strutturale. In rosso le stime con CB e in verde le stime PLS

Bibliografia

- Afthanorhan, A., Awang, Z., Aimran, N. (2020). Five common mistakes for using Partial Least Squares Path Modeling (PLS-PM) in management research. *Contemporary Management Research* **16(4)**, 255-278.
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods and Research* **17(3)**, 303-316.
- Bollen, K. A., Davies, W. R. (2009). Causal indicator models: Identification, estimation and testing. *Structural Equation Modeling* **16(3)**, 498-522.
- Diamantopoulos, A., Riefler, P. (2011). Using formative measures in international marketing models: A cautionary tale using consumer animosity as an example. *Advances in International Marketing* **10(22)**, 11-30.
- Dijkstra, T.K. and Henseler, J. (2015). Consistent partial least squares path modeling. *MIS Quarterly* **39(2)**, 297-316.
- Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. 2nd ed., SAGE, Thousand Oaks, CA.
- Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., Ray, S. (2020). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*. Springer Cham.
- Hair Jr, J. F., Matthews, L. M., Matthews, R. L., Sarstedt, M. (2017). PLS-SEM or CB-SEM: Updated guidelines on which method to use. *International Journal of Multivariate Data Analysis* **1(2)**, 107-123.
- Hair, J. F., Sarstedt, M., Ringle, C. M., Mena, J. A. (2012). An assessment of the use of Partial Least Squares Structural Equation Modelling in marketing research. *Journal of the Academy of Marketing Science* **40(3)**, 414-433.

- Hair, J.F., Sarstedt, M., Ringle, C.M., Thiele, K.O. and Gudergan, S.P. (2016). Estimation issues with PLS and CBSEM: where the bias lies! *Journal of Business Research* **69(10)**, 3998-4010.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine* **2(8)**, e124.
- Rhemtulla, M., van Bork, R., Borsboom, D. (2020). Worse than measurement error: Consequences of inappropriate latent variable measurement models. *Psychological Methods* **25(1)**, 30-45.
- Rigdon, E., Sarstedt, M., Ringle, C. M. (2017). On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations. *Marketing ZFP - Journal of Research and Management* **39**, 4-16.
- Venturini, S., Mehmetoglu, M. (2019). plssem: A Stata package for Structural Equation Modeling with Partial Least Squares. *Journal of Statistical Software* **88(8)**, 1-35.

Ringraziamenti

Ringrazio il prof. Adriano Paggiaro per avermi seguito nella stesura di questa tesi e per la sua grande disponibilità, per me molto preziosa perchè mi ha consentito di conciliare lavoro e studio.

