
Modelli di equazioni strutturali per variabili osservate: una nota sugli indici di adattamento.

Massimiliano Pastore

Università di Padova

Sommario I modelli di equazioni strutturali per sole variabili osservate, pur essendo molto utilizzati nell'ambito di alcune discipline psicologiche, non sempre vengono riconosciuti come caso particolare dei modelli con variabili latenti. Quando tali modelli vengono riconosciuti come appartenenti alla famiglia dei modelli di equazioni strutturali, gli autori tendono a riportare lo stesso tipo di indici di adattamento. In questo lavoro ci siamo occupati di valutare se questi indici siano appropriati quando nel modello mancano le variabili latenti. Abbiamo studiato il comportamento di alcuni noti indici con un esperimento Monte Carlo utilizzando quattro tipologie di modello e manipolando la quota di errore. Dall'analisi è emerso che gli indici di adattamento utilizzati generalmente sono indipendenti dalla varianza d'errore nel modello, cosa che li può rendere non particolarmente informativi. L'unico indice sensibile a tale elemento è il Coefficiente di Determinazione Totale (CD), che sembra anche poco legato alla performance degli altri indici considerati. Pertanto, nel caso di modelli di equazioni strutturali che non contemplano la presenza di variabili latenti, si suggerisce di affiancare l'uso del CD agli altri indici.

Structural Equation models with observed variables: A note on fit indices.

Abstract. Although structural equation models with observed variables are used in some psychological disciplines, they are not always recognized as a special case of models with latent variables. When these models are recognized as belonging to the family of structural equation models, the authors tend to report the same type of fit indices. In this paper, we evaluated whether these indices are appropriate when latent variables are not considered in the model. We studied the behavior of some well-known indices with a Monte Carlo experiment using four types of models and by manipulating the error. The analysis revealed that selected fit indices are generally independent of the error variance in the model, which in turn may render them not particularly informative. The only sensitive index is the Total Coefficient of Determination (CD), which does not seem to be correlated to the performance of other considered indices. Therefore, in structural equation models without latent variables, we propose to use the CD along with other indices.

Parole chiave: structural equation models, observed variables, fit indices

1 INTRODUZIONE

Quando si parla di modelli di equazioni strutturali (MES), si pensa quasi sempre a modelli che comprendono relazioni tra variabili latenti. In realtà, nella famiglia dei MES, si possono considerare a tutti gli effetti anche i modelli che non comprendono l'uso di tali variabili. Ad esempio, sia la regressione lineare (semplice o multipla) che l'Analisi della Varianza possono essere scritti come casi particolari di MES (Jöreskog e Sörbom, 1996; Kline, 2005).

La logica alla base dei MES può essere riassunta in questo modo: 1) Il ricercatore formula una (o più) ipotesi di modello che mette in relazione un'insieme di variabili (che possono essere osservate o latenti); 2) Sulla base del campione di dati produce una matrice di correlazione¹ \mathbf{S} che stima l'intensità delle relazioni tra le variabili osservate; 3) A partire da \mathbf{S} , utilizzando una apposita funzione di adattamento F (basata ad esempio su un algoritmo di massima verosimiglianza), stima i parametri del modello (che indichiamo con $\boldsymbol{\theta}$) e con essi riproduce la matrice di correlazione $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \hat{\boldsymbol{\Sigma}}$; 4) Confrontando la matrice calcolata sui dati osservati \mathbf{S} e quella riprodotta in funzione dei parametri stimati $\hat{\boldsymbol{\Sigma}}$ valuta la plausibilità del modello. Sembra infatti ragionevole che se il modello è adattabile ai dati osservati i suoi parametri debbano esprimere una struttura relazionale simile a quella osservata e pertanto le due matrici di correlazione, osservata e riprodotta dovrebbero essere simili, formalmente: $\mathbf{S} \simeq \hat{\boldsymbol{\Sigma}}$. A questo punto, per valutare la bontà del modello, sono disponibili i cosiddetti indici di adattamento (IA; per una rassegna si veda Tanaka, 1993) alcuni dei quali abbiamo preso in considerazione in questo lavoro e su cui ci soffermeremo nella sezione seguente².

In una rassegna di lavori nel contesto delle scienze della comunicazione, Holbert e Stephenson (2002), hanno individuato che il 57.6% dei lavori con approccio MES, impiegavano solo variabili osservate. Di converso, MacCallum e Austin (2000), analizzando gli articoli in cui si faceva uso di MES in 16 riviste di psicologia tra il 1993 ed il 1997, hanno rilevato solo il 25% di lavori senza variabili latenti. Sembrerebbe però che molti ricercatori non includano questi modelli nella famiglia dei MES, infatti, altre rassegne quali quella di Breckler (1990) nella psicologia e quella di Baumgartner e Homburg (1996) nell'ambito del marketing, hanno considerato come MES solo i casi comprendenti variabili latenti.

Consultando tutti i lavori usciti sulla rivista *Structural Equation Modeling* tra il 1999 ed il 2009, abbiamo rilevato solo 11 articoli che prendevano in considerazione modelli senza variabili latenti, cioè circa il 4% dei lavori pubblicati.

In tabella 1 abbiamo riportato il numero di lavori che utilizzano MES per cinque riviste internazionali, considerate nel periodo 2008-2009. Per semplicità abbiamo considerato solo i casi in cui si faceva espressamente riferimento ai MES oppure quando la tipologia di modello era chiaramente analizzabile con tale metodologia. Dalla lettura della tabella possiamo notare come i lavori che utilizzano solo variabili osservate siano piuttosto rari nelle due riviste di taglio metodologico (*Psychological Methods* e *Multivariate Behavioral Research*) e siano invece piuttosto frequenti nelle riviste di tipo applicativo.

Per quanto riguarda la realtà italiana, abbiamo considerato le annate 2008 e 2009 di tre riviste: *Giornale Italiano di Psicologia*, *Psicologia Sociale* e *Psicologia Clinica dello Sviluppo*. Su 22 lavori che utilizzavano MES, 7 di essi prendevano in considerazione solo variabili osservate (Kastlunger, Martini, Kirchler e Hofmann, 2008; Pierantoni e Prati, 2008; Vieno, Santinello e Pastore, 2008; Affuso e Bacchini, 2009; Chieco, Vieno e Altoè, 2009; Curcuruto, Mariani e Lippert, 2009; Manzi, Brambilla, Regalia e Voci, 2009).

Insomma, sembrerebbe quasi che l'inclusione o meno dei modelli con variabili osservate nella famiglia dei MES sia legata all'ambito disciplinare (Stephenson e Holbert, 2003) e l'impressione che si ricava dalla lettura dei lavori considerati è che tale inclusione o meno comporti differenze

¹ In questo momento intendiamo il concetto di correlazione in senso ampio, ossia come un qualunque tipo di misura dell'associazione tra variabili. Per coerenza con la simbologia che adotteremo in seguito utilizziamo la lettera \mathbf{S} .

² Come opportunamente suggerito da un anonimo referee, va ricordato che nel processo di applicazione dei MES è prevista anche una ulteriore fase di confronto della bontà del modello analizzato con altri modelli alternativi teoricamente plausibili.

	con latenti	senza latenti	MVO/MES
<i>Psychological Methods</i>	6	2	25%
<i>Multivariate Behavioral Research</i>	11	2	15%
<i>European Journal of Social Psychology</i>	12	9	43%
<i>Journal of Experimental Social Psychology</i>	5	14	74%
<i>Child Development</i>	23	16	41%

Tabella 1. Frequenza dei lavori che utilizzano MES con variabili latenti (prima colonna) o senza (seconda colonna) nelle annate 2008-2009 nelle cinque riviste considerate. Sono stati contati solo i lavori che facevano riferimento esplicito ai MES oppure che utilizzavano modelli direttamente assimilabili, escludendo lavori con Modelli *Multilevel* o di *Latent growth analysis*. Nella terza colonna è calcolata la percentuale di lavori che utilizzano MVO sul totale dei lavori considerati.

sul modo di trattare tali modelli e conseguentemente sull'uso dei relativi IA. In particolare, nei lavori di tipo empirico, quando il modello senza latenti è riconosciuto come caso di MES, gli autori tendono a riportare lo stesso tipo di IA utilizzati nei modelli con variabili latenti (si veda, ad esempio Brody, Beach, Philibert, Chen e McBride Murry 2009; Forget-Dubois, Dionne, Lemelin, Pérusse, Tremblay e Boivin, 2009; Hareli, Harush, Suleiman, Cossette, Bergeron, Lavoie, Dugay e Hess, 2009; Hodson, Choma, e Costello, 2009; Knowles, Lowery e Schaumberg, 2009; Newheiser, Tausch, Dovidio e Hewstone, 2009; Okimoto, 2009; Seaton, Yip e Sellers 2009; Wojciszke, Abele e Baryla, 2009; Yau, Tasopoulos-Chan e Smetana, 2009).

Nel presente lavoro, ci siamo interessati all'utilizzo degli IA nei modelli che coinvolgono solo variabili osservate, ed abbiamo cercato di capire se tali indici permettano effettivamente di trarre informazioni adeguate sulla bontà del modello teorico ipotizzato; in particolare, abbiamo voluto analizzare il comportamento di tali IA in relazione all'aumento di errore del modello.

Come primo passo presenteremo in dettaglio i modelli per variabili osservate e descriveremo alcuni dei principali IA utilizzati nel contesto dei MES. Successivamente saranno descritti i modelli scelti ed i dettagli della simulazione Monte Carlo. Infine saranno analizzati i risultati dell'esperimento con riferimento agli IA considerati.

2 MODELLI PER VARIABILI OSSERVATE

In generale, con i modelli per variabili osservate (MVO) si analizzano due tipi di variabili: variabili *single-item* o aggregati di insiemi di item (Stephenson e Holbert, 2003). Già nei lavori di Sewell Wright (1921, 1934) possiamo trovare i primi casi di MVO, esempi di quella che verrà poi definita come *path-analysis* (Blalock, 1971; Pedhazur, 1982). Più recentemente, Wolfle (2003) presenta una bibliografia dettagliata relativa all'introduzione della *path-analysis* nelle scienze sociali.

Nei MVO si assume che le variabili siano misurate direttamente e senza errore. Qualora ciò non fosse vero, risulta evidente che si possono avere delle ripercussioni sulle stime dei parametri del modello (MacCallum, 1995). L'uso più appropriato di tali modelli si avrebbe nel caso di variabili *single-item* oppure aggregati di misure la cui attendibilità sia molto elevata (Hoyle e Kenny, 1999). Se si vuole prendere in considerazione l'errore di misura, esso è ammesso per le sole variabili endogene. In questo modo, si assume che le variabili osservate y e x , coincidano con le corrispondenti latenti η e ξ (Bollen, 1989).

La forma generale di questa classe di modelli è data dall'equazione strutturale

$$y = \alpha + By + \Gamma x + \zeta \tag{1}$$

in cui:

- y indica il vettore delle variabili endogene
- x indica il vettore delle variabili esogene
- ζ indica il vettore degli errori nelle equazioni

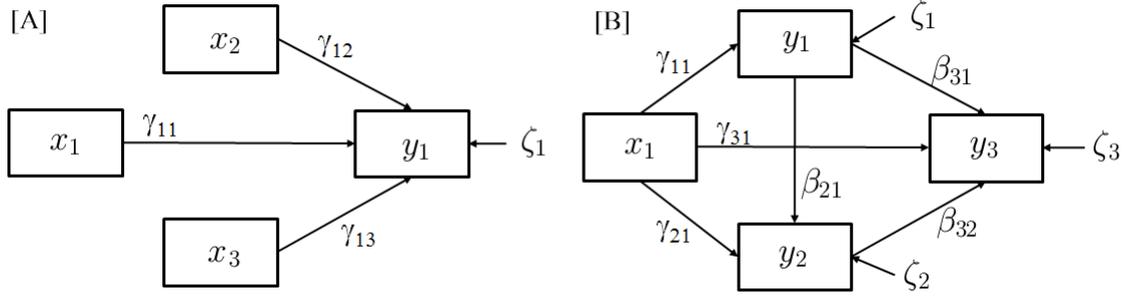


Figura 1. Esempio di due modelli con quattro variabili osservate: [A] Regressione multipla con tre variabili esogene (x_1 , x_2 e x_3) ed una variabile endogena (y_1); [B] Modello strutturale con una variabile esogena (x_1) e tre endogene (y_1 , y_2 e y_3).

- \mathbf{B} indica la matrice dei legami tra le variabili endogene
- $\mathbf{\Gamma}$ indica la matrice dei legami tra le variabili esogene e le endogene

Il vettore α contiene le intercette delle equazioni, ma può essere omesso qualora si analizzino direttamente la matrice di correlazione o di covarianza.

Oltre a \mathbf{B} e $\mathbf{\Gamma}$, vi sono altre due matrici di parametri coinvolte in questo tipo di modelli:

- $\mathbf{\Psi} = Cov(\zeta)$, matrice di covarianza tra gli errori
- $\mathbf{\Phi} = Cov(\mathbf{x})$

A proposito di quest'ultima matrice vanno fatte alcune considerazioni importanti: $\mathbf{\Phi}$ si assume essere una matrice di parametri liberi anche se, qualunque metodo usato per stimarla, produce sempre come risultato i valori della matrice di covarianza osservata $\mathbf{S}_{\mathbf{xx}} = Cov(\mathbf{x})$ (Jöreskog e Sörbom, 1996). Di conseguenza, le varianze e le covarianze delle variabili esogene non sono veri e propri parametri del modello, ma vengono ugualmente contati come tali.

La regressione semplice, multipla o multivariata sono tutti casi particolari derivabili dall'equazione (1). Supponiamo, per esempio, di avere a disposizione quattro variabili osservate; specificando le matrici \mathbf{B} e $\mathbf{\Gamma}$ nel seguente modo:

$$\mathbf{B} = (0) \quad \mathbf{\Gamma} = (\gamma_{11} \ \gamma_{12} \ \gamma_{13})$$

otteniamo il modello di regressione multipla in figura 1A, in cui ci sono tre predittori (x_1 , x_2 e x_3) ed una variabile dipendente (y_1). In questo tipo di modello abbiamo sempre una distinzione netta tra predittori e variabili dipendenti, e la matrice \mathbf{B} è sempre nulla. Ci interessa stimare solo la matrice $\mathbf{\Gamma}$, espressione delle relazioni tra gli uni e le altre (indicati in figura 1 con le frecce).

Se volessimo considerare anche delle relazioni tra predittori, dovremmo ridefinire \mathbf{B} e $\mathbf{\Gamma}$, ad esempio nel seguente modo:

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{pmatrix} \quad \mathbf{\Gamma} = \begin{pmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \end{pmatrix}$$

ottenendo un approccio secondo un modello strutturale (figura 1B), tramite il quale è possibile una scelta più vasta di relazioni possibili. In pratica, l'unico limite consiste nel numero massimo di parametri utilizzabili per specificare il modello, dato da $p(p+1)/2$ in cui p è il numero di variabili osservate.

I modelli in cui la matrice \mathbf{B} può essere scritta in forma triangolare e la matrice $\mathbf{\Psi}$ è diagonale si definiscono modelli *ricorsivi*. Questo implica che le relazioni tra variabili endogene sono unidirezionali e che non vi sono correlazioni tra gli errori. I modelli in cui una o entrambe queste

condizioni vengono meno si definiscono modelli *non-ricorsivi* (Bollen, 1989; Kline, 2005), due esempi dei quali sono rappresentati dai modelli [C] e [D] in figura 2.

Oltre alla possibilità di costruire modelli di relazione più complessi rispetto alla semplice regressione, l'approccio MES consente di avere a disposizione una serie di IA del modello ai dati. Tali indici, molto numerosi, possono essere classificati secondo vari criteri a seconda delle caratteristiche prese in considerazione. Ad esempio possiamo distinguere tra indici che stimano un parametro della popolazione (*population based*) o che descrivono l'adattamento di un campione osservato (*sample based*); oppure distinguere tra indici che misurano la semplicità o al contrario la complessità di un modello; indici normati, i cui valori cadono all'interno di un range definito, e indici non-normati; indici relativi, ossia che utilizzano un modello come riferimento, e indici assoluti (si veda Tanaka, 1993). Obiettivo generale di tutti gli indici è comunque di poter stabilire se il modello definito dal ricercatore si adatta ai dati campionari osservati.

In questo lavoro abbiamo preso in considerazione i seguenti otto indici, sette dei quali risultano tra i più noti ed utilizzati nella letteratura relativa ai MES:

χ^2 . Si ottiene moltiplicando per $(n - 1)$, in cui n è la numerosità campionaria, il valore minimo ottenuto nella funzione di adattamento F per il modello scelto. In generale, minore risulta essere tale valore, migliore risulta il modello; va però tenuto presente che si tratta di una misura molto sensibile alla numerosità campionaria e alla violazione degli assunti e pertanto poco utile come misura di per sè. La ragione per cui lo riportiamo è che molti indici di adattamento vengono calcolati sulla base di tale valore e pertanto ci sarà utile nell'interpretazione dei risultati.

Goodness-of-Fit Index (GFI; Jöreskog e Sörbom, 1984). Tale indice è costruito in modo da assumere valori nell'intervallo 0-1 e può essere interpretato, in maniera simile al tradizionale R^2 , come l'ammontare di variabilità della matrice di covarianza tra le variabili del modello (\mathbf{S}) predetta dalla matrice di covarianza della popolazione stimata ($\hat{\Sigma}$). Pertanto, più tale indice si avvicina ad uno, migliore sarà il modello.

Adjusted Goodness-of-Fit Index (AGFI; Jöreskog e Sörbom, 1984; Tanaka e Huba, 1985). Si tratta di un aggiustamento del GFI che tiene conto dei gradi di libertà del modello, e come l'indice precedente raggiunge il valore massimo di 1 quando $\mathbf{S} = \hat{\Sigma}$.

Comparative Fit Index (CFI; Bentler, 1989, 1990). Tale indice misura la differenza tra il modello in esame (*target model*) e il modello nullo (*null model*) ossia il modello in cui le variabili osservate sono tutte tra loro indipendenti. Pertanto il suo valore varia tra 0, quando le variabili osservate hanno correlazione prossima a zero e 1, quando il valore di χ^2 relativo al modello è inferiore ai suoi gradi di libertà (formalmente $\chi^2_M < gl_M$).

Nonnormed Fit Index (NNFI o TLI; Tucker e Lewis, 1973; Bentler e Bonnett, 1980). Anche questo è un indice comparativo che misura la differenza tra modello *target* e modello nullo. A differenza di CFI, il suo valore può uscire dal range 0-1.

Root-Mean-Square Error of Approximation (RMSEA; Steiger e Lind, 1980, Steiger, 1989). Come i due indici precedenti, utilizza la differenza tra il χ^2 relativo al modello ed i relativi gradi di libertà (formalmente $\chi^2_M - gl_M$) ed aggiusta tale differenza tenendo conto anche della numerosità campionaria. Quando vale 0 indica il migliore adattamento, valori più alti indicano un adattamento peggiore fornendo una misura dell'errore di approssimazione dei dati al modello ipotizzato.

Standardized Root Mean Square Residual (SRMR, Jöreskog e Sörbom, 1984). È un indice che riassume la differenza tra la matrice di covarianza osservata (\mathbf{S}) e la matrice riprodotta sulla base dei parametri del modello ($\hat{\Sigma}$) espressa in termini di residui. In un modello ideale tali residui dovrebbero essere tutti nulli, in un buon modello quindi ci aspettiamo che tale indice sia prossimo a 0. Al crescere dei residui, tale indice cresce, ad indicare un peggioramento nel modello.

Coefficiente di determinazione totale (CD, Jöreskog e Sörbom, 1984). Si tratta di un indice di variabilità spiegata, affine a R^2 , ottenuto come:

$$1 - \frac{|\hat{\Psi}|}{|\hat{\Sigma}_{yy}|}$$

in cui $|\hat{\Sigma}_{yy}|$ è il determinante della matrice di covarianza tra le variabili endogene riprodotta e $|\hat{\Psi}|$ il determinante della matrice di covarianza tra gli errori (Bollen, 1989, p. 118). Diversamente dai tradizionali R^2 , che esprimono la quota di varianza spiegata da un insieme di predittori rispetto ad una sola variabile dipendente, CD fa riferimento alla varianza spiegata da tutto il modello. Il suo valore massimo è 1 quando il modello è senza errore, diminuisce fino a 0 quando la varianza degli errori risulta essere massima. In altri termini, mentre gli R^2 saranno tanti quante le variabili endogene del modello, CD è un unico valore che riassume la proporzione di variabilità spiegata dalle relazioni definite complessivamente nel modello in esame.

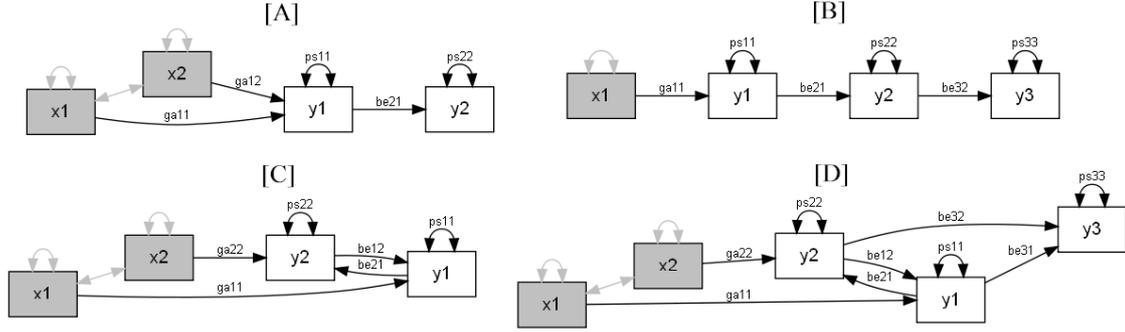


Figura 2. Rappresentazione dei modelli simulati. I rettangoli grigi indicano le variabili esogene, i rettangoli bianchi le variabili endogene. Le frecce nere indicano i parametri del modello, le frecce grige indicano le varianze e le covarianze relative alle variabili esogene. **[A]** Due variabili esogene (x_1 , x_2) e due variabili endogene (y_1 , y_2) con tre legami (γ_{11} , γ_{12} e β_{21}); **[B]** Una variabile esogena (x_1) e tre endogene (y_1 , y_2 , y_3) con tre legami (γ_{11} , β_{21} e β_{32}); **[C]** Due variabili esogene (x_1 , x_2) e due variabili endogene (y_1 , y_2) con quattro legami (γ_{11} , γ_{22} , β_{21} e β_{12}); **[D]** Due variabili esogene (x_1 , x_2) e tre variabili endogene (y_1 , y_2 , y_3) con sei legami (γ_{11} , γ_{22} , β_{21} , β_{12} , β_{31} e β_{32}).

3 SIMULAZIONE

Il programma per la simulazione Monte Carlo è stato scritto in ambiente R (R Development Core Team, 2009). Per valutare il comportamento degli IA presi in considerazione abbiamo simulato i quattro modelli rappresentati in figura 2. I modelli sono stati scelti sulla base delle tipologie incontrate nella letteratura in modo da essere sufficientemente rappresentativi ed al tempo stesso non eccessivamente complessi; i primi due di tipo sono ricorsivo e gli altri di tipo non-ricorsivo.

Il modello **[A]** presenta due variabili esogene (x_1 e x_2) e due endogene (y_1 e y_2). Le variabili esogene hanno un legame diretto con y_1 la quale risulta essere predittore diretto di y_2 . Avremo pertanto 5 parametri: tre dei quali definiscono le relazioni strutturali (γ_{11} , γ_{12} e β_{21}) e due definiscono le varianze degli errori (ψ_{11} e ψ_{22}).

Il modello **[B]** è costituito da una variabile esogena (x_1) e tre endogene (y_1 , y_2 e y_3). In questo modello i parametri sono 6: tre di essi che definiscono le relazioni strutturali (γ_{11} , β_{21} e β_{32}) e tre varianze degli errori (ψ_{11} , ψ_{22} e ψ_{33}).

Il modello **[C]** ha le stesse variabili del modello **[A]** ma con una struttura relazionale diversa definita da 6 parametri: quattro relativi alle relazioni strutturali (γ_{11} , γ_{22} , β_{21} e β_{12}) e due varianze di errori (ψ_{11} e ψ_{22}).

Il modello **[D]** è identico al modello **[C]** con l'aggiunta di una variabile endogena e conseguentemente tre parametri in più (β_{32} , β_{31} e ψ_{33}).

Ogni modello è stato definito in base ai valori dei parametri strutturali, rispettivamente 0.6 i γ e 0.4 i β , scelti in modo da rappresentare degli effetti di grandezza intermedia. Le varianze d'errore

(ossia i parametri ψ) sono state modificate da 0.1 a 1 con incrementi di 0.1; questa variazione traduce di fatto l'effetto di incremento dell'errore e quindi un sostanziale peggioramento nella bontà del modello. I dati sono stati generati con distribuzione normale multivariata.

Per ciascuno dei 10 livelli di errore e ciascun modello abbiamo generato 5000 campioni di numerosità $n = 100$ stimandone i parametri e calcolando gli indici di adattamento (utilizzando la libreria `sem`; Fox, Kramer e Friendly, 2010). In totale abbiamo prodotto $5000 \times 10 \times 4 = 200000$ campioni di dati.

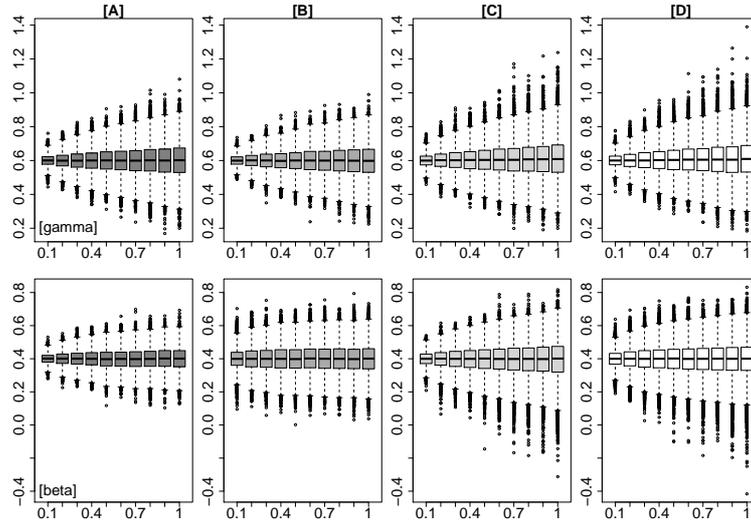


Figura 3. Rappresentazione tramite boxplot (o a scatola e baffi) della distribuzione dei valori di γ (grafici superiori) e β (grafici inferiori) ottenuti sui quattro modelli in funzione della varianza di errore (valori dei parametri ψ). Dato un valore di ascissa, ciascuna scatola permette di identificare mediana (linea interna alla scatola), primo e terzo quartile (rispettivamente il lato inferiore e quello superiore della scatola) della relativa distribuzione.

4 RISULTATI

Per tutti i 200000 campioni di dati generati l’algoritmo di massima verosimiglianza per la stima dei parametri ha raggiunto la convergenza. In nessun caso sono state ottenute soluzioni improprie, ossia casi in cui le varianze stimate hanno valori negativi (*Heywood case*) e pertanto tutte le soluzioni sono state accettate ed analizzate.

Per valutare i risultati ottenuti abbiamo prima di tutto verificato che nei campioni di dati generati i valori dei parametri fossero coerenti con quanto da noi definito. Ricordiamo che per la generazione dei dati i parametri γ erano impostati al valore 0.6 e i parametri β al valore 0.4.

In figura 3 abbiamo rappresentato con una serie di boxplot (o grafico a scatola e baffi) le distribuzioni dei valori ottenuti nelle 5000 replicazioni per ciascuno dei quattro modelli considerati relativamente ai parametri γ e β , rispettivamente nella parte superiore ed inferiore.

Dall’analisi di questi grafici emergono tre aspetti importanti. Primo, la mediana di tutte le distribuzioni risulta essere sempre il valore definito del parametro per la simulazione. Lo stesso vale per la media, non direttamente rappresentata nei grafici. Secondo, il valore dei parametri è indipendente dal modello e dal valore delle varianze d’errore. Detto in altri termini, il valore atteso dei parametri strutturali dei modelli considerati non viene influenzato dalla quota di errore imposta al modello attraverso la manipolazione dei parametri ψ . Terzo, l’aumento nella varianza d’errore comporta un aumento nella variabilità delle distribuzioni relative ai parametri e conseguentemente

nel loro errore standard. Per entrambi i tipi di parametri la varianza totale ottenuta è pari a 0.007; nei γ la varianza cresce da 0.001 con $\psi = 0.1$ a 0.013 con $\psi = 1$, mentre nei β si passa da 0.002 a 0.011.

4.1 Chi quadrato

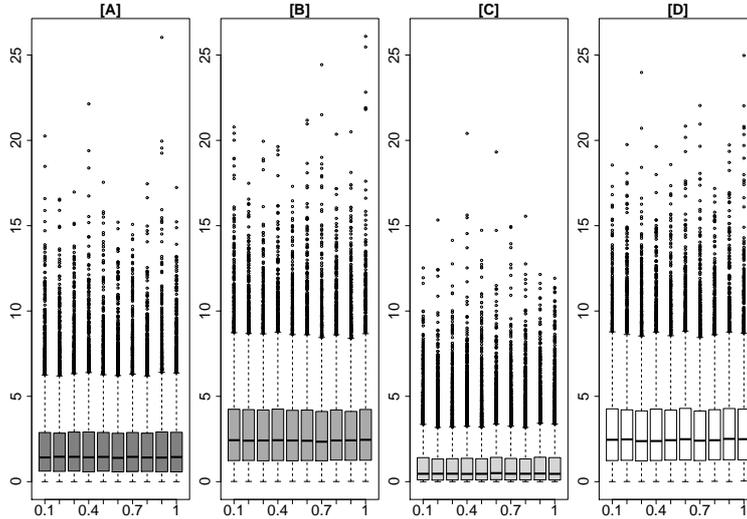


Figura 4. Rappresentazione tramite boxplot della distribuzione dei valori di χ^2 ottenuti sui quattro modelli in funzione della varianza di errore (valori dei parametri ψ). Dato un valore di ascissa, ciascuna scatola permette di identificare mediana (linea interna alla scatola), primo e terzo quartile (rispettivamente il lato inferiore e quello superiore della scatola) della relativa distribuzione.

In figura 4 sono rappresentate le distribuzioni dei valori di χ^2 ottenuti sui quattro modelli in funzione dei valori assegnati alla varianza d'errore (valori dei parametri ψ). Ricordiamo che tale indice viene calcolato sulla base del valore minimo ottenuto nella funzione di verosimiglianza che stima i parametri del modello.

Appare evidente che i valori dei χ^2 sono indipendenti dal livello di errore del modello. In particolare osserviamo che i valori delle mediane rimangono costanti per tutti i livelli di varianza d'errore in ascissa: per il modello **[A]** i valori ottenuti si collocano tra 1.39 e 1.45, per il **[B]** tra 2.34 e 2.45, per il **[C]** tra 0.45 e 0.50 e per il modello **[D]** tra 2.38 e 2.30. In più, le medie ottenute si collocano sempre molto vicine ai valori attesi della distribuzione, ossia con i gradi di libertà del modello, rispettivamente 2 per il modello **[A]** (media osservata 2.07), 3 per il **[B]** (media osservata 3.07), 1 per il **[C]** (media osservata 1.03) e 3 per il modello **[D]** (media osservata 3.09).

In relazione alla variabilità non si ravvisano particolari differenze nei vari livelli di ψ : nel modello **[A]** le varianze relative alle distribuzioni dei χ^2 ottenuti risultano comprese tra 4.02 e 4.46, nel modello **[B]** tra 6.08 e 6.67, nel **[C]** tra 1.94 e 2.31 e per il modello **[D]** tra 6.03 e 6.66.

Considerando che un modello è accettabile quando il valore di χ^2 risulta essere non statisticamente significativo, abbiamo ottenuto in media circa il 90% di modelli accettati, rispettivamente 94% per il modello **[A]**, 86% per il **[B]**, 98% per il **[C]** e 86% per il **[D]**.

4.2 Indici assoluti

In figura 5 sono rappresentate le distribuzioni dei valori di GFI e AGFI, rispettivamente nella parte superiore ed inferiore, in funzione della varianza d'errore. Anche nel caso di questi due indici osserviamo una totale invarianza del risultato, indipendente dal livello di errore e dal modello.

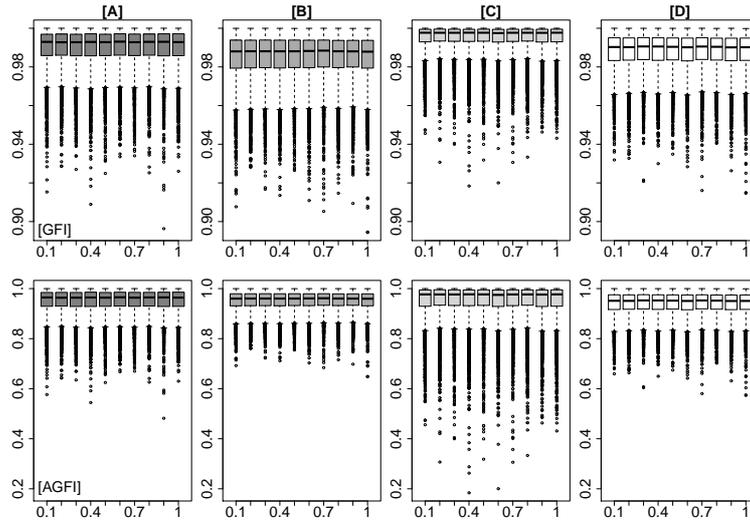


Figura 5. Rappresentazione tramite boxplot della distribuzione dei valori di GFI (grafici superiori) e AGFI (grafici inferiori) ottenuti sui quattro modelli in funzione della varianza di errore (valori dei parametri ψ). Dato un valore di ascissa, ciascuna scatola permette di identificare mediana (linea interna alla scatola), primo e terzo quartile (rispettivamente il lato inferiore e quello superiore della scatola) della relativa distribuzione. L'intervallo di rappresentazione dei valori in ordinata è diverso in quanto diversa la variabilità nei due indici.

L'unica nota di rilievo è data dalla maggiore variabilità presente nell'AGFI, peraltro non legata all'aumento dell'errore. A livello complessivo la varianza relativa all'AGFI è pari a 0.0028, cioè 28 volte quella del GFI, risultata pari a 0.0001. Per mettere in evidenza tale differenza nei due indici, abbiamo adottato intervalli diversi di rappresentazione in ordinata. In pratica i valori di GFI variano tra un minimo di 0.89 (ottenuto nel modello [B] con $\psi = 1$) al valore massimo di 1, mentre per AGFI vale lo stesso valore massimo, ma il minimo valore ottenuto è 0.18 (ottenuto nel modello [C] con $\psi = 0.4$).

Considerando per entrambi gli indici il valore di 0.96 come soglia per l'accettazione di un modello, abbiamo ottenuto complessivamente il 98% dei casi con valori di GFI maggiori o uguali a tale soglia ed il 52% di casi relativamente ad AGFI. Rispetto ai modelli considerati, utilizzando la soglia per GFI abbiamo ottenuto il 98% per [A], il 96% per [B], il 99.6% per [C] e 98% per [D]. Utilizzando AGFI invece il 54% nel modello [A], il 50% nel [B], il 62% nel [C] ed il 42% nel [D].

4.3 Indici incrementali

In figura 6 sono rappresentate le distribuzioni dei valori di CFI e NNFI, rispettivamente nella parte superiore ed inferiore, in funzione della varianza d'errore. Data la evidente differenza tra i due indici in termini di variabilità abbiamo utilizzato due intervalli diversi di rappresentazione in ordinata. La varianza del NNFI valutata complessivamente è risultata circa 10 volte superiore a quella del CFI (rispettivamente 0.0021 e 0.0002). Bisogna però tenere conto che CFI non può mai superare il valore 1 mentre NNFI sì; quest'ultimo indice produce il suo valore massimo (1.28) nel modello [B] con $\psi = 1$. Relativamente al valore minimo abbiamo invece rilevato 0.70 per il CFI (ottenuto nel modello [B] con $\psi = 0.9$) e 0.39 per il NNFI (ottenuto nel modello [C] con $\psi = 0.9$).

L'aspetto comunque rilevante è che, anche nel caso di questi due indici, osserviamo una sostanziale invarianza dei valori mediani e medi, indipendenti dal livello di errore nel modello. Differentemente dagli indici precedenti però, la variabilità del CFI e del NNFI risulta essere sensibile all'incremento delle varianze di errore.

Adottando anche per questi due indici la soglia di 0.96 come valore per l'accettazione del modello (Hu e Bentler, 1999), abbiamo ottenuto il 97% di valori di CFI e l'87% di valori di NNFI maggiori o uguali a tale soglia. CFI ha raggiunto la soglia nel 97% dei casi nel modello [A], il 92% nel [B], il 99.5% nel [C] ed il 99.6% nel [D]. NNFI invece ha raggiunto la soglia nell'86% dei casi in [A], 83% in [B], 87% in [C] e 91% in [D]

4.4 RMSEA e SRMR

In figura 7 sono rappresentate le distribuzioni dei valori di RMSEA e SRMR, rispettivamente nella parte superiore ed inferiore, in funzione della varianza d'errore. Anche in questo caso gli intervalli rappresentati in ordinata sono diversi ed in relazione agli intervalli di valori ottenuti di ciascun indice.

In relazione al primo dei due indici osserviamo un comportamento simile agli indici già presentati sino a qui, ossia la totale invarianza dall'errore. Le mediane del RMSEA si collocano infatti sempre a zero, il valore minimo ottenibile; le medie sono all'incirca uguali nei quattro modelli, intorno al valore di 0.03. Le varianze delle distribuzioni relative ai valori ottenuti di RMSEA sono leggermente diverse tra i quattro modelli ma costanti in funzione dei valori di ψ : 0.003 nel modello [A], 0.002 nei modelli [B] e [D], 0.004 nel modello [C].

Considerando 0.05 come valore soglia per l'accettabilità del modello (Hu e Bentler, 1999) sono stati ottenuti complessivamente circa il 71% di casi accettabili, rispettivamente 70% in [A], [B] e [D], 73% in [C].

Relativamente al secondo indice osserviamo invece una caratteristica nuova rispetto a tutti gli altri indici considerati, ossia che aumentando la varianza d'errore SRMR tende a crescere indicando un peggioramento dell'adattamento. Nel modello [A] la mediana della distribuzione dei valori di SRMR passa da 0.014, quando $\psi = 0.1$, a 0.027, con $\psi = 1$; nel modello [B] da 0.026 a 0.041, nel modello [C] da 0.007 a 0.018, nel modello [D] da 0.010 a 0.027. La stessa crescita si osserva anche sui valori medi, rispettivamente da 0.015 a 0.029 nel modello [A], da 0.030 a 0.043 nel modello [B], da 0.008 a 0.021 nel [C] e da 0.011 a 0.029 nel modello [D]. Anche le varianze delle distribuzioni

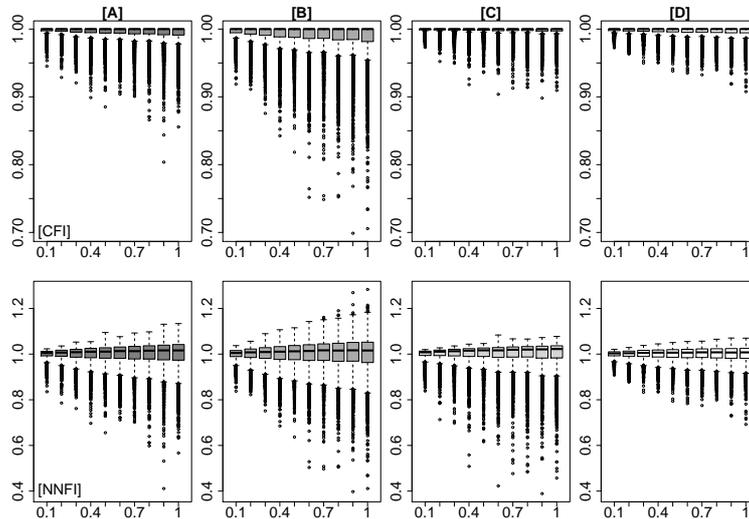


Figura 6. Rappresentazione tramite boxplot distribuzione dei valori di CFI (grafici superiori) e NNFI (grafici inferiori) ottenuti sui quattro modelli in funzione della varianza di errore (valori dei parametri ψ). Dato un valore di ascissa, ciascuna scatola permette di identificare mediana (linea interna alla scatola), primo e terzo quartile (rispettivamente il lato inferiore e quello superiore della scatola) della relativa distribuzione. L'intervallo di rappresentazione dei valori in ordinata è diverso in quanto diversa la variabilità nei due indici.

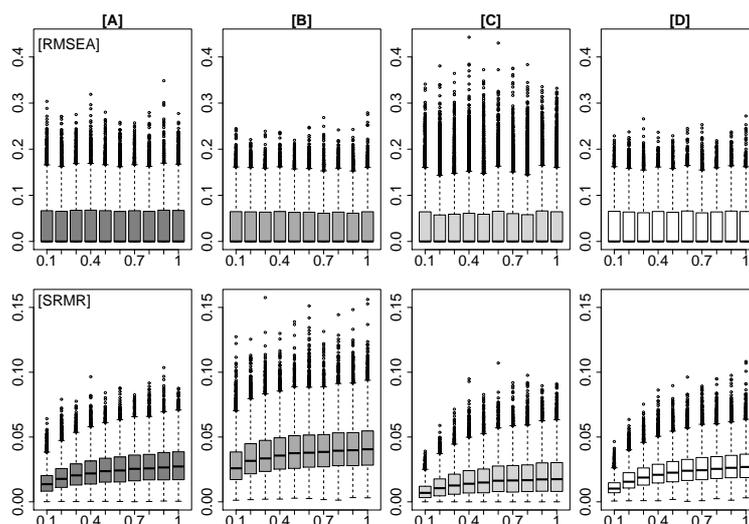


Figura 7. Rappresentazione tramite boxplot distribuzione dei valori di RMSEA (grafici superiori) e SRMR (grafici inferiori) ottenuti sui quattro modelli in funzione della varianza di errore (valori dei parametri ψ). Dato un valore di ascissa, ciascuna scatola permette di identificare mediana (linea interna alla scatola), primo e terzo quartile (rispettivamente il lato inferiore e quello superiore della scatola) della relativa distribuzione. L'intervallo di rappresentazione dei valori in ordinata è diverso in quanto diversa la variabilità nei due indici.

risultano aumentare in funzione dei parametri ψ , da 0.0001 a 0.0002 nei modelli [A] e [D], da 0.003 a 0.004 nel [B] e da meno di 0.001 a 0.003 nell modello [C].

Considerando 0.06 come valore soglia per l'accettabilità del modello (Hu e Bentler, 1999) sono stati ottenuti complessivamente circa il 96% di casi accettabili, rispettivamente 98% in [A], 86% in [B] 99% in [C] e [D].

4.5 Coefficiente di determinazione totale

In figura 8 sono rappresentate le distribuzioni del CD nei quattro modelli, in funzione della varianza d'errore. Pur essendoci lievi differenze nei valori assoluti tra i quattro modelli considerati, risultano evidenti una diminuzione del valore mediano e contemporaneamente un aumento della variabilità nella distribuzione di tale indice. Nel modello [A] il valore della mediana passa da 0.91 a 0.49, nel modello [B] da 0.78 a 0.27, nei modelli [C] e [D] da 0.97 a circa 0.61. Le varianze delle distribuzioni dei parametri invece aumentano rispettivamente da 0.0003 a 0.0052 in [A], da 0.0015 a 0.0060 in [B] e da 0.0001 a 0.0050 nei modelli [C] e [D].

A differenza degli altri IA considerati, per CD non è definito un valore soglia per l'accettabilità del modello, in generale ci si aspetta che tale indice risulti il più alto possibile. Abbiamo quindi analizzato questo indice in relazione agli altri considerati precedentemente. In figura 9 abbiamo riportato la percentuale di casi accettati utilizzando un macro-criterio che tiene conto di tutti gli indici nel seguente modo: χ^2 non significativo, GFI, AGFI, CFI e NNFI maggiori o uguali a 0.96, RMSEA minore o uguale a 0.05, SRMR minore o uguale a 0.06. Tra questi casi abbiamo considerato accettati quelli per cui CD era minore ad una soglia variabile tra 0.1 e 1. Riportando questi valori scelti di CD in ascissa e la percentuale di casi accettati in ordinata possiamo osservare in sostanza quante volte il modello in esame viene accettato anche a fronte di un valore di CD al di sotto di una soglia definita. In generale risulta ovvio che aumentando il valore soglia di CD aumenta anche la percentuale di casi accettati, si osservano però delle sensibili differenze tra i quattro modelli considerati. In particolare, nel modello [B], anche considerando casi con CD piuttosto basso, la percentuale di modelli che verrebbe accettata utilizzando gli altri indici risulta consistente. Ad esempio, per valori di CD minori o uguali a 0.40, che ricordiamo indica una

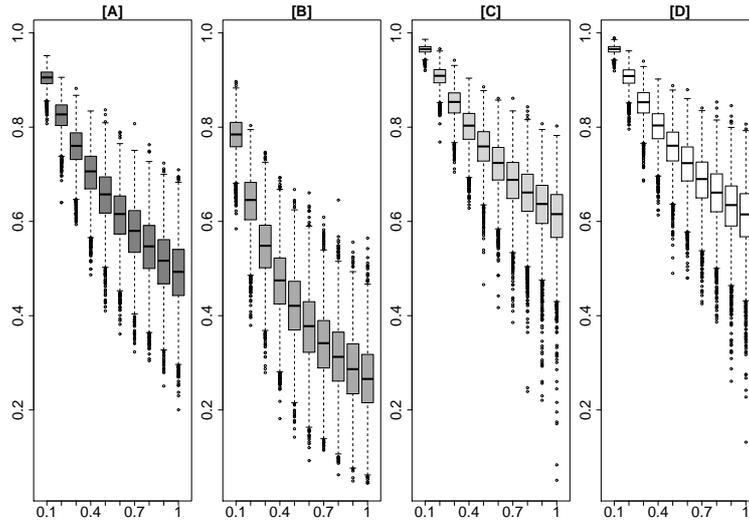


Figura 8. Rappresentazione tramite boxplot distribuzione dei valori di CD ottenuti sui quattro modelli in funzione della varianza di errore (valori dei parametri ψ). Dato un valore di ascissa, ciascuna scatola permette di identificare mediana (linea interna alla scatola), primo e terzo quartile (rispettivamente il lato inferiore e quello superiore della scatola) della relativa distribuzione.

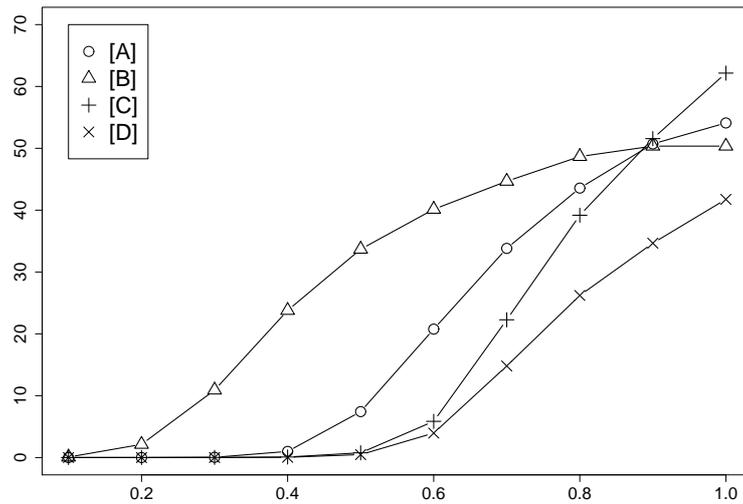


Figura 9. Relazione tra CD e quota di modelli accettabili sulla base degli altri IA. I modelli vengono considerati accettati se presentano i seguenti requisiti: χ^2 non significativo, GFI, AGFI, CFI e NNFI maggiori o uguali a 0.96, RMSEA minore o uguale a 0.05, SRMR minore o uguale a 0.06. In ascissa sono riportati i valori del CD utilizzati come soglia. In ordinata sono rappresentate, separatamente per i quattro modelli [A], [B], [C] e [D], le percentuali di casi accettati in cui il valore osservato di CD è minore di quanto riportato in ascissa.

proporzione di varianza spiegata inferiore al 40%, abbiamo più del 20% di casi in cui gli altri IA portano all'accettazione del modello. Negli altri tre modelli osserviamo percentuali simili quando consideriamo valori soglia di CD compresi tra 0.60 e 0.80.

In sostanza, dato che CD dipende dalla quota di errore del modello, espressa nel determinante della matrice Ψ , risulta evidente che tenda ad abbassarsi con l'aumento dei valori in tale matrice rilevando, di conseguenza, un minore adattamento del modello ai dati.

5 CONCLUSIONI

Nell'esperimento che abbiamo presentato, sono stati presi in esame 4 diversi MVO. Tenendo conto della forma generale dei MVO (vedi formula 1), sono stati fissati i valori dei parametri strutturali (γ e β) e manipolato la varianza relativa agli errori del modello (parametri ψ). In questo modo abbiamo potuto valutare il comportamento di alcuni indici di adattamento utilizzati solitamente in letteratura in funzione dell'aumento dell'errore nei modelli. A livello generale sono emersi due risultati che ci sembrano particolarmente interessanti: il primo relativo ai parametri dei modelli ed il secondo, che era obiettivo di questo lavoro, relativo agli IA.

Dall'analisi dei parametri (vedi Fig. 3) abbiamo osservato che la variabilità associata ad essi risulta direttamente proporzionale alla quota di errore. Anche se si tratta di un risultato atteso, va considerato che tale variabilità riduce di fatto la forza delle relazioni analizzate. In altri termini e considerando uno scenario empirico, questo potrebbe implicare che una proporzione molto alta nella varianza d'errore del modello comporti più difficilmente una significatività statistica dei parametri.

Il secondo risultato è che molti IA risultano poco o per nulla sensibili alla quota di errore e questo li rende di conseguenza non sufficientemente informativi. Compito degli IA sarebbe quello di permettere una valutazione sulla bontà del modello ipotizzato dal ricercatore in relazione ai dati rilevati; come riportano Gerbing e Anderson (1993) un indice ideale dovrebbe indicare il grado di adattamento lungo un ipotetico *continuum* in cui 1 indica il perfetto adattamento e 0 il peggiore (o viceversa come negli indici basati sui residui). Pertanto ci dovremmo aspettare che una quota rilevante di errore nel modello (da noi manipolata aumentando i valori dei relativi parametri) incida su di essi, con un conseguente peggioramento degli IA. Nei casi che abbiamo considerato, i valori della statistica χ^2 risultano totalmente indipendenti dalla varianza d'errore. Di conseguenza tutti gli indici che dipendono direttamente dalla funzione di minimizzazione (GFI e AGFI) o da χ^2 , che ne è una diretta trasformazione (CFI, NNFI e RMSEA), si comportano in maniera insoddisfacente. Tre di questi indici (GFI, AGFI e RMSEA) non mostrano neppure un incremento della variabilità legato all'aumento dell'errore. In pratica, come abbiamo osservato, la percentuale di modelli accettabili risulta essere complessivamente elevata, senza relazione alcuna con l'aumento dei valori nei parametri di varianza d'errore.

Gli unici indici che si comportano come vorremmo sono quelli basati sulla matrice di covarianza (SRMR e CD) e che, a differenza degli altri, tengono conto anche delle varianze degli errori. A proposito di SRMR va detto però che, come gli altri IA, comporta quasi sempre un valore che porterebbe all'accettazione del modello. Di conseguenza anche questo indice sembra poco utile nei casi in cui il modello dovesse presentare una consistente quota di errore.

In sostanza, l'unico indice tra quelli da noi considerati che risulta evidentemente sensibile all'aumento di varianza d'errore è il CD. A proposito di questo indice va fatta una considerazione: nell'ambito delle discipline psicologiche e sociali sembra totalmente ignorato. Nei lavori da noi presi in considerazione non viene mai riportato né citato. Jöreskog e Sörbom (1996), nel manuale relativo all'uso di LISREL, uno dei software utilizzati per i MES, lo descrivono ma al tempo stesso il CD non viene prodotto tra gli IA nell'output del programma. Altri noti software dedicati ai MES quali ad esempio AMOS (Arbuckle, 2009) o Mplus (Muthén e Muthén, 2007) non ne fanno neppure menzione. Tra i limiti del CD riportiamo che esso è sensibile a valori molto bassi nella diagonale principale di $\hat{\Psi}$ (Bollen, 1989) e a valori statisticamente diversi da zero nelle posizioni fuori diagonale (Bagozzi e Yi, 2008), ed in generale va interpretato con cautela nei modelli ricorsivi o calcolato in maniera diversa (Jöreskog, 2000). Un ulteriore limite di questo indice consiste nella mancanza di valori di soglia riconosciuti per cui la sua interpretazione può risultare meno immediata.

L'impressione ricavata dal nostro esperimento è comunque che l'informazione apportata dal CD possa essere utile nella valutazione di un modello e in particolare quando si tratta di un MVO. In sintesi, se affiancato agli altri indici più usati, esso permetterebbe di dare una migliore interpretazione del modello. Ci sembra evidente che avere degli ottimi IA e contemporaneamente una

bassa quota di variabilità spiegata non possa portare serenamente all'accettazione di un modello³. Nel caso invece in cui buoni IA siano accompagnati anche da valori congrui del CD avremo invece una migliore accettabilità del modello.

Riferimenti bibliografici

- AFFUSO, G., BACCHINI, D. (2009). La reciproca influenza fra la qualità delle relazioni familiari e problemi comportamentali e depressivi in adolescenza. *Psicologia Clinica dello Sviluppo*, 13, 119-137.
- ARBUCKLE, J. L. (2009). *Amos 18 User's Guide*. Chicago: SPSS Inc.
- BAGOZZI, R., P., YI, Y. (2008). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16, 74-94.
- BAUMGARTNER, H., HOMBURG, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing*, 13, 139-161.
- BENTLER, P. M. (1989). *EQS structural equations program manual*. Los Angeles, CA: BMDP Statistical Software.
- BENTLER, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- BENTLER P. M., BONETT D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- BLALOCK, H. M., Jr. (Ed.) (1971). *Causal models in the social sciences*. Chicago, Aldine.
- BOLLEN, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.
- BRECKLER, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, 107, 260-273.
- BRODY, G. H., BEACH, S. R. H., PHILIBERT, R. A., CHEN, Y., MCBRIDE MURRY, V. (2009). Prevention effects moderate the association of 5-HTTLPR and youth risk behavior initiation: Gene \times environment hypotheses tested via a randomized prevention design. *Child Development*, 80, 645-661.
- CHIECO, F., VIENO, A., ALTOÈ, G. (2009). Il ruolo del sostegno sociale nella relazione tra eventi di vita negativi e solitudine in adolescenza. *Psicologia Sociale*, 3, 437-453.
- CURCUTO, M., MARIANI, M. G., LIPPERT, S. K. (2009). La fiducia nei sistemi informatici. Contributo alla validazione italiana di un modello. *Psicologia Sociale*, 2, 255-275.
- FORGET-DUBOIS, N., DIONNE, G., LEMELIN, J., PÉRUSSE, D., TREMBLAY, R. E., BOIVIN, M. (2009). Early child language mediates the relation between home environment and school readiness. *Child Development*, 80, 736-749.
- FOX, J., KRAMER, A., FRIENDLY, M. (2010). *sem: Structural Equation Models*. R package version 0.9-21. <http://CRAN.R-project.org/package=sem>
- GERBING, D.W., ANDERSON, J.C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K.A. Bollen, J.S. Long (eds.), *Testing Structural Equation Modeling*. Newbury Park, CA: SAGE.
- HARELI, S., HARUSH, R., SULEIMAN, R., COSSETTE, M., BERGERON, S., LAVOIE, V., DUGAY, G., HESS, U. (2009). When scowling may be a good thing: The influence of anger expressions on credibility. *European Journal of Social Psychology*, 39, 631-638.
- HODSON, G., CHOMA, B. L., COSTELLO, K. (2009). Experiencing alien-nation: Effects of a simulation intervention on attitudes toward homosexuals. *Journal of Experimental Social Psychology*, 45, 974-978.
- HOLBERT, R. L., STEPHENSON, M. T. (2002). Structural equation modeling in the communication sciences, 1995-2000. *Human Communication Research*, 28, 531-551.
- HOYLE, R. H., KENNY, D. A. (1999). Sample size, reliability, and test of statistical mediation. In R.H. Hoyle (ed.), *Statistical strategies for small sample research* (pp. 195-222). Thousand Oaks, CA: Sage.
- HU, L., BENTLER, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- JÖRESKOG, K. G. (2000). Interpretation of R^2 revisited. Technical documents, Scientific Software. Disponibile on-line: <http://www.ssicentral.com/lisrel/techdocs/r2rev.pdf>.
- JÖRESKOG, K. G., SÖRBOM, D. (1984). *LISREL VI user's guide* (3rd ed.). Mooresville, IN: Scientific Software.

³ Secondo quanto riportato da Bagozzi e Yi (2008) potrebbe verificarsi anche la situazione opposta, ovvero alto valore di CD ma IA scadenti. Nel nostro esperimento questa situazione non si è mai verificata, ma sembra logico che anche in questi casi il modello non dovrebbe essere accettato.

- JÖRESKOG, K. G., SÖRBOM, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software.
- KASTLUNGER, B., MARTINI, M., KIRCHLER, E., HOFMANN, E. (2008). Impegno, soddisfazione e fiducia del cliente bancario: un'analisi empirica a Roma e in Sardegna. *Psicologia Sociale*, 2, 307-324.
- KLINE, R.B. (2005). *Principles and Practice of Structural Equation Modeling* (2nd. edition). New York: Guilford Press.
- KNOWLES, E. D., LOWERY, B. S., SCHAUMBERG, R. L. (2009). Anti-egalitarians for Obama? Group-dominance motivation and the Obama vote. *Journal of Experimental Social Psychology*, 45, 965-969.
- MACCALLUM, R. C. (1995). Model specification: Procedures, strategies, and related issues. In R.H. Hoyle (ed.), *Structural equation modeling: Concept, issues, and applications* (pp. 16-36). Thousand Oaks, CA: Sage.
- MACCALLUM, R. C., AUSTIN, J. T. (2000). Application of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201-226.
- MANZI, C., BRAMBILLA, M., REGALIA, C., VOICI, A. (2009). Prenditi cura di me: ipotesi di contatto e riduzione del pregiudizio all'interno del contesto familiare. *Psicologia Sociale*, 1, 115-135.
- MUTHÉN, L. K., MUTHÉN, B., O. (2007). *Mplus Statistical Analysis With Latent Variables User's Guide*. Los Angeles, CA: Muthén & Muthén.
- NEWHEISER, A., TAUSCH, N., DOVIDIO, J. F., HEWSTONE, M. (2009). Entitativity and prejudice: Examining their relationship and the moderating effect of attitude certainty. *Journal of Experimental Social Psychology*, 45, 920-926.
- OKIMOTO, T. G. (2009). The moderating and mediating role of group identification in observers' reactions to intragroup disrespect. *European Journal of Social Psychology*, 39, 973-990.
- PEDHAZUR, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd ed.). New York: Holt, Rinehart & Winston.
- PIETRANTONI, L., PRATI, G. (2008). Mediatori e moderatori nella ricerca psicosociale: una riflessione metodologica. *Psicologia Sociale*, 2, 325-335.
- R DEVELOPMENT CORE TEAM (2009). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>
- SEATON, E. K., YIP, T., SELLERS, R. M. (2009). A longitudinal examination of racial identity and racial discrimination among african american adolescents. *Child Development*, 80, 406-417.
- STEIGER, J. H. (1989). *EzPATH: A supplementary module for SYSTAT and SYGRAPH*. Evanston, IL: SYSTAT.
- STEIGER, J. H., LIND, J. C. (1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- STEPHENSON, M. T., HOLBERT, R. L. (2003). A Monte Carlo simulation of observable versus latent variable structural equation modeling techniques. *Communication Research*, 3, 332-354.
- TANAKA, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K.A. Bollen, J.S. Long (eds.), *Testing Structural Equation Modeling*. Newbury Park, CA: SAGE.
- TANAKA, J. S., HUBA, G. J. (1985). A fit index for covariance structural models under arbitrary GLS estimation. *British Journal of Mathematical and Statistical Psychology*, 42, 233-239.
- TUCKER L. R., LEWIS C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.
- VIENO, A., SANTINELLO, M., PASTORE, M. (2008). Lo sviluppo del comportamento antisociale durante la fase preadolescenziale. *Giornale Italiano di Psicologia*, 35, 177-192.
- WOJCISZKE, B., ABELE, A. E. BARYLA, W. (2009). Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology*, 39, 69-81.
- WOLFLE, L. M. (2003). The introduction of path analysis to the social sciences, and some emergent themes: An annotated bibliography. *Structural Equation Modeling*, 10, 1-34.
- WRIGHT, S. (1921). Correlation and causation. *Journal of Agriculture Research*, 20, 557-585.
- WRIGHT, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5, 161-215.
- YAU, J. P., TASOPOULOS-CHAN, M., SMETANA, J. G. (2009). Disclosure to parents about everyday activities among american adolescents from mexican, chinese, and european backgrounds. *Child Development*, 80, 1481-1498.