Incorporating Expert Knowledge in Structural Equation Models: Applications in Psychological Research

Integrare il Parere degli Esperti con i Modelli di Equazioni Strutturali: Applicazioni nella Ricerca Psicologica

Gianmarco Altoè, Claudio Zandonella Callegher, Enrico Toffalini and Massimiliano Pastore

Abstract Structural Equation Modeling (SEM) is used in psychology to model complex structures of data. However, sample sizes often cannot be as large as ideal for SEM, leading to a problem of insufficient power. Bayesian estimation with informed priors can be beneficial in this context. Our simulation study examines this issue over a real case of a mediation model. Parameter recovery, power and coverage were considered. The advantage of a Bayesian approach was evident for the smallest effects. The correct formalization of the theoretical expectations is crucial, and it allows for increased collaboration among researchers in Psychology and Statistics.
Abstract *I Modelli di Equazioni Strutturali (SEM) sono spesso utilizzati in psicologia. Tuttavia, campioni limitati portano ad un problema di insufficiente potenza. Il nostro studio di simulazione esamina i vantaggi dell'approccio bayesiano con prior informative nel caso di un modello di mediazione. Sono state considerate la stima dei parametri, il coverage e la potenza. Il vantaggio dell'approccio Bayesiano è risultato evidente per gli effetti minori. La formalizzazione delle aspettative teoriche è cruciale e favorisce una fruttuosa collaborazione tra i ricercatori.*

Key words: Expert elicitation, Informative Priors, Structural Equation Models (SEM), Small sample sizes, Psychological research

---

Gianmarco Altoè
Department of Developmental Psychology and Socialisation, University of Padova
e-mail: `gianmarco.altoe@unipd.it`

Claudio Zandonella Callegher
Department of Developmental Psychology and Socialisation, University of Padova
e-mail: `claudio.zandonellacallegher@phd.unipd.it`

Enrico Toffalini
Department of General Psychology, University of Padova
e-mail: `enrico.toffalini@yahoo.it`

Massimiliano Pastore
Department of Developmental Psychology and Socialisation, University of Padova
e-mail: `massimiliano.pastore@unipd.it`

1 Introduction

Structural Equation Modeling (SEM) encompasses a range of multivariate statistical techniques. SEMs are composed of two parts: a *measurement model* and a *structural model* (see Fig. 1). The measurement model defines unobserved constructs (*latent variables*, circles in Fig. 1) according to a set of measured outcomes (*observed variables*, squares in Fig. 1), whereas the structural model describes the relationships between latent variables. SEMs are widely used in psychology to model complex relations between different latent psychological constructs. However, as the complexity of the model increases, more data are required to obtain accurate parameter estimates and model fit statistics [10]. Nevertheless, in many research settings, the number of participants may be limited and appropriate statistical techniques are required to enhance the reliability of the results.

Often in the literature, the Bayesian approach is suggested over frequentist estimation when limited data are available [1]. The inclusion of prior information can help in the parameter estimation, but researchers have to carefully consider priors choice. However, most of the studies rely on default software prior settings. A recent review, underlined that the use of diffuse default priors can result in severely biased estimates, and this bias can be decreased only by incorporating informative priors [8].

Informative priors allow researchers to include in the analysis relevant knowledge in the field. Researchers could also consider to include opinions of experts. On the base of their experience in the field, experts can evaluate relevant information and help researchers in the definition of a plausible range of values and priors choice. *Elicitation* is a structured procedure that allows experts to express their knowledge and uncertainty about quantities of interest in the form of probability distributions [4, 3]. Elicitation can be used to define priors according to experts' judgement.

The remainder of this article is structured as follows. In Sec. 2, we present a simulation study to evaluate the influence of different prior specifications in the case of SEM with small sample size. For the sake of simplicity (but without losing generalizability) we present a mediation model in which the measurement model is not considered. Following common procedures, all variables were standardized (i.e., mean = 0 and a standard deviation = 1) before fitting all models. In Sec. 3, we discuss the obtained results.

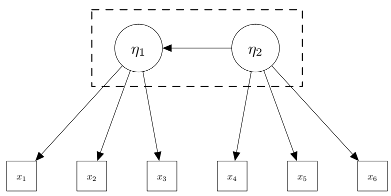

Fig. 1 A structural equation model. Within the dashed box is the structural model, outside is the measurement model. Circles for latent variables; rectangles for observed variables.

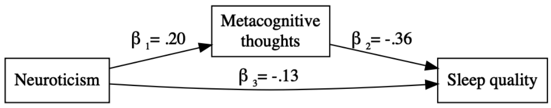

Fig. 2 Mediation model from [7]. The relation between Neuroticism and sleep quality is mediated by metacognitive thoughts.

2 Simulation

We considered a mediation model from [7] presented in Fig. 2. The study evaluated the relationship between participants' self-reported sleep quality (*Sleep quality*), participants' tendency to become anxious (*Neuroticism*), and negative beliefs about sleeping problems (*Metacognitive thoughts*). In particular, the association between *Neuroticism* and *Sleep quality* ($\beta_3 = -.13$) is mediated by *Metacognitive thoughts*. In other words, people with higher levels of distress and anxiety tend to have dysfunctional beliefs and attitudes about sleep ($\beta_1 = .20$) that, in turns, induce them to perceive and report a worse-quality sleep ($\beta_2 = -.36$).

2.1 Simulation details

The simulation was carried in `R` version 3.6.2 [5] using R-packages `lavaan` [6] and `blavaan` [2]. In the simulation, we considered as parameters of interest the regression coefficients ($\beta_1,\ \beta_2,\ \beta_3$). We compared the performance of Maximum Likelihood (ML) estimation and Bayesian estimation under four different sample size conditions (i.e., 20, 50, 100, 500).

Three different prior distribution specifications were used (see Fig. 3):

1. *Default prior* - $\beta_i \sim N(0, 10)$. These are intended to be non-informative.
2. *Reasonable prior* - $\beta_1 \sim N(.20,\ .50)$, and $\beta_{2;3} \sim N(-.20,\ .50)$. These are moderately informative to exclude excessively large values that are not reasonable within psychology research. Moreover, the mean of each prior is set to reflect the direction of the main results in the literature.
3. *Experts prior* - $\beta_1 \sim N(.20,\ .20)$, $\beta_2 \sim N(-.40,\ .20)$, and $\beta_3 \sim N(-.20,\ .20)$. These are intended to be highly informative representing experts' judgement.

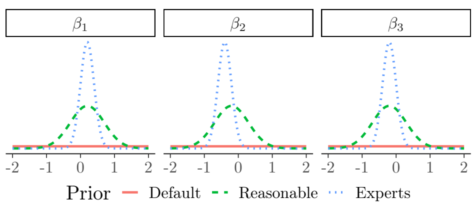

Fig. 3 Prior distribution in the three different settings. Default priors are intended to be non-informative. Reasonable priors are intended to exclude implausible values. Experts priors represent experts' judgement.

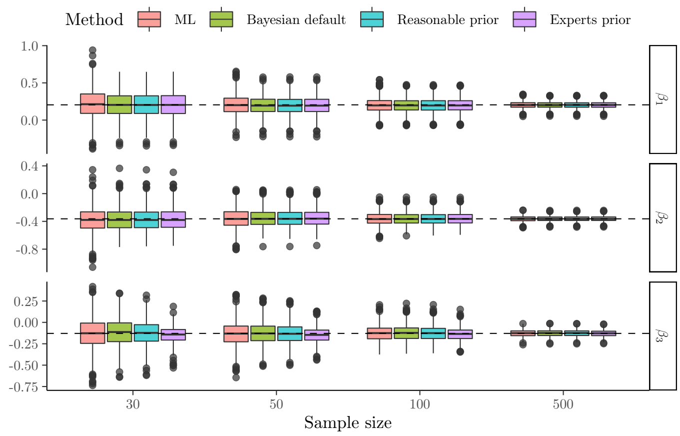

Fig. 4 Estimates distribution for each parameter across the different condition. Dashed lines represent the true population values.

Relative mean bias, relative median bias, mean square error (MSE), coverage and power were considered [9]. The relative mean bias (or median bias) evaluates the relative difference between mean estimate ($\bar{\theta}$; or median estimate $\widetilde{\theta}$) across replications and the population value (θ). Relative bias included between -.10 and .10 are considered acceptable [9]. MSE takes into account variability as well as bias of the estimates: $MSE = \sigma^2 + (\bar{\theta} - \theta)^2$, where σ is the standard deviation of the estimates across replications and $\bar{\theta}$ is the mean. Coverage is the proportion of replications in which the population value is included in the 95% confidence interval (CI; for the ML estimation) or 95% highest posterior density interval (HPD; for the Bayesian estimation). Instead, power is the proportion of replications in which the value zero is not included in the 95% CI or 95% HPD. Analyses were conducted considering the standardized parameters and for each condition 1000 replications were considered.

2.2 Results

The tables with detailed results for each parameter and condition are available at `https://osf.io/hwj8d/`. To interpret the results of relative mean and median bias, we considered the distribution of the estimated parameters (see Fig. 4). Only with very small sample sizes ($n = 30$) it is possible to observe some differences between estimation methods: Maximum likelihood approach produces the widest distributions, whereas Bayesian approach with experts prior has narrower distributions. However, differences between methods are noticeable for the parameter β_3 (i.e., the parameter with the smallest population value) but are less evident for the other parameters and, as the sample size increases, estimation methods perform similar to each other.

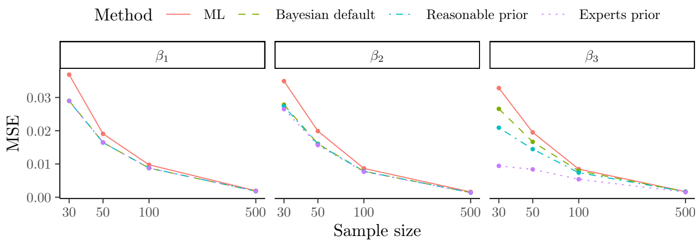

Fig. 5 Mean Squared Error (MSE) values for each parameter across the different conditions.

Considering the MSE (see Fig.5), we have the same results pattern. Differences between methods are bigger in the case of very small samples ($n = 30$), where Bayesian approach with experts prior performs better. However, differences between prior specification are noticeable only for the parameter β_3.

Finally, the result of coverage and power are presented in Fig. 6. Coverage reaches adequate levels in all conditions with sample size equal to or greater than 100. With smaller sample sizes, Bayesian approach with experts prior showed excessive coverage in the case of the parameter β_3. Power is extremely low when sample sizes are small. ML estimation performs slightly better in terms of power across all conditions, except for the parameter β_3 where is outperformed by Bayesian approach with experts prior. However, adequate levels of power are reached for all parameters only with large sample sizes ($n = 500$).

3 Discussion and conclusions

In the simulation, we evaluated the different estimation methods in the case of SEMs with small sample size. Overall, results indicate that informative priors are useful in

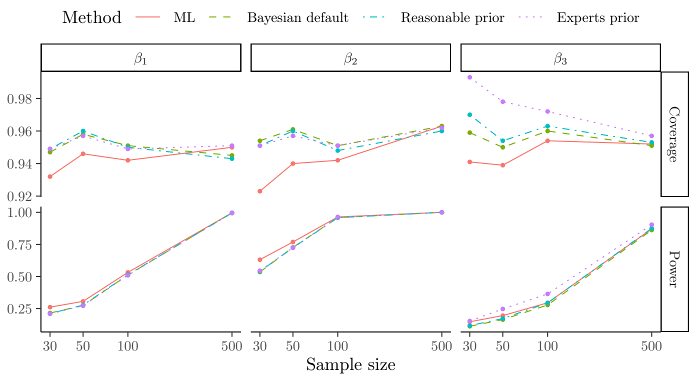

Fig. 6 Coverage and power values for each parameter across the different conditions.

the case of limited sample sizes and when the true population values are small. Parameter estimates were more stable across replications and extreme values were less likely. When the sample size increases the difference between estimation methods becomes less evident.

However, results are not consistent for all the parameters. In most conditions, Bayesian approach performs better than ML but results are very similar between the different prior specifications. Only in the case of small true population parameter values the Bayesian approach with expert priors performs much better than the other prior specifications. Future studies should focus on the role of prior definition in SEMs with different levels of complexity (e.g., also taking the measurement part into account) and in which the effect sizes vary on a larger range. Another important aspect that future studies should evaluate is the impact of prior knowledge misspecification, in particular in situations with small sample sizes.

Finally, we want to highlight that expert knowledge elicitation is not only useful to inform prior distributions but it can help also in other aspects of the analysis. Experts can help and inform researchers in the design of the experiments, definition of the models, interpretation of the results and make reasonable and informed choices along all the research process. Thus, the collaboration between different experts is a crucial point that should be encouraged in any applied research field.

References

1. McNeish, D.: On Using Bayesian Methods to Address Small Sample Problems. Structural Equation Modeling: A Multidisciplinary Journal **23**(5), 750–773 (2016). DOI 10.1080/10705511.2016.1186549
2. Merkle, E.C., Rosseel, Y.: *Blavaan* : Bayesian Structural Equation Models via Parameter Expansion. Journal of Statistical Software **85**(4) (2018). DOI 10.18637/jss.v085.i04
3. O'Hagan, A. (ed.): Uncertain Judgements: Eliciting Experts' Probabilities. Statistics in Practice. John Wiley & Sons, London ; Hoboken, NJ (2006)
4. O'Hagan, A.: Expert Knowledge Elicitation: Subjective but Scientific. The American Statistician **73**(sup1), 69–81 (2019). DOI 10.1080/00031305.2018.1518265
5. R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing (2018)
6. Rosseel, Y.: *Lavaan* : An *R* Package for Structural Equation Modeling. Journal of Statistical Software **48**(2) (2012). DOI 10.18637/jss.v048.i02
7. Sella, E., Carbone, E., Toffalini, E., Borella, E.: Personality traits and sleep quality: The role of sleep-related beliefs. Personality and Individual Differences **156**, 109,770 (2020). DOI 10.1016/j.paid.2019.109770
8. Smid, S.C., McNeish, D., Miočević, M., van de Schoot, R.: Bayesian Versus Frequentist Estimation for Structural Equation Models in Small Sample Contexts: A Systematic Review. Structural Equation Modeling: A Multidisciplinary Journal **27**(1), 131–161 (2020). DOI 10.1080/10705511.2019.1577140
9. Smid, S.C., Rosseel, Y.: Sem with Small Samples. In: R. van de Schoot, M. Miočević (eds.) Small Sample Size Solutions, first edn., pp. 239–254. Routledge (2020). DOI 10.4324/9780429273872-20
10. Wolf, E.J., Harrington, K.M., Clark, S.L., Miller, M.W.: Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. Educational and Psychological Measurement **73**(6), 913–934 (2013). DOI 10.1177/0013164413495237