

Preregistration: Comparing a Structural After Measurement (SAM) Approach to Standard Structural Equation Model (SEM) Estimation

Valentin Kriegmair

2024-06-04

Version: 0.1.0

Last updated: 2023-10-31

Preregistration template designed by: Björn S. Siepe, František Bartoš, Tim P. Morris, Anne-Laure Boulesteix, Daniel W. Heck, and Samuel Pawel

1. General Information

1.1. What is the title of the project?

Comparing a Structural After Measurement (SAM) Approach to Standard Structural Equation Model (SEM) Estimation

1.2. Who are the current and future project contributors?

Valentin Kriegmair

1.3. Provide a description of the project.

The studies registered here are part of an adversarial collaboration project. Here the aim is to conceptually (only in part) replicate the results obtained by Dhaene and Rosseel (2023) and Rosseel and Loh (2022). I will evaluate the performance of a Structural After Measurement (SAM) approach for estimating structural equation models (SEM) in comparison to standard SEM estimation methods. This will serve as the basis for the adversarial collaboration with another researcher who will evaluate the same research question from the perspective of a conceptual replication of the (in part contradicting) results obtained by Robitzsch (2022). However, the following only describes the first (conceptual) replication.

1.4. Did any of the contributors already conduct related simulation studies on this specific question?

No prior related simulation studies have been conducted by the contributors.

2. Aims

Structural After Measurement (SAM) is an estimation method for structural equation models that consists of a stepwise estimation of the measurement and structural parts of a model. The aims of the current simulation are:

1. How do SAM and traditional SEM methods (including ML and ULS) compare in terms of bias, Mean Squared Error (MSE), and convergence rates in small to moderate samples?
2. What is the impact of model misspecifications, such as residual correlations and cross-loadings, on the performance of SAM compared to traditional SEM methods?

3. Data-Generating Mechanism

3.1. How will the parameters for the data-generating mechanism (DGM) be specified?

3.2. Study 1

In study 1 (conceptually replicating Rosseel and Loh (2022)) data will be generated parametrically. Four different population structural equation models (SEM) with latent variables and continuous indicators based on the following matrices will be simulated:

- B as $M \times M$ matrix representing latent regression coefficients with all $b = 0.1$.
- Model 1.1, 1.2, and 1.3:

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0.1 & 0 & 0.1 & 0 \\ 0.1 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0.1 & 0 \end{bmatrix}$$

- Model 1.4 (reversed effect between η_3 and η_4):

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0.1 & 0 & 0 & 0 \\ 0.1 & 0.1 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0.1 & 0 \end{bmatrix}$$

- Ψ as $M \times M$ as diagonal matrix (0 on the off diagonal) representing variances of the factors with $1 - kb^2$ on the diagonal where k is the number of latent regressor per factor and b the regression coefficients (0.1): so the variance for f3 is $1 - 2 * 0.1^2 = 0.98$ and for f4 and f5 $1 - 2 * 0.1^2 = 0.97$.

- Models 1.1 - 1.3:

$$\Psi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.97 & 0 & 0 \\ 0 & 0 & 0 & 0.98 & 0 \\ 0 & 0 & 0 & 0 & 0.98 \end{bmatrix}$$

- Model 1.4: (reversed effect between η_3 and η_4)

$$\Psi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.97 & 0 & 0 \\ 0 & 0 & 0 & 0.97 & 0 \\ 0 & 0 & 0 & 0 & 0.98 \end{bmatrix}$$

- Λ as $P \times M$ matrix representing factor loadings.

- Model 1.1, 1.3 and 1.4:

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.7 & 0 & 0 & 0 & 0 \\ 0.7 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0.7 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0.7 \\ 0 & 0 & 0 & 0 & 0.7 \end{bmatrix}$$

- Model 1.2: cross loadings will be set to be 10% lower than the factor loadings: $\Lambda_{ik,jk} = 0.63 = 0.9 \times 0.7$. They will be generated by the following elements in Λ : (2, 2), (5, 3), (8, 4), (11, 5), (14, 1).

- Θ as a $P \times P$ matrix representing the residual variances and covariances of the indicators.

- Model 1.1, 1.2 and 1.4: The diagonal generated as:

$$\Theta^* = \text{Var}(\eta) \Lambda^T \times \frac{1}{r-1}$$

(where r is the reliability of the indicators) and 0 on all off-diagonal elements

- Model 1.3:

* Θ^* on the diagonal.

- * Correlated residuals generated between specific indicator pairs: for $i = (2, 5, 8, 11, 14)$ and $i' = (3, 6, 9, 12, 15)$, and for each $k = 1, \dots, 4$ and $l = k + 1, \dots, 5$, the entries (i_k, i'_l) and (i'_l, i_k) in Θ are set to $0.6 \times \min \Theta^*$, ensuring correlated errors among selected indicator pairs without exceeding a 0.6 correlation coefficient.

3.2.2. Study 2:

For study 2, again, different population models will be generated parametrically. Further, the different models of study 2 will be used for different simulation settings resulting in two sub-studies 2.1 and 2.2 (see simulation settings).

- B as $M \times M$ matrix representing latent regression coefficients with varying parameter size defined by two conditions of endogenous factor variance explained by the exogenous factors (low: $R^2 = 0.1$ or medium: $R^2 = 0.4$ see below under factor):

– Model 2.1 and 2.2:

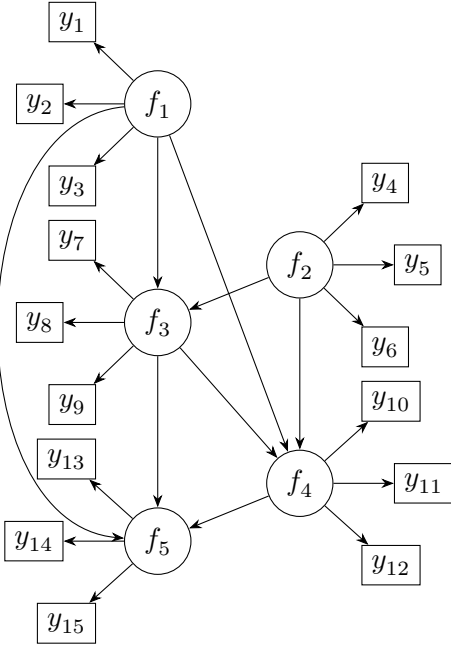
$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \beta_{\eta_4, \eta_1} & \beta_{\eta_4, \eta_2} & 0 & 0 & 0 \\ 0 & \beta_{\eta_5, \eta_2} & \beta_{\eta_5, \eta_3} & \beta_{\eta_5, \eta_4} & 0 \end{bmatrix}$$

- Λ as $P \times M$ matrix representing factor loadings of indicators on the latent factors.

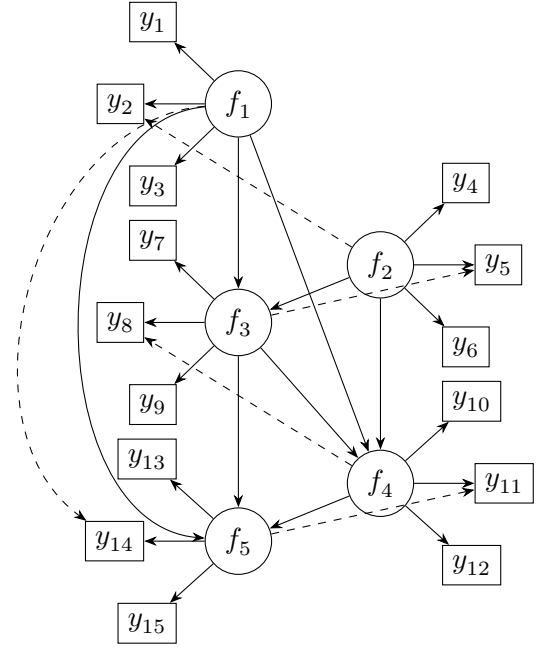
– Model 2.1:

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.8 \end{bmatrix}$$

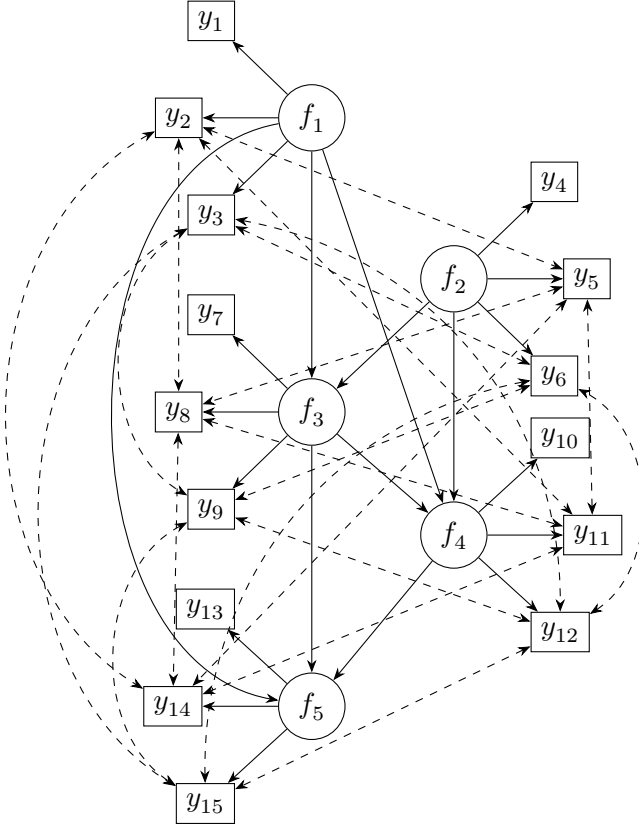
- Model 2.2 with cross-loadings either in the exogenous ($\lambda_{6,3}$), endogenous ($\lambda_{12,5}$) or both parts of the model. Which cross loading is present depends on the misspecification simulation factor. The specific magnitude of the endogenous ($\lambda_{12,5}$) loading depends on



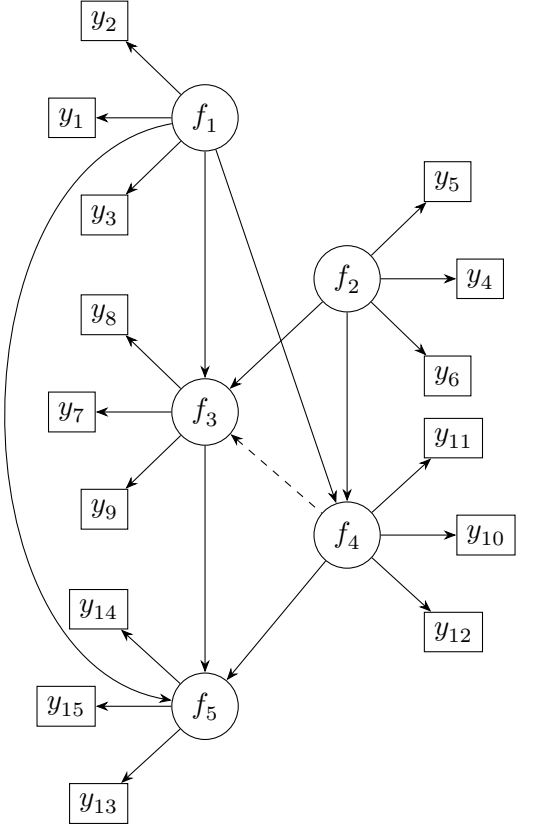
Model 1.1 Note. Error terms are not explicitly shown in the figure.



Model 1.2 Note. Error terms are not explicitly shown in the figure.



Model 1.3 Note. Error terms are not explicitly shown in the figure.



Model 1.4 Note. Error terms are not explicitly shown in the figure.

R^2 (see under 3.2.2):

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & \lambda_{6,3} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 0.8 & \lambda_{12,5} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.8 \end{bmatrix}$$

- Θ as a $P \times P$ matrix representing the residual variances and covariances of the indicators. This is computed as the portion of the indicator's total variance that is not explained by the latent factors, after accounting for the strength and reliability of its relationship to these factors (factor loadings), as well as the effects of regressions between the latent factors themselves.

– Model 2.1: The diagonal of Θ generated as:

$$\Theta^* = \text{Var}(\eta) \Lambda^T \times \frac{1}{r-1}$$

(where r is the reliability of the indicators) and 0 on all off-diagonal elements

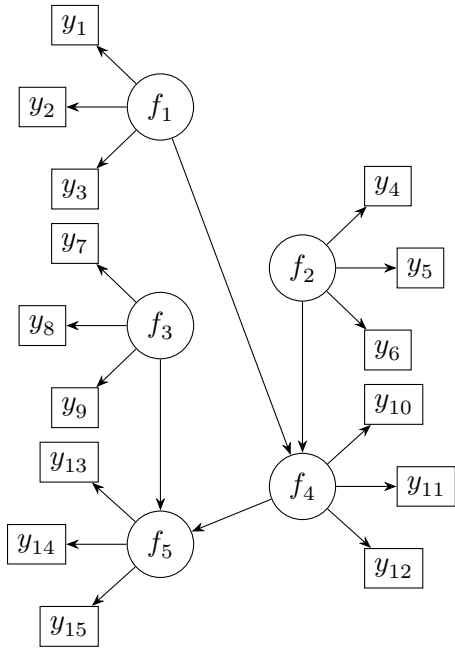
– Model 2.2:

- * Θ^* on the diagonal.
- * Correlated residuals generated between specific indicator pairs in either the endogenous, exogenous or both parts of the model.

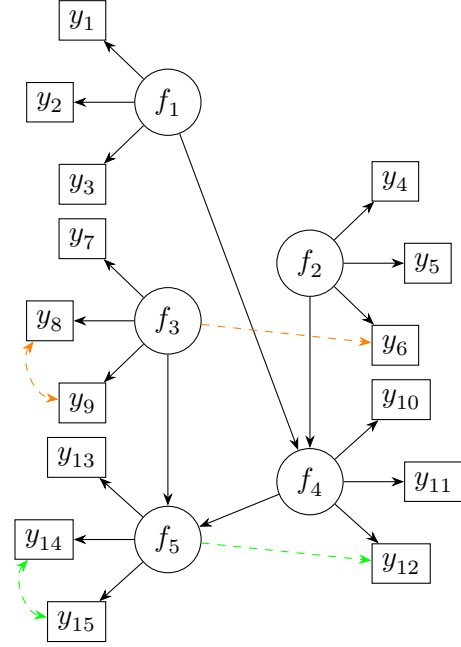
Thus depending on the simulation setting either:

- * $\Theta_{8,9}$, $\Theta_{9,8}$ (exogenous part)
- * $\Theta_{14,15}$ and $\Theta_{15,14}$ (endogenous part)
- * $\Theta_{8,9}$, $\Theta_{9,8}$, $\Theta_{14,15}$ and $\Theta_{15,14}$ (both parts)

are set $0.6 \times \min \Theta^*$, ensuring correlated errors among selected indicator pairs without exceeding a 0.6 correlation coefficient:



Model 2.1 Note. Error terms are not explicitly shown in the figure.



Model 2.2 Note. Error terms are not explicitly shown in the figure. Orange lines represent misspecifications in the endogenous part of the model. Green lines represent misspecifications in the exogenous part of the model. These types of misspecifications result in different realisations of model 2.2 when they are modulated as factors in study 2 but are subsumed under one model here.

3.2. What will be the different factors of the data-generating mechanism?

3.2.1. Study 1

The first study will modulate the following factors:

- Different misspecifications of the population model where the population model varies between the different models (1.1, 1.2, 1.3, 1.4) as described above, while the analysis model remains specified as model 1.1.
- Sample sizes of small ($N = 100$), medium ($N = 400$), or large ($N = 6400$)
- Indicator reliability of low (.3), moderate (.5), or high (.7)

3.2.2. Study 2:

The second study will modulate the following factors if the data generating process across both studies:

- **Sample sizes of small ($N = 100$)**, medium ($N = 400$), or large ($N = 6400$)
- **Variance explained (R^2)** of the endogenous factor variance explained by the exogenous factors: low ($R^2 = 0.1$) or medium ($R^2 = 0.4$)

In two sub-studies 2.1 and 2.2, the following additional factors will be modulated, respectively:

- **Study 2.1:**
 - **Indicator reliability** of three indicators per factor: *all high* (.8), *all low* (.5), *average low* (.5) varying between .7 to .3 with the highest reliability for the scaling indicator.
 - **Sample sizes** of small ($N = 100$), medium ($N = 400$), or large ($N = 6400$)
 - **Distribution:** either normal or non-normal with non-normal data generated by skewing factors and residuals with a kurtosis of -2 and 8, respectively.
- **Study 2.2**
 - **Structural misspecifications** by varying the population model as described and thus omitting either
 - * a residual covariance in the exogenous part of the model
 - * a factor loading in the endogenous part of the model
 - * a factor loading in the exogenous part of the model
 - * a structural path in the endogenous part of the model
 - * a structural path in the analysis model.
 - **Number of measurement blocks** (how many separate measurement models are fitted in the first step of SAM):
 - * Separate measurement model per latent variable ($b = k = 5$)
 - * Joint measurement model for all exogenous variables ($b = 3$)

Explanation of the chosen factor values:

The aim is a conceptual replication of the results obtained by Dhaene and Rosseel (2023) in study 1 and Rosseel and Loh (2022) study 2. The values were chosen accordingly. No detailed explanation was provided by the authors.

If there is more than one factor: How will the factor levels be combined and how many simulation conditions will this create?

- Study 1:
 - 4 population models x 2 model specifications x 3 sample sizes x 3 reliabilities = 72 conditions
 - * Note: For each of the 4 population models (1.1, 1.2, 1.3, 1.4), two scenarios are considered:
 - Correctly specified analysis model matching the population model.
 - Incorrectly specified analysis model, corresponding to model 1.1, with specific misspecifications for models 1.2 (omitted cross-loadings), 1.3 (correlated item residuals), and 1.4 (structural misspecification).
- Study 2.1:
 - 2 population models (2.1, 2.2) x 3 sample sizes x 3 reliabilities x 2 distributions = 36 conditions
- Study 2.2:
 - 2 population models x 3 sample sizes x 5 misspecifications = 30 conditions

4. Estimands and Targets

Estimands

Structural model parameters (path coefficients)

5. Methods

Both studies will compare four different estimation methods for SEMs:

- **“Vanilla” SEM:** (structural and measurement model estimated simultaneously) (rationale: the current standard approach in SEM estimation serving as a baseline with maximum likelihood (ML)):
- **SAM:** (separating the estimation of the measurement and structural model to alleviate the potential for propagation of bias from (e.g. misspecified) measurement part to the structural part of the model)
 - **Local SAM** (Uses summary statistics from the measurement model to derive the model-implied mean vector and variance-covariance matrix of latent variables. These statistics are then utilized to estimate the structural parameters. A mapping matrix (M) is used to transform the observed data into the latent variable space. It can be estimated using different methods.)
 - * **With ML mapping matrix** (Akin to a factor score approach (M. S. Bartlett (1937), Maurice S. Bartlett (1938)))

- * **With ULS mapping matrix** (Uses the Moore-Penrose pseudoinverse, suitable for scenarios with complex or underdetermined systems, where the K matrix is rank-deficient but requires adjustments for structural constraints.)
- **Global SAM** (rationale: Fixing the parameters obtained from the measurement model in the first step, and then using them as constants in the full SEM during the second step. Suitable for models where local SAM is impractical due to higher-order latent variables or rank deficiencies in λ .)

“Vanilla” SEM as well as both steps in the SAM approach will be estimated using Maximum Likelihood (ML) using the `lavaan` (Rosseel (2012)) package in R (R Core Team (2023)).

6. Performance Measures

Across both studies the following performance measures will be captured:

- Convergence rates: Proportions of observed data sets that successfully converged for each estimation method.
- Empirical relative biases: Average difference between an estimate and its true value, normalized by the true value, assessed across all path coefficients.
- Empirical coverage levels of 95% confidence intervals (CIs): Proportion of observed data sets where the constructed CIs included the true value.
- Root Mean Squared Errors (RMSE): Calculated as the square root of the average squared difference between an estimate and its true value, evaluated under conditions of model misspecification.

How will Monte Carlo uncertainty of the estimated performance measures be calculated and reported?

Monte Carlo uncertainty will be calculated using the `simhelpers` (“Helper Functions for Simulation Studies” (n.d.)) package in R as Monte Carlo Standard Errors (MCSEs):

For the bias:

$$\sqrt{S_T^2/K}$$

For the RMSEA:

$$\sqrt{\frac{K-1}{K} \sum_{j=1}^K (RMSE_{(j)} - RMSE)^2}$$

$$S_T^2 = \frac{1}{K-1} \sum_{k=1}^K (T_k - \bar{T})^2$$

How many simulation repetitions will be used for each condition?

- Replicating Rosseel and Loh (2022) study 1 will consist of 5000 repetitions per condition.
- Replicating Dhaene and Rosseel (2023) study 2 will consist of 10000 repetitions per condition.

How will missing values due to non-convergence or other reasons be handled?

As mentioned above the convergence rates will be captured and evaluated.

How do you plan on interpreting the performance measures? (optional)

The results will be interpreted in comparison to the results and interpretations obtained by Dhaene and Rosseel (2023) and Rosseel and Loh (2022).

Which statistical software/packages do you plan to use?

The simulation will be set up in R Core Team (2023) using the `lavaan` package for generating data based on the population models as well as for applying SEM estimation methods (Rosseel (2012)) as well as the `furrr` (Vaughan and Dancho (2022)) package for parallel simulation execution.

Which computational environment do you plan to use?

The simulation will run on a high performance computing cluster using “TARDIS Core Package Documentation — Tardis” (n.d.) with “A Future for r: A Comprehensive Overview” (n.d.) and Heidilohr (n.d.)

Which other steps will you undertake to make simulation results reproducible? (optional)

The code of the simulation will be made available on GitHub
(<https://github.com/valentinkm/AdversarialSimulation>)

References

- “A Future for r: A Comprehensive Overview.” n.d. Accessed April 23, 2024. <https://cran.r-project.org/web/packages/future/vignettes/future-1-overview.html>.
- Bartlett, M. S. 1937. “The Statistical Conception of Mental Factors.” *British Journal of Psychology. General Section* 28 (1): 97–104. <https://www.proquest.com/docview/1293463650/citation/9FE530638F5041E0PQ/1>.
- Bartlett, Maurice S. 1938. “Methods of Estimating Mental Factors.” *Nature*.
- Dhaene, Sara, and Yves Rosseel. 2023. “An Evaluation of Non-Iterative Estimators in the Structural After Measurement (SAM) Approach to Structural Equation Modeling (SEM).” *Structural Equation Modeling: A Multidisciplinary Journal* 30 (6): 926–40. <https://doi.org/10.1080/10705511.2023.2220135>.
- Heidilohr. n.d. “Azure Virtual Desktop Documentation.” Accessed April 23, 2024. <https://learn.microsoft.com/en-us/azure/virtual-desktop/>.
- “Helper Functions for Simulation Studies.” n.d. Accessed April 23, 2024. <https://meghapsimatrix.github.io/simhelpers/>.

- R Core Team. 2023. “R: A Language and Environment for Statistical Computing.” Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Robitzsch, Alexander. 2022. “Comparing the Robustness of the Structural After Measurement (SAM) Approach to Structural Equation Modeling (SEM) Against Local Model Misspecifications with Alternative Estimation Approaches.” *Stats* 5 (3): 631–72. <https://doi.org/10.3390/stats5030039>.
- Rosseel, Yves. 2012. “Lavaan: An r Package for Structural Equation Modeling.” *Journal of Statistical Software* 48 (May): 1–36. <https://doi.org/10.18637/jss.v048.i02>.
- Rosseel, Yves, and Wen Wei Loh. 2022. “A Structural After Measurement Approach to Structural Equation Modeling.” *Psychological Methods*, No Pagination Specified–. <https://doi.org/10.1037/met0000503>.
- “TARDIS Core Package Documentation — Tardis.” n.d. Accessed April 23, 2024. <https://tardis-sn.github.io/tardis/>.
- Vaughan, Davis, and Matt Dancho. 2022. “furrr Furrr: Apply Mapping Functions in Parallel Using Futures.” 2022.