EAM2025

XI Conference

23RD - 25TH JULY 2025

Spain Tenerife
Canary Islands

European Association of Methodology

Standard Error Estimation in the Local Structural-After-Measurement (LSAM) Approach

Seda Can and Yves Rosseel

Significance of Standard Errors (SEs) in SEM

- Parameter Estimation: Extensively studied in SEM (e.g., distributions, estimators)
- Standard Errors (SEs): Critical but underexplored (Deng et al., 2018)
- Why SEs Are Important:
  - Reflect Sampling Variability: Essential as SEM relies on sample data, not full populations
    - Indicator of Precision:
    - Smaller SEs = Higher precision of parameter estimates
    - Larger SEs = Greater uncertainty
  - Role in Statistical Testing:
    - Key for z-scores, t-scores, and significance testing
    - Ensures robust and reliable conclusions in population inferences (Yuan & Hayashi, 2006)

Significance of Standard Errors (SEs) in SEM

- Two Approaches in SEM:
  - Standard Estimation (Joint ML):
    - Uses maximum likelihood (ML) to estimate parameters simultaneously.
  - Structural-After-Measurement (SAM) Approach:
    - Local SAM (LSAM): A SAM approach dividing estimation into separate measurement and structural components (Rosseel & Loh, 2022).
- Purpose: Systematically evaluate SE estimation under varying conditions within the LSAM approach to SEM, focusing specifically on continuous (and complete) data
- Research Gap:
  - While LSAM point estimates have been examined (Dhaene & Rosseel, 2023), SE behavior in LSAM remains unexplored.

Comparison of Standard Estimation and LSAM

Aspect	Joint ML	Local SAM (LSAM)
Estimation Method	Joint, **system-wide** estimation of all parameters (Bollen, 1996)	**Two-stage** approach: estimates measurement first, then structural part (Rosseel & Loh, 2022)
Optimization	Relies on **iterative procedures**	Allows **noniterative estimators** for measurement part (Dhaene & Rosseel, 2021)
Addressing Model Issues	Prone to **nonconvergence** and improper solutions with small samples (e.g., negative variances)	**Mitigates interpretational confounding** and convergence issues by separating measurement from structural estimates
Flexibility for Complex Models	Limited flexibility for complex models	Enables **implementation of some complex models** (e.g., multi-group mixtures) that are not possible with joint estimation (Perez Alonso et al., 2024)

Different approaches to compute SEs: standard estimation

1. Calculate the Unit Information Matrix:
   - Can be either observed or expected (Savalei, 2010)
2. Derive the Variance-Covariance Matrix:
   - Invert the unit information matrix.
   - Divide by sample size (N) to reflect variability in estimates
3. Take the square root of diagonal elements of the variance-covariance matrix to get SEs

- Robust SEs in ML Framework:
  - To protect against deviations from normality and correct model specification

Different approaches to compute SEs: standard estimation cont.

- **Two-Step SE Estimation**
- Using the analytic procedure below, a joint information matrix $\boldsymbol{I}$ is computed for all parameters in the full model, which is then partitioned as follows:

$$I = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}$$

  - where the 1-index corresponds to the measurement part, and the 2-index corresponds to the structural part

- The two-step corrected variance-covariance matrix for the structural parameters ($\Sigma_{2(1)}$) is then computed as follows, where Σ_{11} represents the variance-covariance matrix obtained from Step 1:

$$\Sigma_{2(1)} = I_{22}^{-1} + I_{22}^{-1} I_{21} \Sigma_{11} I_{12} I_{22}^{-1}$$

- This procedure follows the method outlined in Equation 17 from Bakk et al. (2014) to manage these uncertainties, building on the work of Gong and Samaniego (1981) and refined further by Parke (1986)

Different approaches to compute SEs: standard estimation

- Under nonnormality, using robust standard errors is critical to avoid biased inference.
- **Two-step corrected SEs** account for measurement-model uncertainty but assume:
  - Correct model specification
  - Multivariate normality
- To address this limitation, **Yuan & Chan (2002)** proposed a **robust version** of the two-step SE correction (See Equations (4a) and (4b) in their paper)
- Implementation available via sam() function in **lavaan** (v0.6-20+)
- Our study compares both versions (**standard vs. robust**) under:
  - Normal and nonnormal data
  - Within the LSAM estimation framework

Bootstrapping for Robust SEs in SEM

- Bootstrapping in SEM:
- Widely used to obtain robust SEs without relying on distributional assumptions
  - Effective in small samples

- Key Findings from Research:
  - Larger SEs under skewed data: Boomsma (1986) showed bootstrap SEs > ML SEs with skewed data
  - Improved accuracy in SEM: Bollen & Stine (1990) demonstrated bootstrap SEs for direct/indirect effects.
  - Consistency in nonnormal data: Nevitt & Hancock (2001) found bootstrapping effective with $n \geq 200$ for bias reduction.
  - SE Consistency with Misspecification: Yuan & Hayashi (2006) observed bootstrap SEs remained consistent under model misspecifications, unlike analytic SEs.

Nonparametric vs. Parametric Bootstrapping in SE Estimation

- Nonparametric Bootstrapping:
  - Widely Used: Popular for its flexibility, avoiding strict distributional assumptions.
  - Works well with small samples, directly resampling from observed data
- Parametric Bootstrapping:
  - Assumes a Specific Distribution: Useful when there is a reasonable assumption about data distribution
  - Reduced Variability: By relying on a specified distribution, it may lead to more accurate SE estimates
  - Advantages in Small Samples: Offers reliable inferences when limited data may not represent the full sampling distribution (Hestenberg, 2015)

Aim and Design of the Current Study

- This study aims to assess the performance of both analytic and resampling-based SEs derived in LSAM approaches, including two-step, as well as nonparametric and parametric bootstrapping.
- To evaluate SE estimation under varying conditions, two simulation studies were conducted, differing primarily in the models employed.
- A few characteristics were common to both studies:
    - Estimation Methods:
        - LSAM (Two-step, Robust two-step, nonparametric, and parametric bootstrapping)
        - Joint ML SEM (standard and robust)
    - Sample Sizes: 50, 100, 200, 500, 1000
    - The outcome measure of interest (SE)

Aim and Design of the Current Study

- **Design and Conditions:**
  - Estimation Methods:
    - LSAM (**Two-step, Robust two-step, nonparametric, and parametric bootstrapping**)
    - Joint ML SEM (**standard and robust**)
  - Sample Sizes: **50, 100, 200, 500, 1000**
  - Model Specification: **Correctly specified** and **misspecified model**
  - Distributions: **Normal and nonnormal**

Study 1

- In Study 1, the misspecification condition was introduced by omitting two residual covariances between the second and third indicators within each latent variable from the analysis model, which were specified as 0.40 in the population model.
- Nonnormal latent scores with skewness of −2 and excess kurtosis of 8

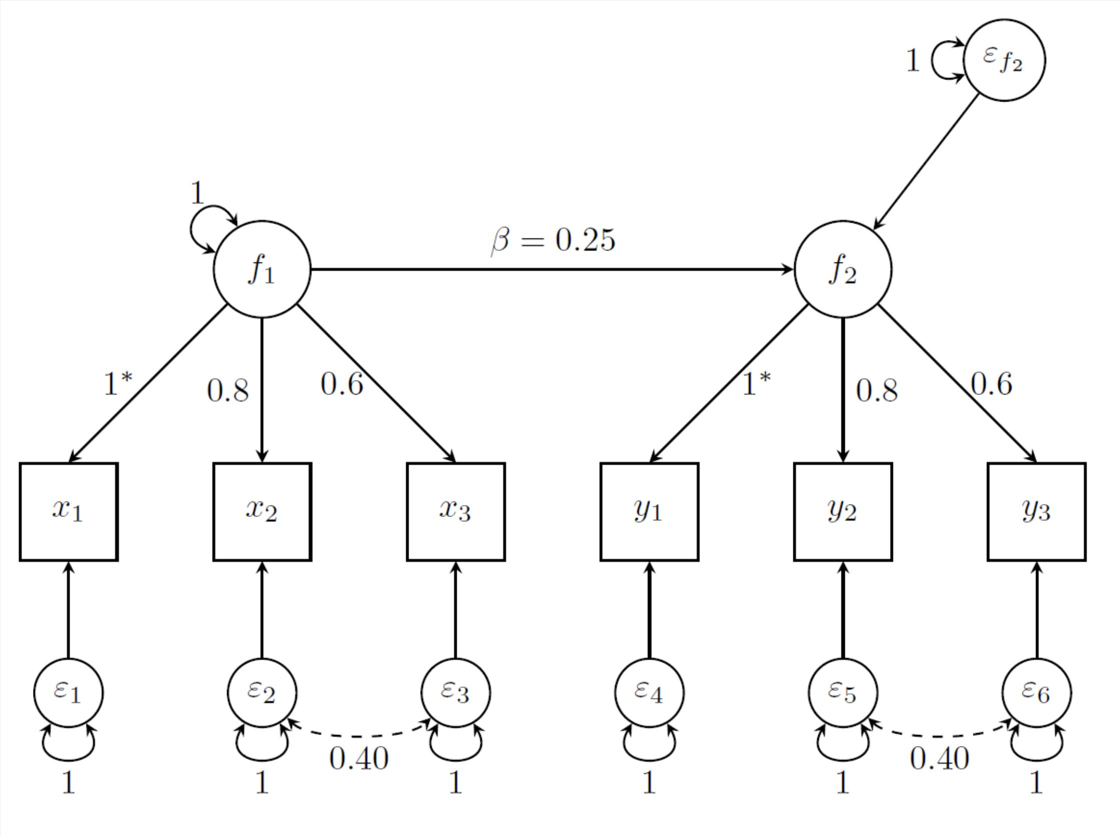


Figure 1
The model and unstandardized population values used in the simulations for Study 1. Residual covariances (dashed double-headed arrows) are included in the population model but omitted in the analysis model under the misspecified condition. For scaling purposes, the first factor loading of each latent variable is fixed to 1 (denoted by 1 in the figure).*

Study 2

- Study 2 introduced misspecification in the structural part by removing the path from f1 to f2.
- Study 2 extended nonnormality to include exogenous variables, disturbances, and residuals.
  - Exogenous vars: skew = −2, kurtosis = 8
  - Nonnormal disturbances (ζ1 and ζ2) were generated using centered exponential distributions, with rate 1 and variances set to 0.91 and 0.71, respectively.
  - Residuals were generated using centered exponential distributions, with λ as the specified rate parameter

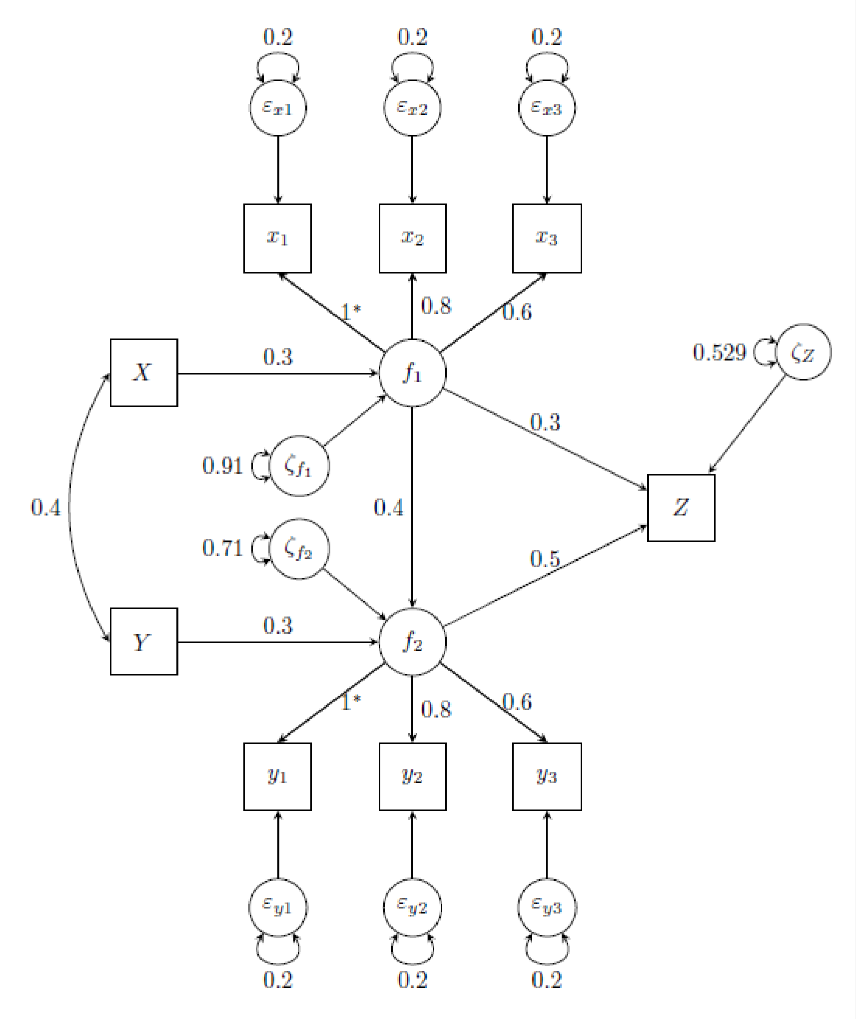


Figure 2
The model and unstandardized population values used in the simulations for Study 2. For scaling purposes, the first factor loading of each latent variable is fixed to 1 (denoted by 1 in the figure).*

Bias Assessment

- Empirical SEs were calculated as the standard deviation of point estimates across replications:
  - 10,000 replications were used for non-resampling methods (e.g., standard and two-step approaches),
  - 1,000 for resampling-based methods (i.e., nonparametric and parametric bootstrap)
- Model-based SEs were obtained by:
  - averaging the SE estimates provided by each method across these replications
- SE Bias: the ratio of the model-based SE to the empirical SE for each method
  - Ratio = 1: Unbiased SE estimate.
  - Ratio > 1: Indicates SE overestimation
  - Ratio < 1: Indicates underestimation

Study 1 Results (LSAM methods)

SAM Two-step yielded accurate SEs under normal conditions, but showed increasing underestimation under nonnormality, especially in misspecified models.

SAM Robust improved performance under nonnormal/correct conditions as sample size grew, but still underestimated SEs when the model was misspecified.

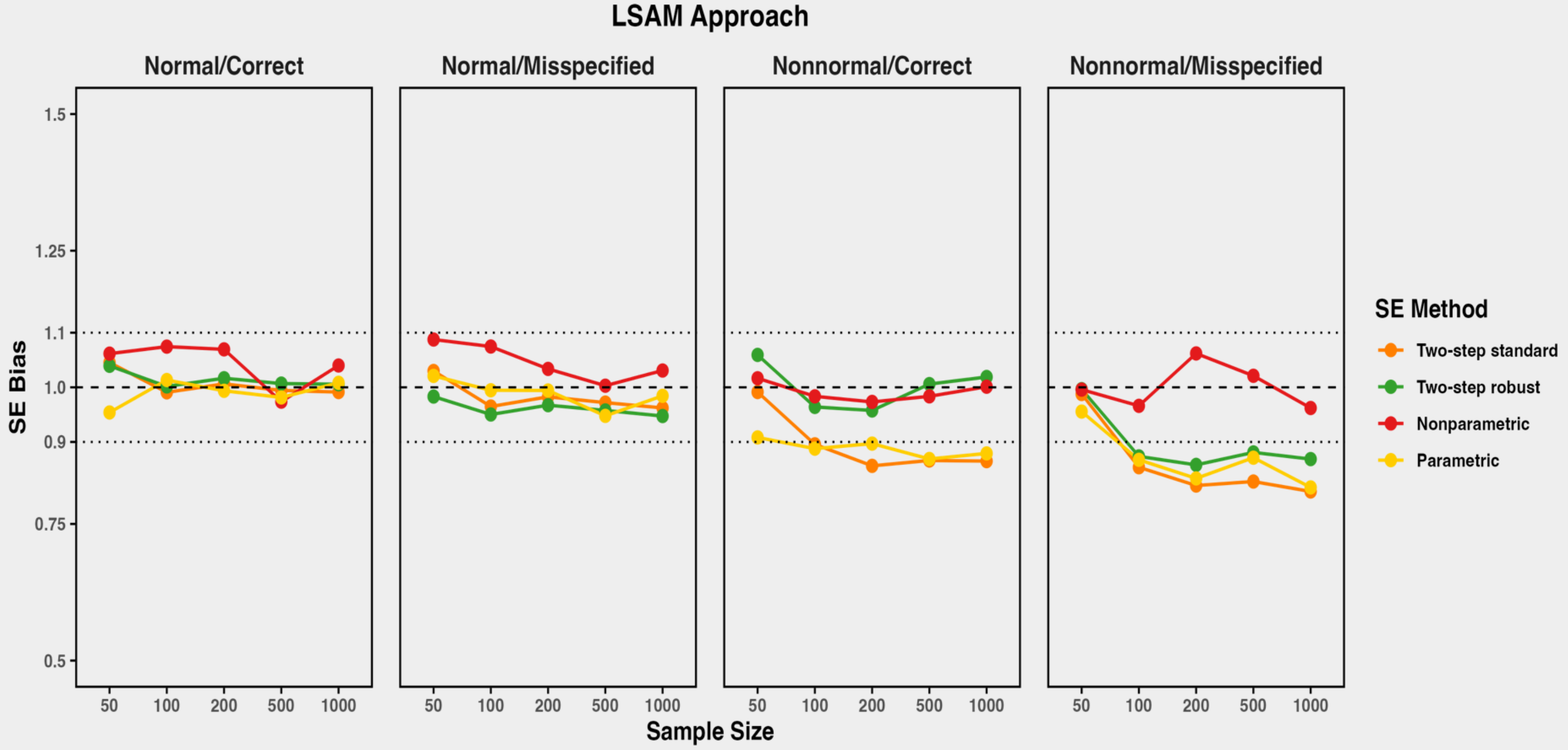


Figure 3. *Bias in SEs across various sample sizes and SE methods under different conditions in Study 1.*

Study 1 Results (LSAM methods)

SAM Nonparametric performed consistently well under nonnormality, even at small sample sizes, though it slightly overestimated under normality.

SAM Parametric was accurate under normality but showed notable SE underestimation under nonnormal/misspecified conditions.

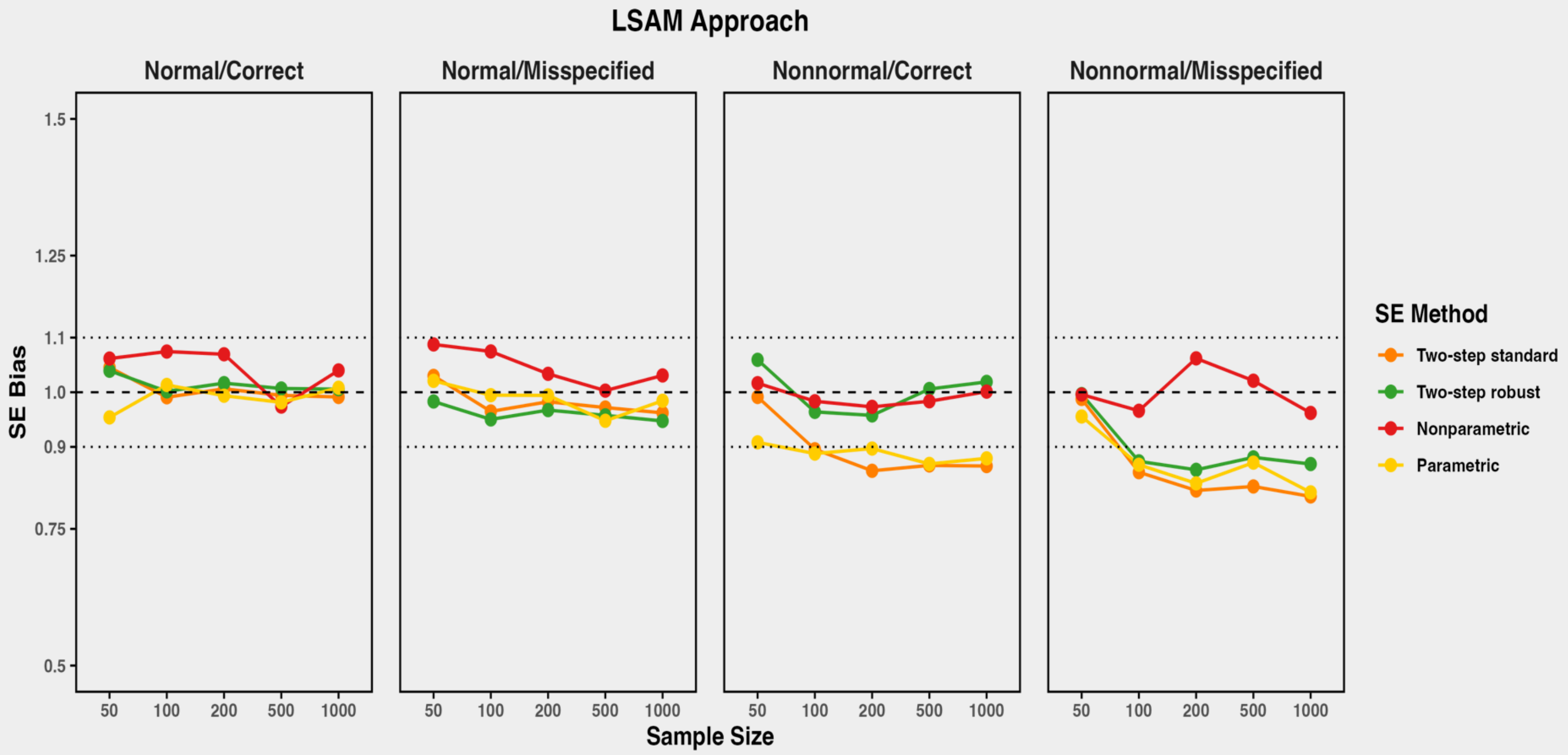
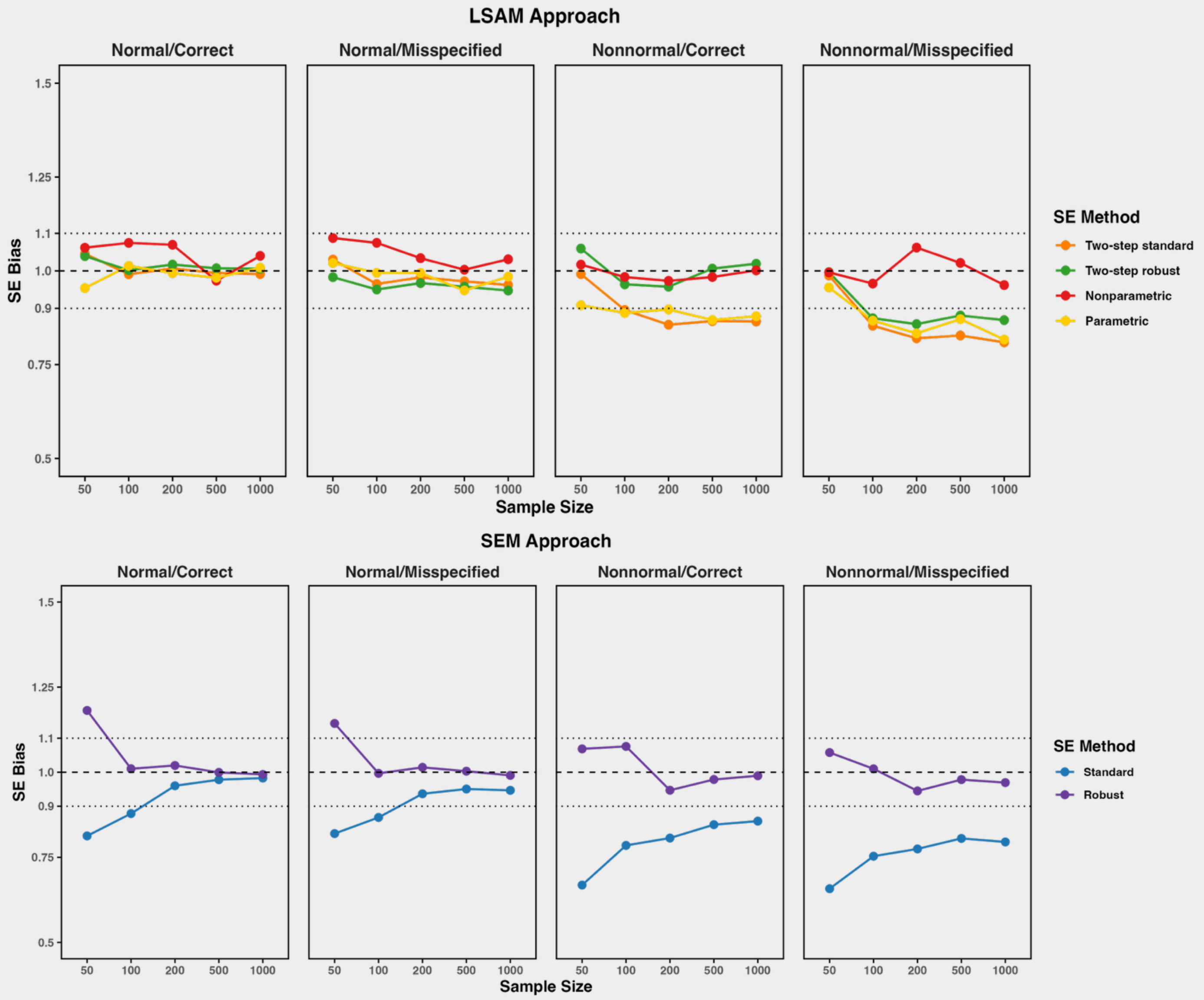


Figure 3. *Bias in SEs across various sample sizes and SE methods under different conditions in Study 1.*

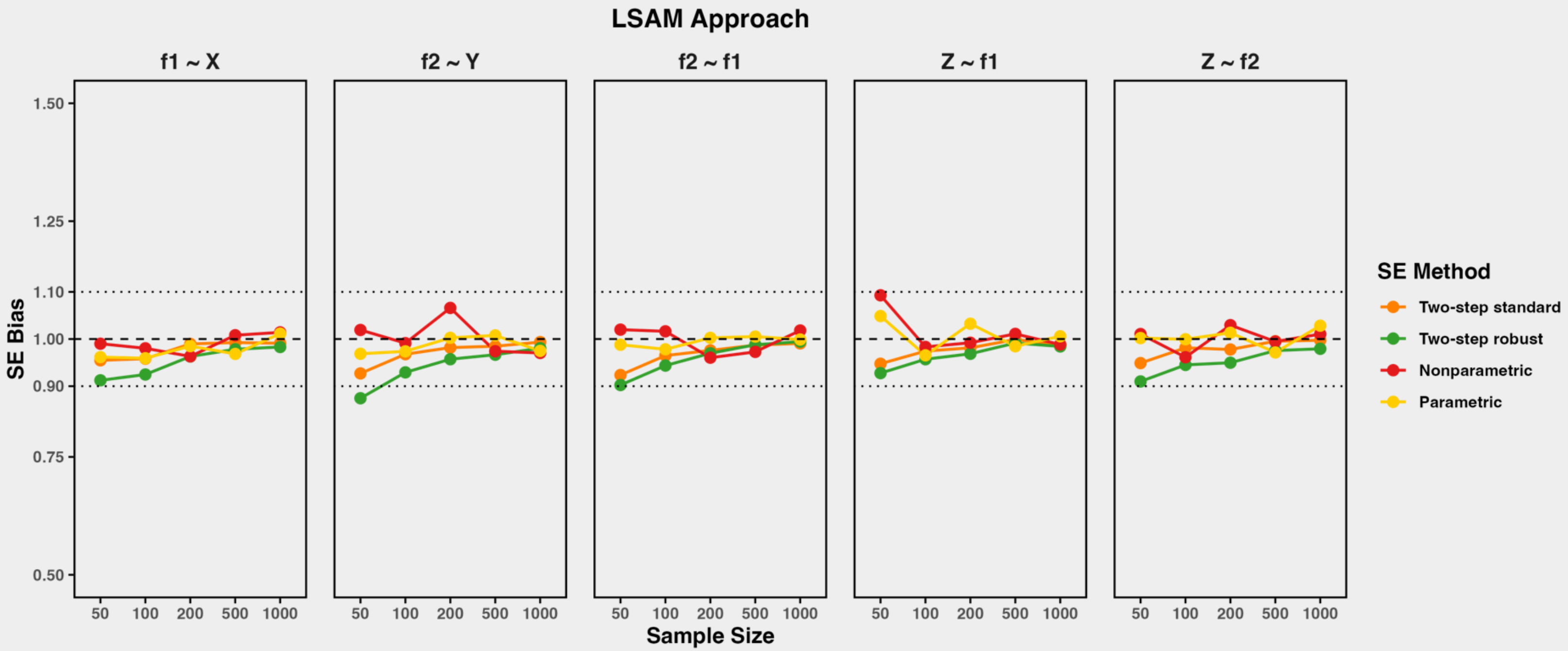
Study 1 Results (SEM methods)

SEM Standard consistently underestimated SEs, especially under nonnormal/misspecified conditions.

SEM Robust initially overestimated in small normal samples but became more accurate with larger *N*. Under nonnormality, SEM Robust outperformed SEM Standard in producing more reliable SEs.



Study 2 Results For Correct Model with Nonnormal Data (LSAM methods)



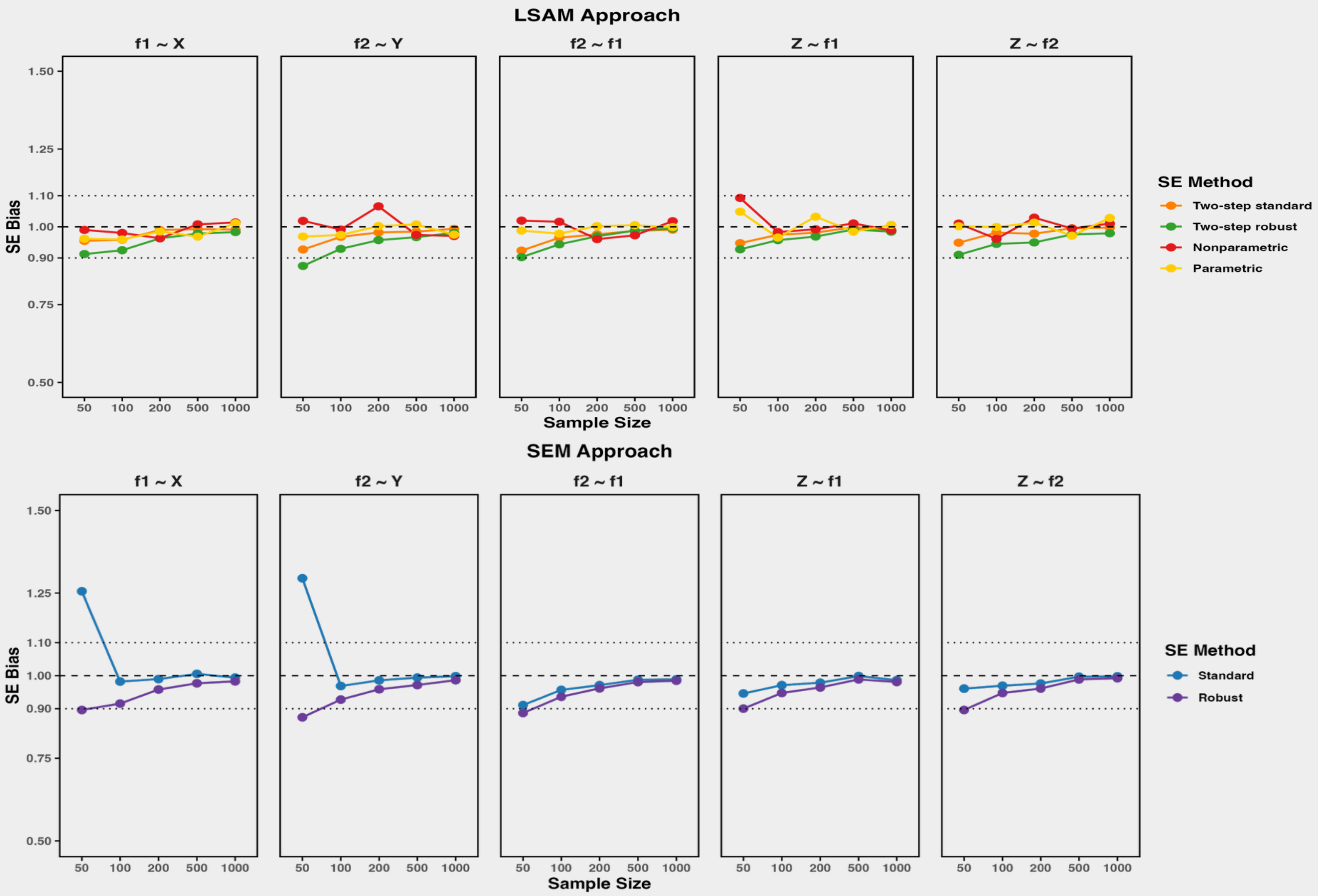
SAM Nonparametric delivered near-unbiased SEs across all regression coefficients

SAM Two-step & **SAM Parametric** showed slightly more variability, depending on parameter and sample size

SAM Robust showed the greatest variability, especially at small sample sizes

Study 2 Results For Correct Model with Nonnormal Data (SEM methods)

Under nonnormal data, **SEM Robust** outperformed **SEM Standard**, which showed more variability, particularly **for** f1 ~X and f2 ~Y in small samples.



Study 2 Results For Misspecified Models

Under normally distributed data:

LSAM Methods

For **f1~X** and **f2~Y:**

All LSAM methods yielded **SE bias values close to 1**

For **paths involving** Z:

SAM Nonparametric provided the **most accurate SEs** across sample sizes

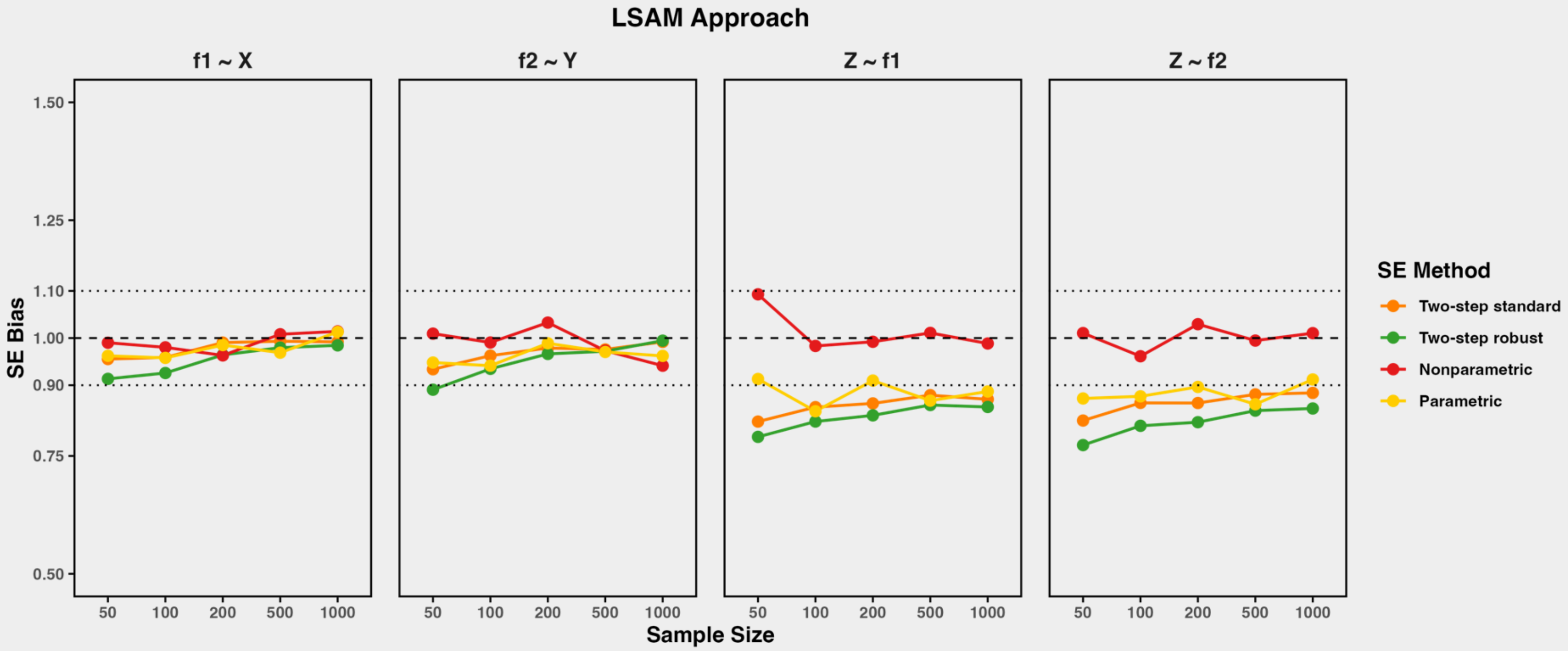
Other LSAM methods showed **underestimation**, with **bias = 0.81-0.91** (≈19%-9% bias)

SEM Methods

SEM Standard had **higher bias**, especially for **Z~ f1** in small samples

SEM Robust outperformed SEM Standard across all paths and sample sizes

Study 2 Results For Misspecified Model with Nonnormal Data (LSAM methods)



For **paths involving** Z:

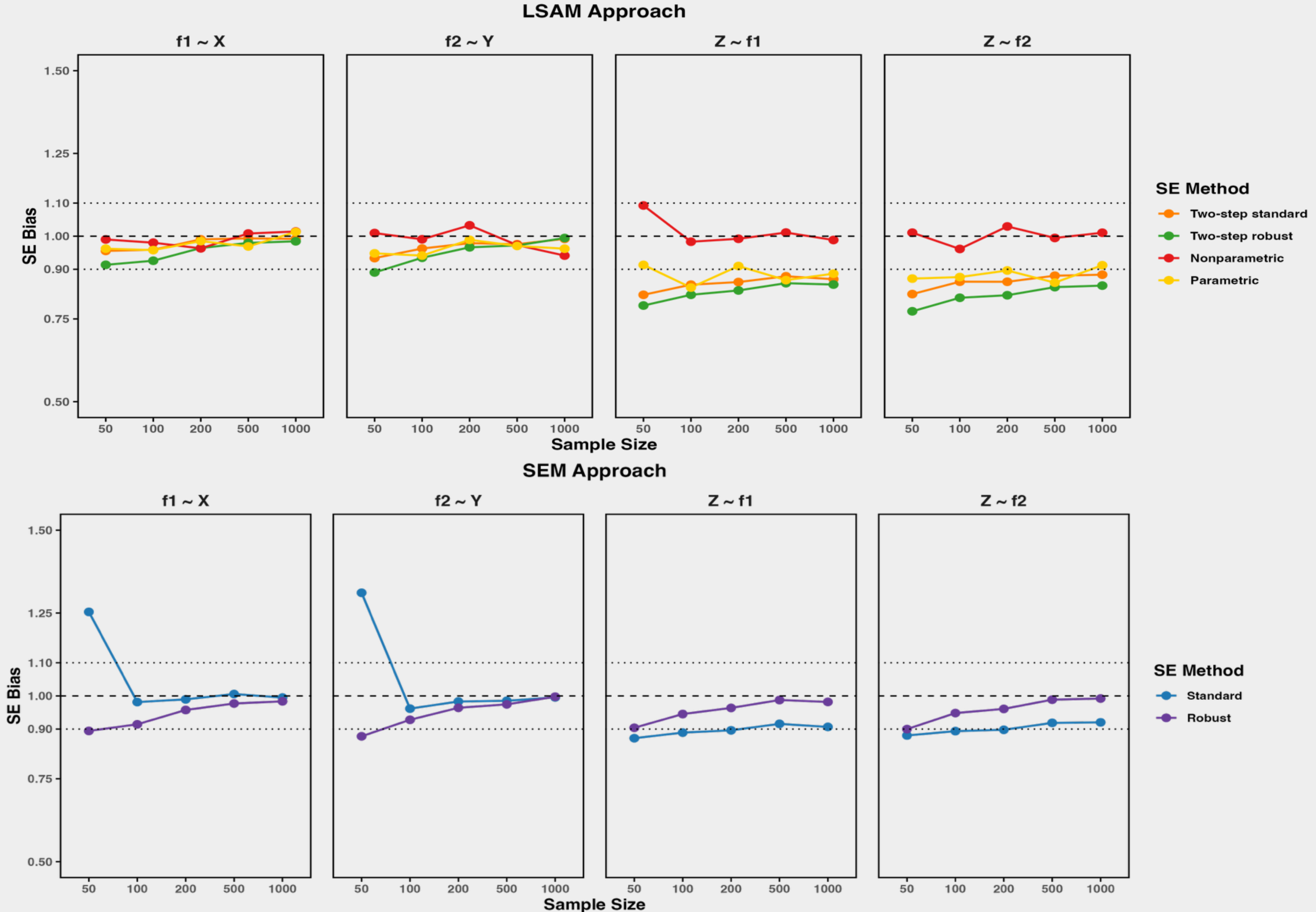
SAM Nonparametric again **performed best**

Other LSAM methods underestimated SEs, bias ranged from 0.81 to 0.91

Study 2 Results For Misspecified Model with Nonnormal Data (SEM methods)

SEM Standard showed the largest biases in small samples e.g., 25% and 31% overestimation for **f1~X** and **f1~Y** at *N = 50*

SEM Robust improved upon SEM Standard but still showed **~10% underestimation** at small *N*. Bias values improved with increasing sample size



Discussion

- First study to evaluate **SE estimation within LSAM** using:
  - Two-step, Robust Two-step, Nonparametric, and Parametric Bootstrap
- **Study 1**: Simple SEM with measurement misspecification and nonnormal latent scores
- **Study 2**: Complex SEM with structural misspecification and three layers of nonnormality

Discussion

- SAM Nonparametric: Most robust and accurate, especially under nonnormality and misspecification
- SAM Parametric: Best under normality, even with misspecification
- SAM Two-step: Accurate in normal/correct; more bias in nonnormal/small-*N* cases
- SAM Robust: Improvement over Two-step in nonnormal/large-*N* scenarios
- SEM Standard: Prone to bias, particularly under nonnormality
- SEM Robust: More stable but not bias-free in small samples
- Our results align with prior bootstrap literature in SEM (Bollen & Stine; Nevitt & Hancock)
- First study to apply parametric bootstrap in SEM → promising but computationally intensive

Discussion: Limitations & Future Work

- The findings are specific to the conditions manipulated in our simulations.
- Study 1 examined a simple two-factor SEM, while Study 2 extended the model by incorporating observed exogenous and endogenous variables, expanding the scope to include more latent variables or exploring complex models, such as latent growth models, could improve the generalizability of these findings and provide greater support to applied researchers.
- Regarding two-step SE estimation in LSAM, the current approach relies on returning to the global model to compute the joint information matrix, which is somewhat incompatible with local SAM. Future research should focus on developing a method that eliminates the need to switch back to a global perspective.

References

Bakk, Z., Oberski, D. L., & Vermunt, J. K. (2014). Relating latent class assignments to external variables: Standard errors for correct inference. *Political Analysis*, 520-540. https://doi.org/10.1093/pan/mpu003

Bollen, K. A. (1996). An alternative two stage least squares (2sls) estimator for latent variable equations. *Psychometrika*, *61(1*), 109-121. https://doi.org/10.1007/BF02296961627

Bollen, K. A., & Stine, R. A. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. In C. C. Clogg (Ed.), *Sociological methodology* (pp. 115-140). Blackwell.

Boomsma, A. (1986). On the use of bootstrap and jackknife in covariance structure analysis. In N. L. F. De Antoni & A. Rizzi (Eds.), *Compstat 1986: Proceedings in computational statistics* (pp. 205-210). Physica.

Deng, L., Yang, M., & Marcoulides, K. M. (2018). Structural equation modeling with many variables: A systematic review of issues and developments. *Frontiers in Psychology*, *9*, 580. https://doi.org/10.3389/fpsyg.2018.00580666

Dhaene, S., & Rosseel, Y. (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). https://doi.org/10.1080/10705511.2023.22201356735

Gong, G., & Samaniego, F. J. (1981). Pseudo maximum likelihood estimation: Theory and applications. *The Annals of Statistics*, *9*, 861-869.

Hestenberg, T. C. (2015). What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum. *The American Statistician*, *69(4*),371-386. https://doi.org/10.1080/00031305.2015.1089789691

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, *8*(3), 353-377.732 https://doi.org/10.1207/S15328007SEM0803_2733

Parke, W. R. (1986). Pseudo maximum likelihood estimation: The asymptotic distribution. *The Annals of Statistics*, *14(*1), 355-357.738

Perez Alonso, A. F., Rosseel, Y., Vermunt, J. K., & De Roover, K. (2024). Mixture multigroup structural equation modeling: A novel method for comparing structural relations across many groups [Advance online publication]. *Psychological Methods*. https://doi.org/10.1037/met0000667

Savalei, V. (2010). Expected versus observed information in sem with incomplete normal and nonnormal data. *Psychological Methods*, *15*(4), 352-367. https://doi.org/10.1037/a0020143757

Yuan, K.-H., & Chan, W. (2002). Fitting structural equation models using estimating equations: A model segregation approach. *British Journal of Mathematical and Statistical Psychology*, *55*(1), 41-62. https://doi.org/10.1348/000711002159699779

Yuan, K.-H., & Hayashi, K. (2006). Standard errors in covariance structure models: Asymptotics versus bootstrap. *British Journal of Mathematical and Statistical Psychology*, *59*, 397-417. https://doi.org/10.1348/000711005X85896785