A comparison of dimensionality assessment methods in factor analysis under second order, bifactorial, and oblique factor models

P.-O. Caron
Univeristé TÉLUQ

87th Annual National Convention of the Canadian Psychological Association

6th June 2026

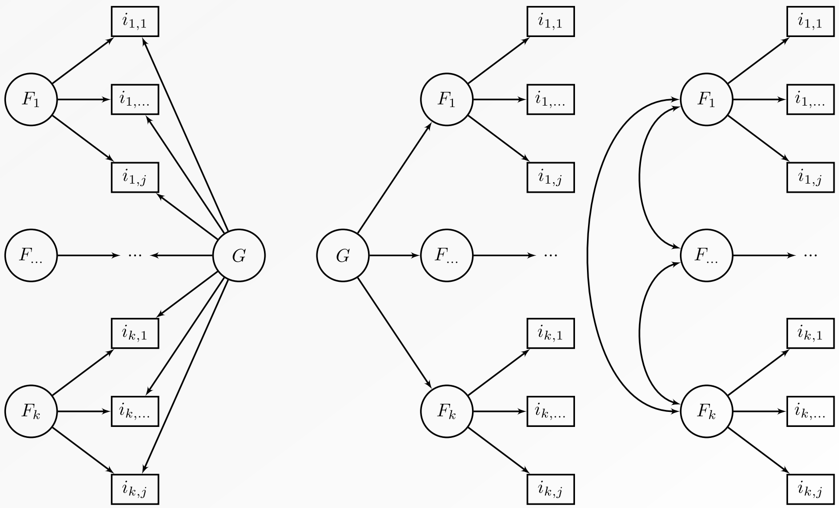
2026-06-06

Introduction

- Determining the number of factors has remained a challenging issue in factor analysis
- Several stopping rules have been proposed
- Several promising recent ones deserved careful comparison, especially in hard, but realistic conditions.

Introduction

- Lack of studies on *general* factor model
  - Many on oblique structures
  - No on 2ND
  - BIF (Jimenez et al., 2023)



Bifactor model

Second order factor model

Oblique factor model

Purpose

- To compare the performance (accuracy and bias) of seven recommended stopping rules under the second order (2ND), the bifactorial (BIF) and the oblique factor (OBL) models
  - OBL : Correlated pair of common factors
  - 2^{ND} : A single second order factor for common factor
  - BIF : A single general factor for items

METHOD

Stopping rules

- Parallel analysis (PA50, PA95)
  - Resample the eigenvalues of surrogated datasets with no factor (Identity matrix) and the same characteristics as the target dataset (same number of variables and subjects).
  - Based on PCA.
  - The first k^{th} empirical eigenvalues greater than 95^{th} or 50^{th} percentile simulated eigenvalues are retained.

Stopping rules

- Empirical Kaiser Criterion (EKC)
  - Derive expected eigenvalues from the Marchenko-Pastur distribution.

$$\lambda_j = \max\left(\frac{p\sum_{i=0}^{j}\lambda_i}{p-j-1}\left(1+\sqrt{\frac{p}{n}}\right)^2, 1\right)$$

$$\lambda_0 = \left(1+\sqrt{\frac{p}{n}}\right)^2$$

  - The first k^{th} empirical eigenvalues over 1 are retained.

Stopping rules

- Minimum Average Partial Correlation (MAP)
  - Remove factors step by step and examine the remaining (partial) correlations between variables.
  - The optimal number of factors is the number that minimizes the average squared partial correlations.

Stopping rules

- Sequential χ^2 model test (SMT)
  - Sequentially fit *k*-factor models using maximum likelihood (ML).
  - For each model, compare the model-implied covariance matrix to the sample covariance matrix using a χ^2 test.
  - Retain the first model that is not significantly different (non-significant χ^2).

Stopping rules

- Next Eigenvalue Sufficiency Test (NEST)
  - Resample the eigenvalues of surrogated datasets with the previous *k* factors
    - When $k = 0$, it is equivalent to PA95
  - Retain the k^{th} factor if the empirical eigenvalue is greater than the corresponding 95^{th} percentile simulated eigenvalues
  - Stop when the next factor's eigenvalue is not distinguishable from simulated ones (below the 95^{th} percentile)
  - Package Rnest

Stopping rules

- Out-of-sample prediction error (OSPE)
  - Split the date in halves (a training set and a test set)
  - Fit the *k*-factor model to the training set
  - Derive the model-implied covariance matrix
  - Cross-validate the model implied-covariance matrix to the test set
  - Choose the number of factors that minimizes out-of-sample prediction error
  - Packages fspe

Stopping rules

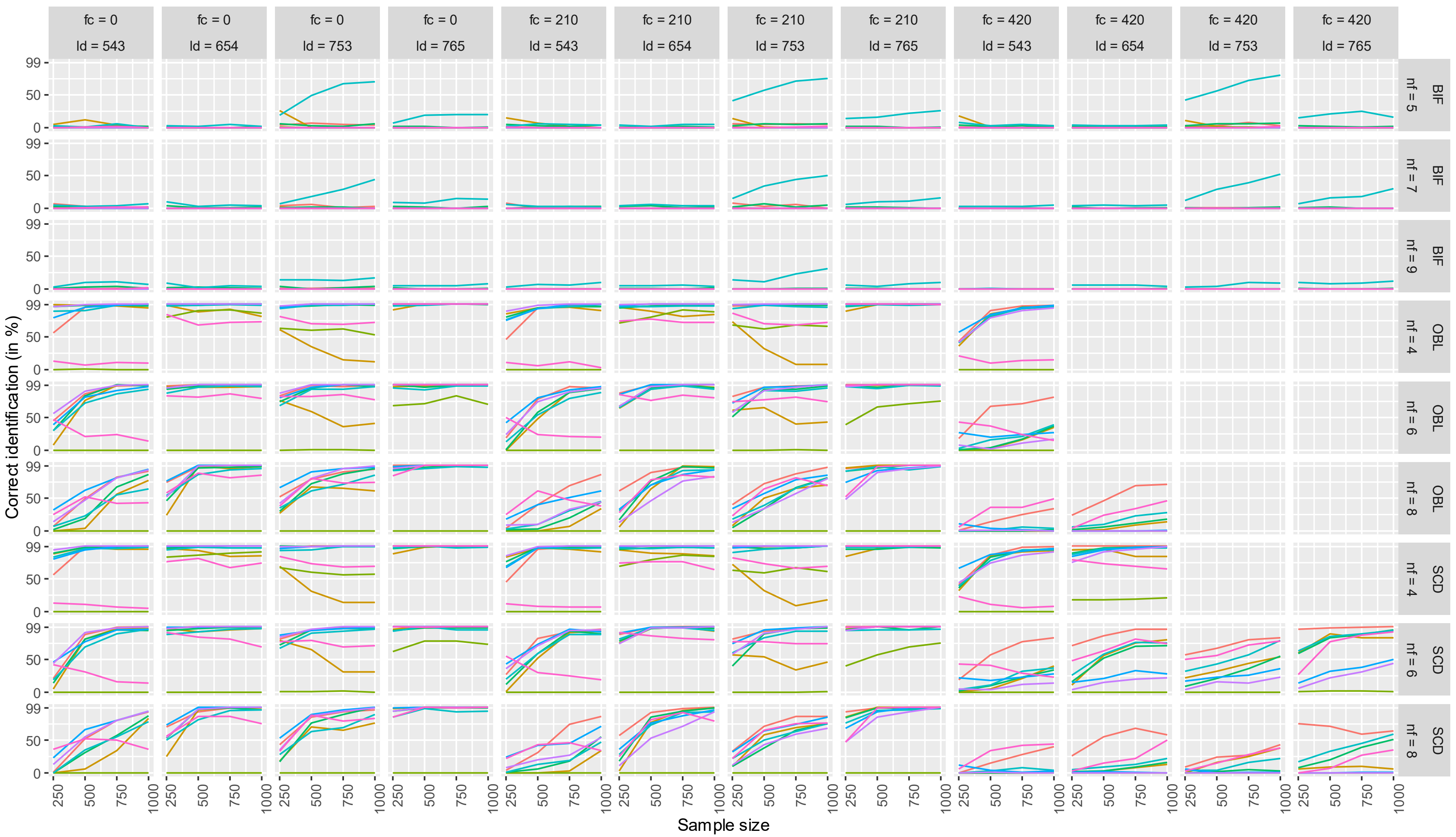
- Exploratory Graph Analysis (EGA)
  - Estimate the number of factors using a network approach
    - Variables are represented as nodes, and relationships as edges (partial correlations)
  - Use community detection algorithms to identify clusters of variables
  - The number of detected communities = number of factors
  - Package EGAnet

Conditions

- A Monte Carlo simulation
  - Sample size N = 250, 500, 750, 1000
  - 24 items with nf = 4, 6 and 8 factors
  - Combination of loadings ld = (.4, .5, .6), (.3,.5,.7), (.5,.6,.7), (.5,.4,.3)
  - Three population factorial model (BIF, 2ND, OBL) with there defining parameters, fc = (0), (.2,.1,0) and (.4,.2,0)
  - Each of the 432 conditions is repeated 100 times.
- Performance is evaluated in terms of accuracy (correct identification of dimensionality) and bias (tendency to over- or under-estimate dimensionality)

RESULTS

All results

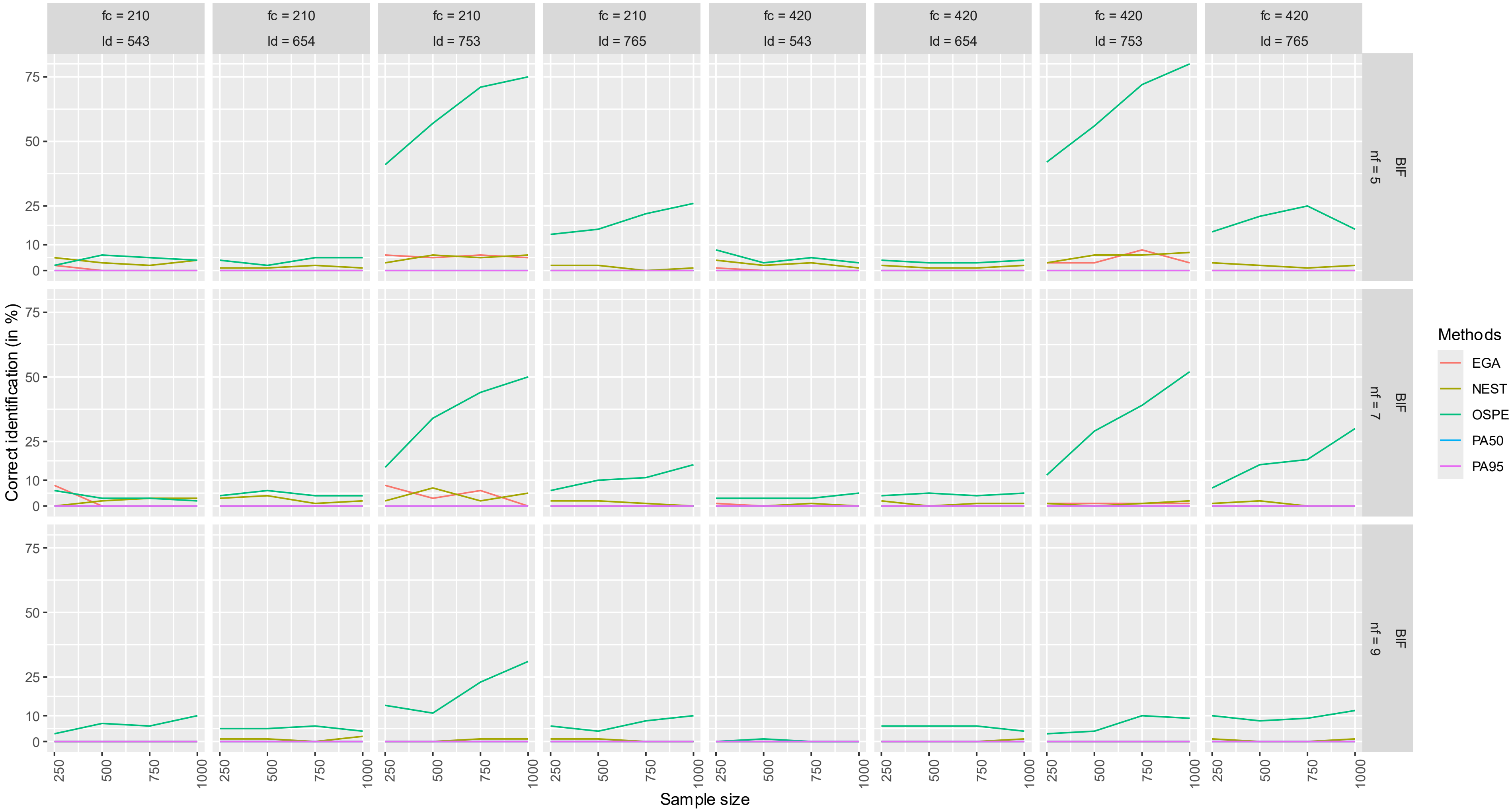


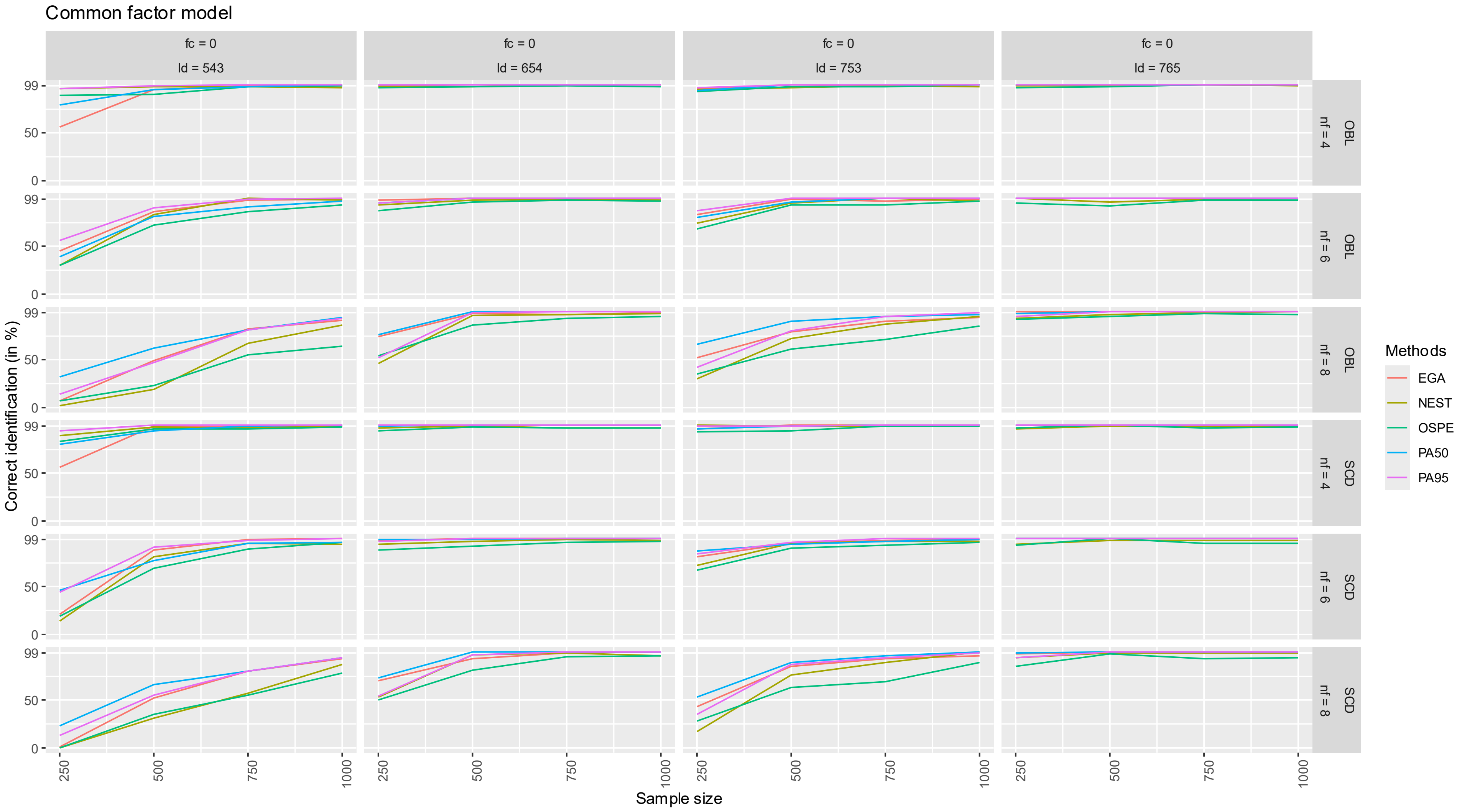
fc = 0
Id = 543
Id = 654
Id = 753
Id = 765
fc = 210
fc = 420
BIF
OBL
SCD
nf = 5
nf = 7
nf = 9
nf = 4
nf = 6
nf = 8
99
50
0
250
500
750
1000
Correct identification (in %)
Sample size
Methods
EGA
EKC
MAP
NEST
OSPE
PA50
PA95
SMT

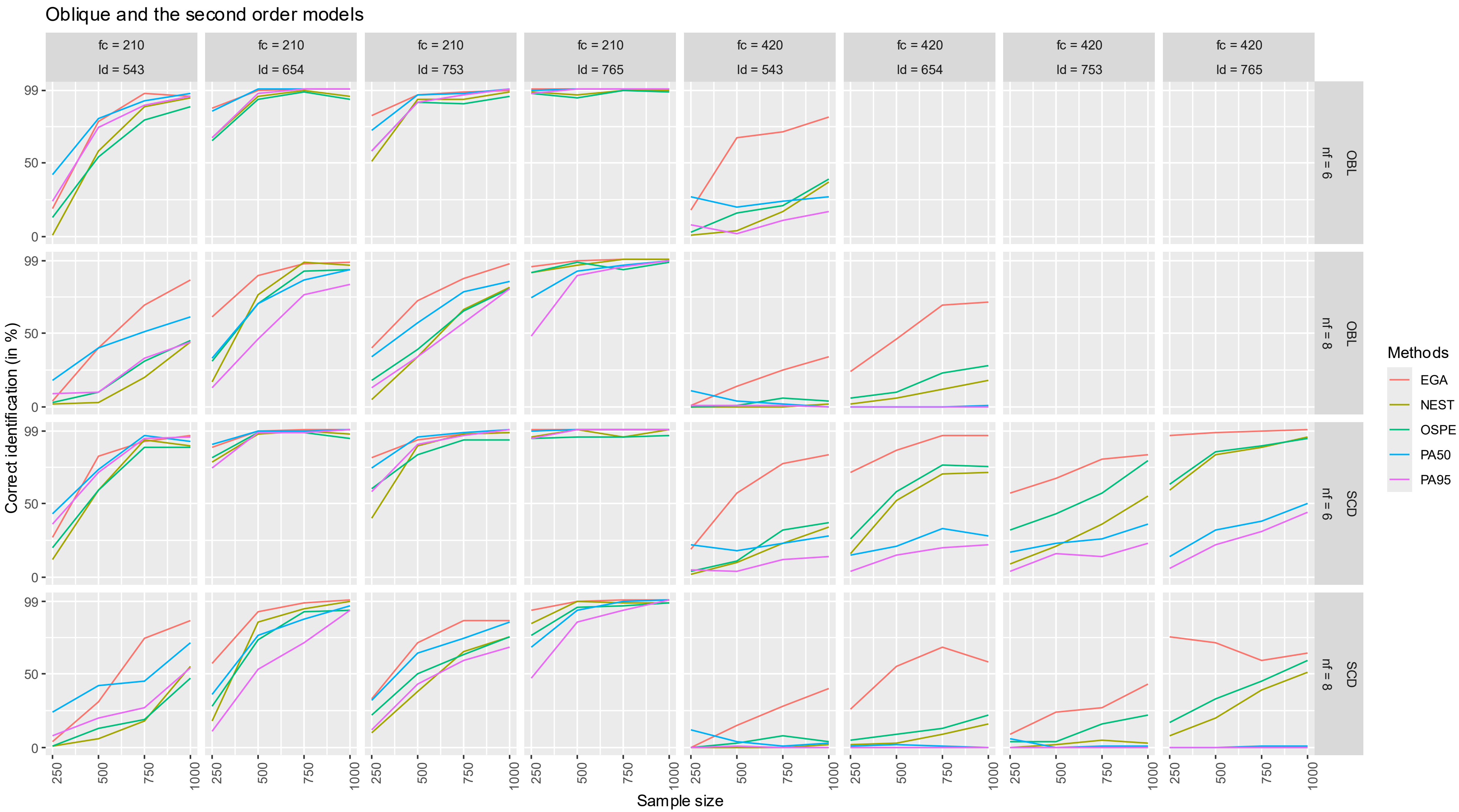
Results

- Poor performances of EKC, MAP, and SMT
- Unexpected results in bifactor models
  - OSPE performed the best
  - All other methods performed very pooly
- EGA performed the best overall
- PA50 and PA95 *performed well* in the fc = 0

Bifactor model

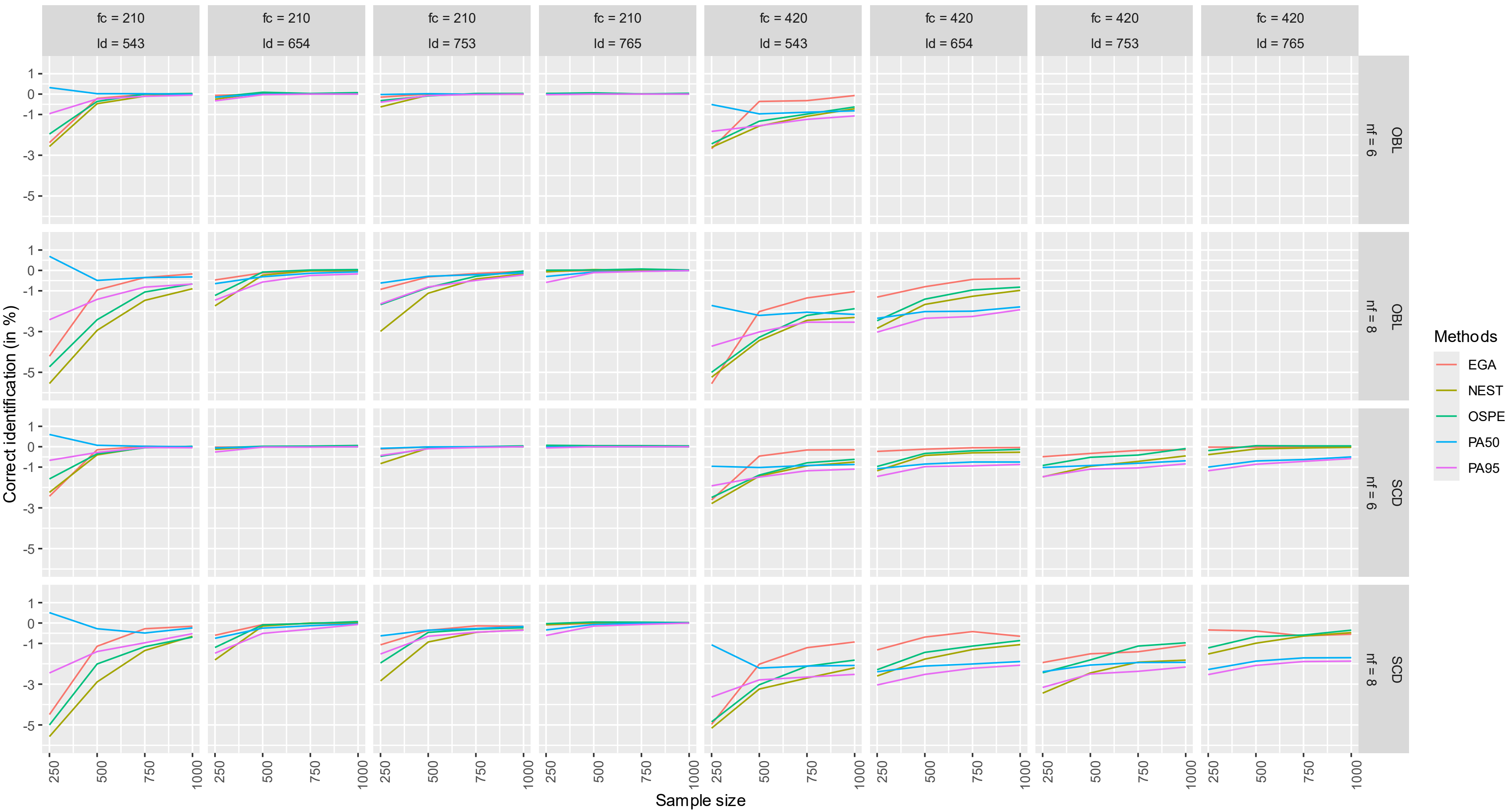



Common factor model
fc = 0
ld = 543
fc = 0
ld = 654
fc = 0
ld = 753
fc = 0
ld = 765
OBL
nf = 4
OBL
nf = 6
OBL
nf = 8
SCD
nf = 4
SCD
nf = 6
SCD
nf = 8
99
50
0
Correct identification (in %)
250
500
750
1000
Sample size
Methods
EGA
NEST
OSPE
PA50
PA95



Oblique and the second order models
fc = 210
ld = 543
fc = 210
ld = 654
fc = 210
ld = 753
fc = 210
ld = 765
fc = 420
ld = 543
fc = 420
ld = 654
fc = 420
ld = 753
fc = 420
ld = 765
OBL
nf = 6
OBL
nf = 8
SCD
nf = 6
SCD
nf = 8
Correct identification (in %)
Sample size
Methods
EGA
NEST
OSPE
PA50
PA95

Bias in oblique and the second order models



DISCUSSION

Discussion

- Limits
  - Factor structures with the same *expected* loadings
  - Equal *expected* eigenvalues for all factors
  - No cross-loadings

Recaps

- Poor performances of EKC, MAP, and SMT
- Unexpected results in bifactor models
  - OSPE performed the best
  - All other methods performed very pooly
  - Need more investigation
- EGA performed the best overall because there was no cross-loading
- PA50 and PA95 *performed opportunisticly well* when fc = 0

THANK YOU

References

- Caron, P.-O. (2025). Rnest : An R package for the Next Eigenvalue Sufficiency Test for factor analysis. *Multivariate Behavioral Research, 60*(5), 1062-1068. https://doi.org/10.1080/00273171.2025.2512343
- Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. PLOS ONE, 12(6), e0174035.
- Haslbeck, J. M. B., & van Bork, R. (2024). Estimating the number of factors in exploratory factor analysis via out-of-sample prediction errors. *Psychological Methods, 29*(1), 48-64. https://doi.org/10.1037/met0000528
- Jiménez, M., Abad, F. J., Garcia-Garzon, E., Golino, H., Christensen, A. P., & Garrido, L. E. (2025). Dimensionality assessment in bifactor structures with multiple general factors: A network psychometrics approach. *Psychological Methods, 30*(4), 770-792. https://doi.org/10.1037/met0000590