

# Multilevel SEM with pseudobalanced groups and small samples



Joop Hox  
Cora Maas

Utrecht University

2nd Amsterdam Multilevel Conference  
Amsterdam, March 30-31, 1999

# Structural Equation Modeling (SEM)



- Multivariate normality
  - data summed up perfectly by covariance matrix  $S$ 
    - + means  $M$
- Model for  $\Sigma$  contains parameters  $\theta$ 
  - factor loadings, path coefficients
  - estimated by Maximum Likelihood
  - standard errors, overall significance test
- Software
  - e.g., Lisrel, Eqs, Amos
- Submodel: Confirmatory Factor Analysis (CFA)



# Two level SEM

- Assume a population that contains groups
- Sampling at two levels:  
groups and individuals within groups
- Decompose population Total scores into (disaggregated)  
group means  $B$  and individual deviations  $W$
- In the population we have  $\Sigma_T = \Sigma_W + \Sigma_B$
- Decompose the sample scores
- In the sample we have  $S_T = S_W + S_B$



# Multilevel SEM

- Decompose the sample scores  $T$  into  $W$  (Within) and  $B$  (Between)
- Compute pooled within groups covariances  $S_{PW}$
- Compute between groups covariances  $S_B$
  
- Muthén has shown that
- $S_{PW} = \Sigma_W$  and  $S_B = \Sigma_W + n\Sigma_B$ 
  - with  $n$  equal to common group size

# Estimating multilevel SEM with standard SEM software



- $S_{PW} = \Sigma_W$ 
  - model it with Within model
- $S_B = \Sigma_W + n\Sigma_B$ 
  - model it with Within model + (scale factor n) \* Between model
- Use two-group option of standard SEM software
  - group I = Within, group II = Between
  - if groups are unequal, ignore that problem  
use average group size, call it pseudobalanced solution



# Our questions

- How accurate is the pseudobalanced solution?
  - parameter estimates
  - standard errors
  - overall model test
- Given unbalanced data
  - varying number of groups
  - varying group sizes
  - varying proportion of group level variance

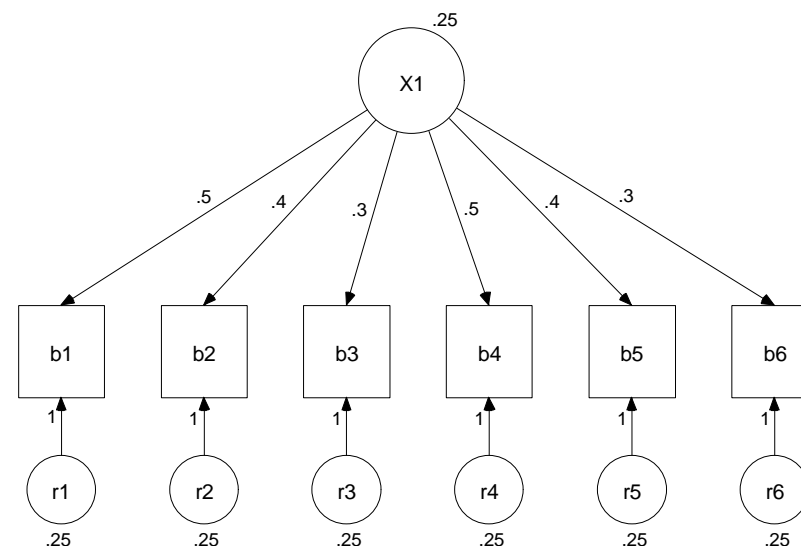
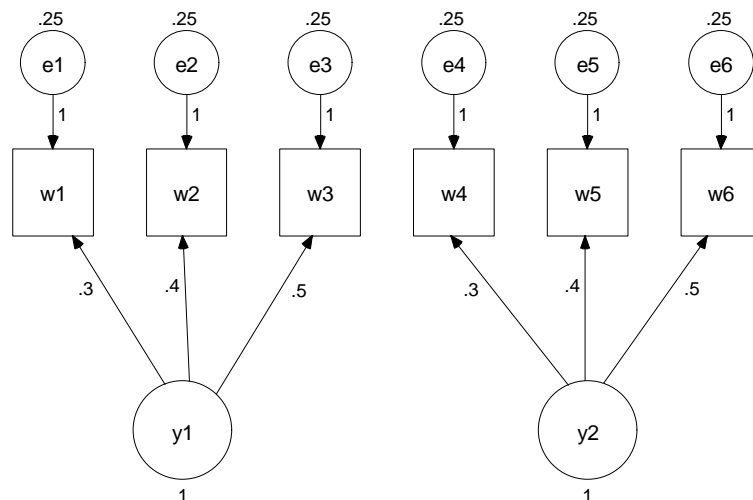


# The model

■ Two level Confirmatory Factor Analysis

■ Within model

■ Within model





# Simulation procedure

- Simulated conditions
  - *number of groups NG*
    - | 50 - 100 - 200
  - *group size (average) GS*
    - | 5/15 (10) - 10/30 (20) - 25/75 (50)
  - *observed variables group level variance (average ICC) ICC*
    - | variance of between factor is 0.25/0.50 (.20/.33)
- Simulation procedure: do per condition 1000 times
  - generate group and individual data, from model + condition
  - combine into total scores
  - compute  $S_{PW}$  and  $S_B$
  - estimate model: store estimates, s.e.'s and chi-square

# Results:

## Within model



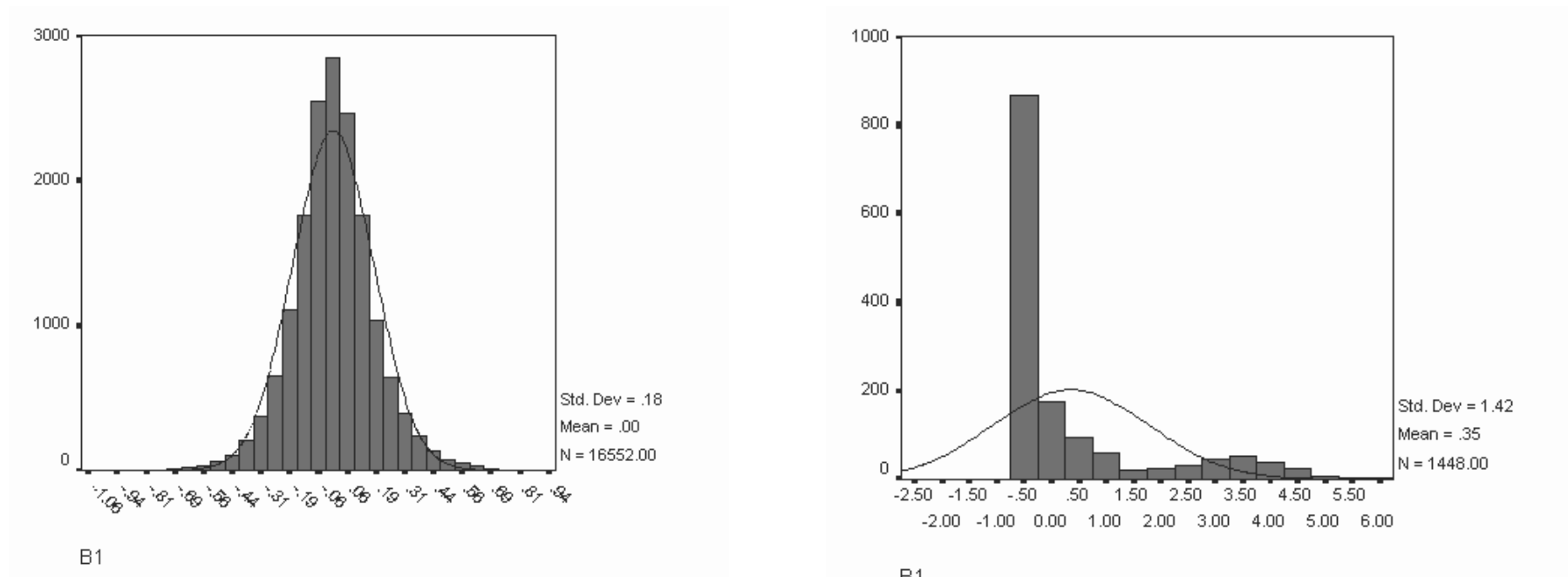
- No inadmissible solutions
  - all variance estimates  $> 0$
- Factor loadings (.3 .4 .5) mean bias 0.00005 (zero)
  - no differences across conditions
- Variances (.25) mean bias -0.0005 (zero)
  - 'significant' effect of sample sizes
  - e.g., for NG 50/100/200: -0.00095/-0.00049,-0.00018
- Coverage of 95% confidence intervals
  - smallest 95.13%, largest 96.03%
- Conclusion: No problem!

# Results: Between model



- Some inadmissible solutions
  - 8% of simulations have variance estimates  $< 0$
  - most of these  $\lll 0$
  - also factor loadings  $\ggg 1$
  - Note: program always converged
  
- Inspect distribution of inadmissible solutions
  
- Analyze effects of NGroups, GSize and ICC for admissible solutions

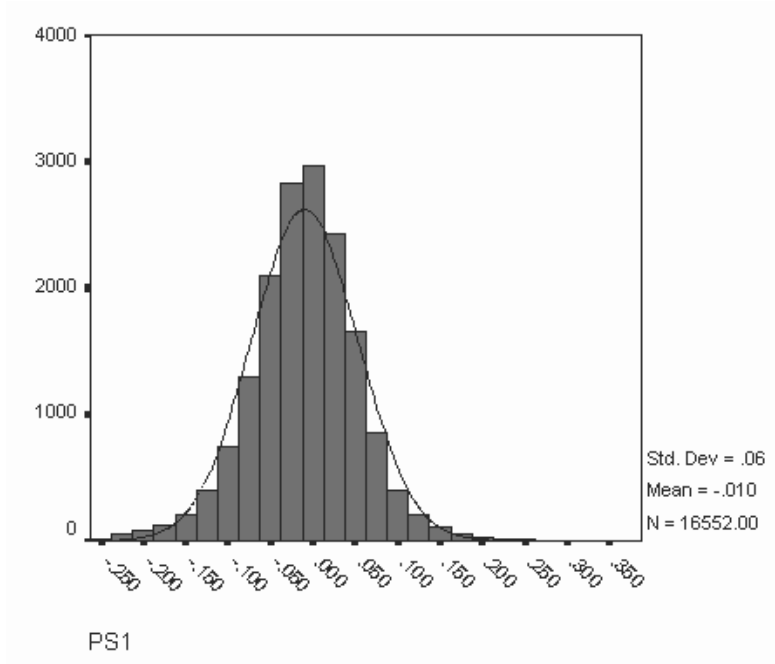
# Distribution of between loading b1



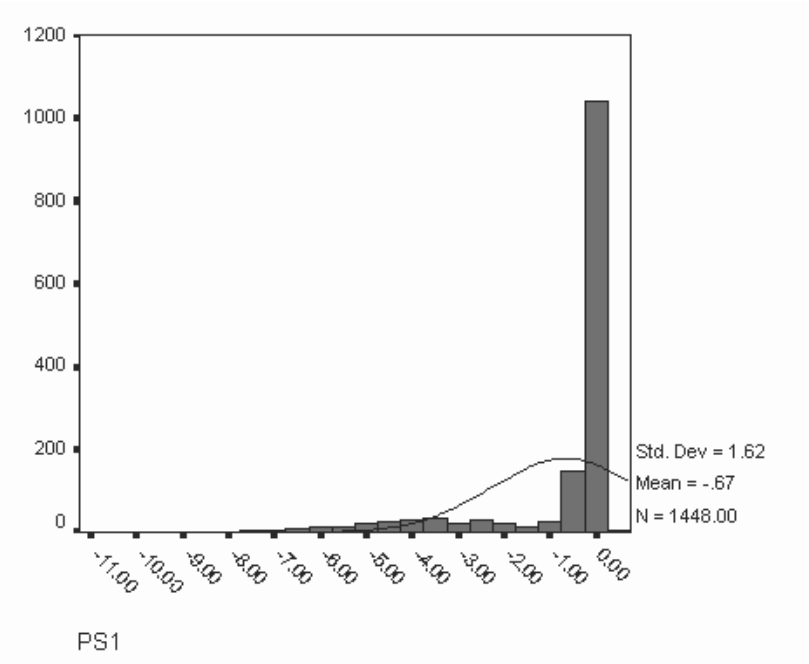
Bias of loading b1 for  
admissible solutions

Bias of loading b1 for  
inadmissible solutions

# Distribution of between variance ps1



Bias of variance ps1 for admissible solutions



Bias of variance ps1 for inadmissible solutions

# Between model: distribution of inadmissible solutions



## ■ Inadmissible solutions for different NG, GS and ICCMain effects:

■ NG	50/100/200:	18%/1%/0%
■ GS	10/20/50:	10%/0%/0%
■ ICC	lo/hi	14%/2%

## ■ Interactions:

■ NG/ICC lo	29%/10%/1%
■ NG/ICC hi	6%/1%/0%
■ GS/ICC similar pattern, less dramatic	
■ NG/GS and 3-way interaction significant but negligible	

## ■ Conclusion: we have a problem!

- Largest effect: NG, large GS and ICC compensate in part

# Results Between model, admissible solutions only: bias



- Factor loadings (.3 .4 .5) mean bias 0.004
  - effect of NG (bias 0.007/0.004/0.005)
  - effect of ICC (bias 0.01/0.00)
- Variances (.25) mean bias -0.01
  - effect of ICC (bias -0.013/-0.008)
  - effect of NG (bias -0.011/-0.010/-0.010)
- Conclusion: No problem!

# Results Between model, admissible solutions only: coverage



- Coverage of 95% confidence interval
  
- Factor loadings
  - smallest 89%, largest 94%
  - small effects of NG, ICC and GS, no real improvements
    - e.g., NG 50/100/200: 89%/90%/91%
  
- Variances
  - smallest 90%, largest 90%
  - small effects of NG, GS
    - e.g., NG 50/100/200: 88%/91%/91%
  
- Conclusion: standard errors too small
  - but not totally misleading

# Results Between model, all solutions, overall chi-square test



- Mean chi-square 19.6 (expected 18): too high
  - mean  $\chi^2$  significantly  $> 18$  in all conditions
  - smallest difference when NG=50, GS=10, ICC low
- Overall percentage rejections 8.9%
  - admissible solutions 9.0%, inadmissible solutions 7.3%
  - effects of ICC and NGroup\* ICC
    - ICC lo/hi: 7.9%/9.9%
    - ICC lo: NG 50/100/200 6.6%/7.9%,9.1%
    - ICC hi: NG 50/100/200 9.8%/10.5%/9.4%
- Conclusion: too many rejections
  - large NG, GS and ICC do not help here



# Conclusions

- Within model, no problems
- Between model, problems:
  - inadmissible solutions (usually grossly wrong)
  - estimates unbiased, but
  - standard errors too small
  - chi-squares too high
  - result: models too complex
- Especially with small number of groups
- But results fairly close
  - e.g. if we choose  $\alpha=0.25$ , operating  $\alpha \approx 0.05$



# Some references

- Muthén, B. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, 22, 376-398.
- McDonald, R.P. (1994). The bilevel reticular action model for path analysis with latent variables. *Sociological Methods & Research*, 22, 399-413.
- Hox, J.J. (1993). Factor analysis of multilevel data. Gauging the Muthén model. In: J.H.L. Oud & R.A.W. van Blokland-Vogeleesang (eds.). *Advances in longitudinal and multivariate analysis in the behavioural sciences*. Nijmegen, NL: ITS.
- Hox, J.J. (1995). *Applied multilevel modeling*. Amsterdam: TT Publikaties.

(Hox references downloadable from <http://www.fsw.ruu.nl/ms/jh>)