MEASURING EXTREME RESPONSE STYLE USING NONLINEAR SEM

MEIKE MORREN
VU UNIVERSITEIT AMSTERDAM, AMSTERDAM NETHERLANDS

Contents

- ERS
  - Quadratic pattern
  - Current methods
- Nonlinear SEM
  - Unconstrained (LMS method: Moosbrugger et al 2009)
  - Constrained approaches (Product indicators: Algina & Moulder 2001, Marsh et al 2004)
- Results
  - Simulated data (multinominal logit model)
  - Model fit & parameter estimates
  - Bias by ERS
- Recommendations & future research

INTRODUCTION

Extreme Response Style

- Tendency to select (or avoid) the endpoints on a Likert scale

TA	A	N	D	TD

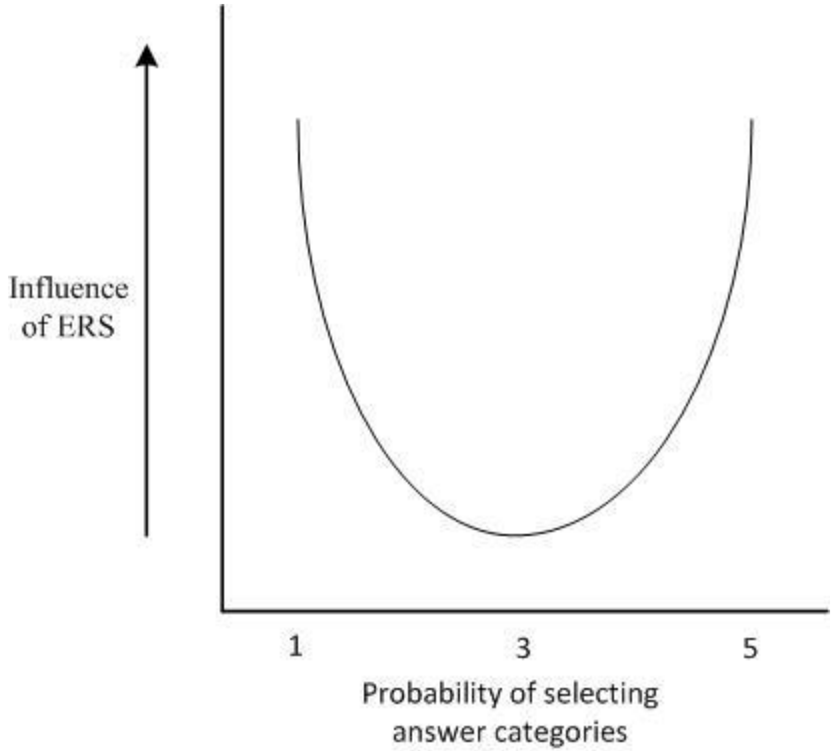
- Totally agree (TA) or totally disagree (TD)

Extreme Response Style

- If systematically across many items that differ in content,
  - i.e. the reason for selecting these categories is unlikely to be related to content
- The tendency to select the extremes results in quadratic relationship between observed response pattern and latent response style factor

TA	A	N	D	TD
TA	A	N	D	TD
TA	A	N	D	TD
TA	A	N	D	TD
TA	A	N	D	TD
TA	A	N	D	TD

Extreme Response Style



The more one is subject to ERS, the more one prefers the endpoints and avoids the midpoint/adjacent categories

Assumption: the items are unrelated in content!

Current methods to detect & correct for ERS

- Latent variable methods
  - Response style is measured by latent variable relating to two (or more) unrelated subsets of items

    CFA: Billiet & McClendon (2000); Cheung & Rensvold (2000)

    LCFA: Moors (2004), Kieruj & Moors (2010), Kankaras & Moors (2009), Morren, Gelissen & Vermunt (2011)

    LCA: Eid & Rauber (2000), Aichholzer (2013), Austin et al (2006)

    IRT: De Jong et al (2008), Bolt & Johnson (2009), Rossi et al (…)

- Sum score methods
  - Response style is measured by summing extreme responses across many unrelated item

    Steenkamp & Baumgartner (1998), Baumgartner & Steenkamp (2001), Johnson et al (2006), Harzing (2006), He & Vijver (2013)

- RIRSMACS
  - Response style is measured by three sumscores calculated across sets of unrelated items (the random indicators)

    Weijters (2006,2008), Weijters, Schillewaert & Geuens (2008), Thomas, Abts & Vander Weyden (2014, 2013), Meade & Craig (2012)

Why another method?

- All previous methods rely on assumption of unrelated items or unrelated item subsets which requires:

  a) To include a set of unrelated items in the survey (RIRSMACS) (27 items)
  b) To recode the items (sum-score method) (No longer possible to measure the substantive trait)
  c) To search for two or more unrelated subsets (latent variable methods)

The direct estimation of the quadratic effect caused by ERS and the linear effect of substantive factor facilitates estimation of (and correction for) ERS in related items or a single subset of items

  a) Making it useful for applied researchers
  b) Easier to apply to secondary datasets
  c) Applicable to surveys measuring related subjects, such as the Big Five

NONLINEAR SEM

Literature nonlinear SEM

- Constrained approach

  Nonlinear parameter constraints necessary

  (i.e. product indicator parameters are functions of the loadings and variances of the linear indicators)

  - Product indicators
    - Kenny & Judd, 1984; Joreskog & Yang, 1996, Jaccard & Wan, 1995 ; Algina & Moulder, 2001, Wall & Amemiya, 2000
  - Two-step approach
    - Ping, 1995, 1996

- Unconstrained approach
  - Product indicators – mean centering (Marsh, Wen & Hau, 2004)

- Latent Moderated Structural Equations (LMS, Klein & Moosbrugger, 2000, 2008)

- Quasi-Maximum Likelihood (QML, Klein & Muthén, 2007)

Literature review

- Comparison among approaches reveal that:
  - LMS and Joreskog & Yang's approach outperform two-step approach in terms of biased estimators (even if latent variables are highly correlated) (Kelava et al, 2008)
  - LMS and QML outperform unconstrained and constrained approaches in terms of standard errors, unconstrained overestimates latent variances (Moosbrugger, Schermelleh-Engel, Kevala, & Klein, 2009)
  - Unconstrained approach by Marsh shows least amount of bias, closely followed by LMS. However, when small sample size & poor reliability, then LMS outperforms unconstrained approach (Harring, Weiss & Hsu, 2012)

Contributions of this study

To measure quadratic factor-to-items relationships

Using LMS:

- Current approaches focus on relationships among latent variables and either only interaction terms (Joreskog & Yang; Marsh, Algina & Moulder, Wall & Amemiya), interaction and quadratic effects (Kevala et al, Moosbrugger et al., Lee, Song & Poon) or only quadratic effects (Harring, Weis & Hsu)

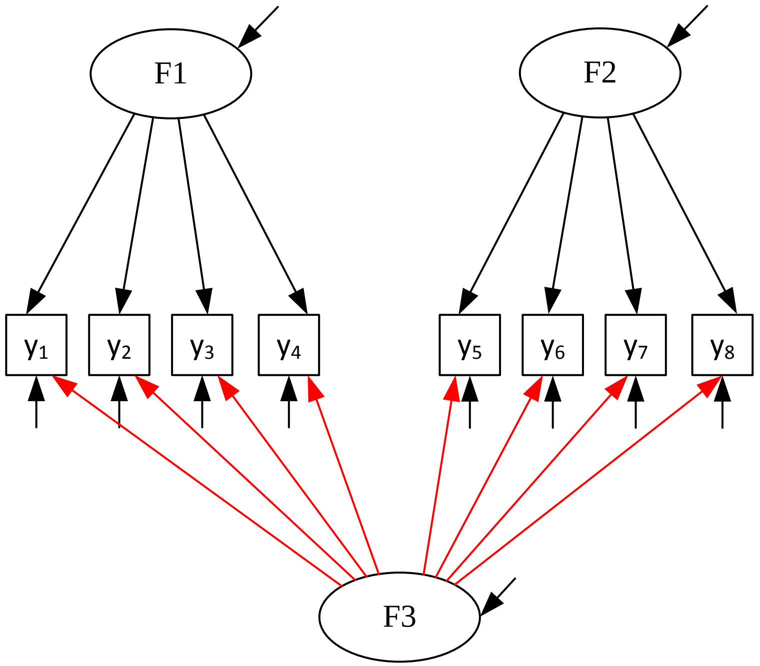
Using product indicator approach:

- Constrained approaches focus on relationships between latent and observed variables but only take into account interaction effects (Marsh et al, Algina & Moulder, Joreskog & Yang) or include only a quadratic relationship between one of the items in the model (Bauer, 2005)

GENERATED DATASET

Model: graphical representation

- The red arrows indicate quadratic relationships
- Identification by setting latent variances to 1
- Three conditions:
  - The substantive factor has a stronger influence than style factor
  - The substantive factor has a equal influence to style factor
  - The substantive factor has a weaker influence than style factor



Model: Nominal response model (Bock)

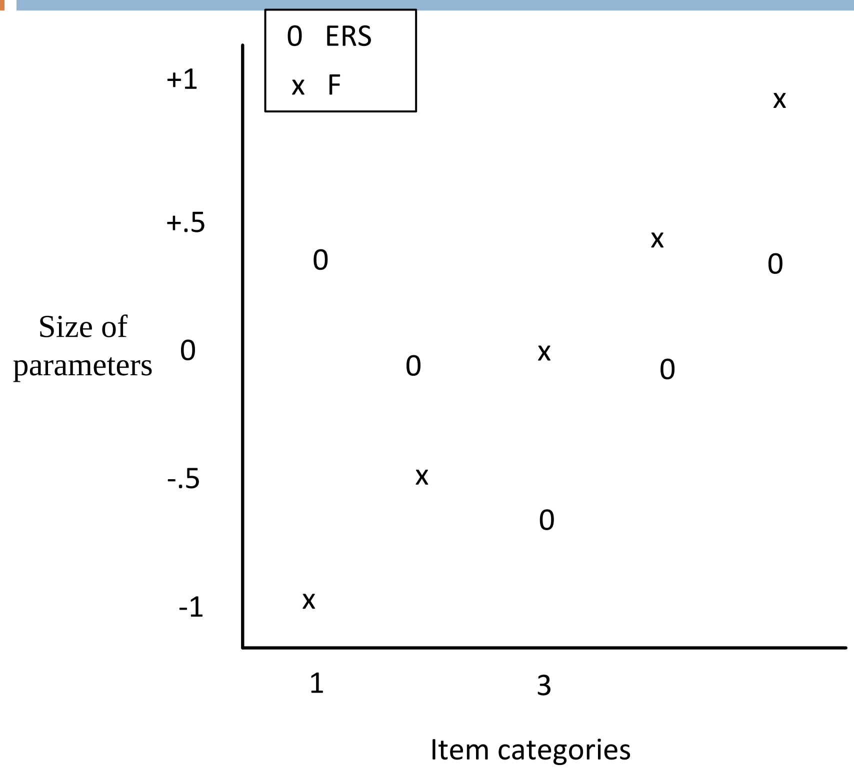
- Three continuous, normally distributed latent factors
- 8 observed variables, each having 5 categories
- Uncorrelated latent factors

- Latent Gold 4.5, N=4000 for each generated dataset

- Nominal response model:

$$P(Y_{ij} = c|F1, F2, ERS) = \frac{\exp\left(\alpha_{0jc} + \beta_{1jc}\, F_{1i} + \beta_{2jc}\, F_{2i} + \beta_{3jc} ERS_i\right)}{\sum_{\mathbf{d=1}}^{\mathbf{C}} \exp\left(\alpha_{0jd} + \beta_{1jd}\, F_{1i} + \beta_{2jd}\, F_{2i} + \beta_{3jd} ERS_i\right)}$$

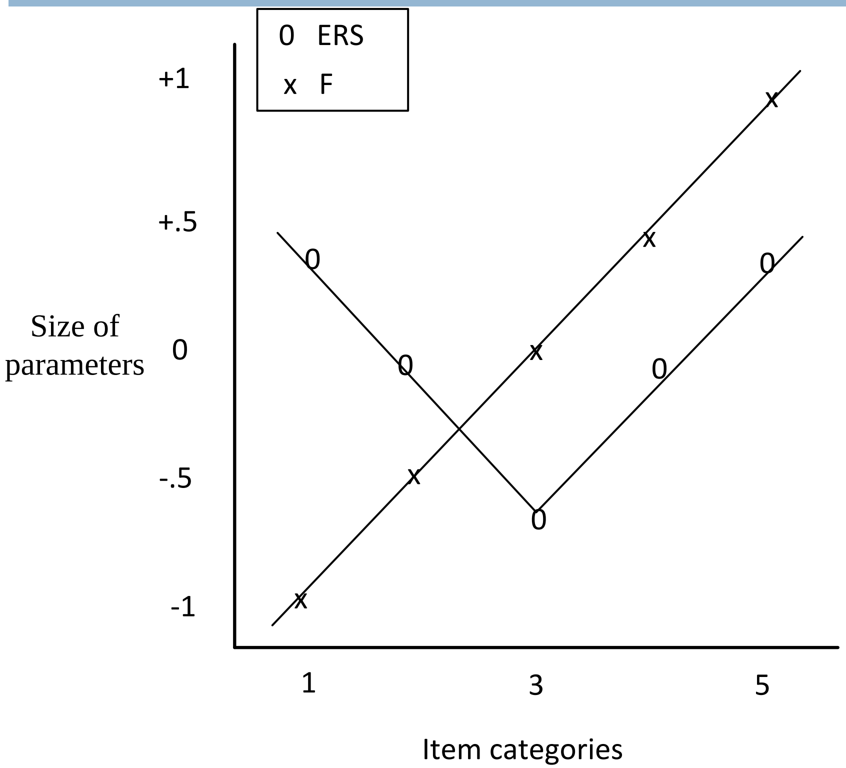
Simulated parameter values



	Cat	F < ERS	F = ERS	F > ERS
F	1	-.6	-1	-1
	2	-.3	-.5	-.5
	3	0	0	0
	4	.3	.5	.5
	5	.6	1	1
ERS	1	.4	.4	.25
	2	-.1	-.1	-.05
	3	-.6	-.6	-.4
	4	-.1	-.1	-.05
	5	.4	.4	.25

A CFA model on these data would result in factor loadings: .55 in condition 1; .75 in condition 2 and 3

Simulated parameter values



A CFA model on these data would result in factor loadings: .55 in condition 1; .75 in condition 2 and 3

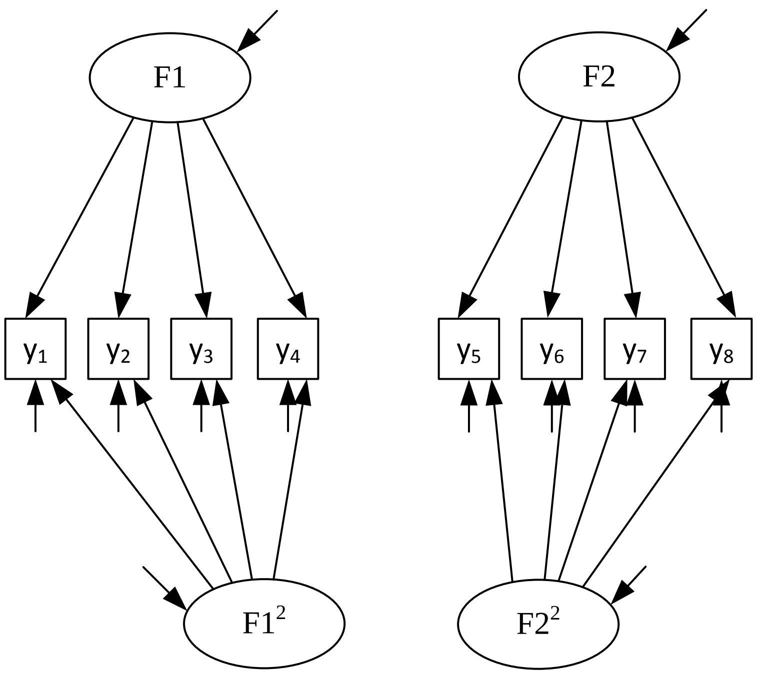
	Cat	F < ERS	F = ERS	F > ERS
F	1	-.6	-1	-1
	2	-.3	-.5	-.5
	3	0	0	0
	4	.3	.5	.5
	5	.6	1	1
ERS	1	.4	.4	.25
	2	-.1	-.1	-.05
	3	-.6	-.6	-.4
	4	-.1	-.1	-.05
	5	.4	.4	.25

RESULTS

Generated dataset

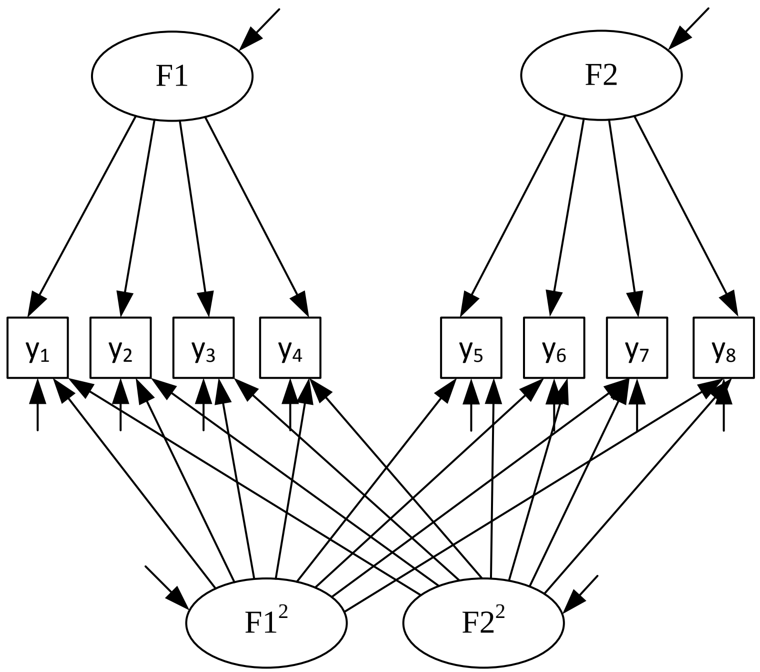
LMS approach (Moosbrugger et al 2009)

- Two quadratic latent variables are calculated, each relate to one item subset
- No correlation between the two latent quadratic terms was possible in Mplus => misspecification
- Other modeling approaches
  - Quadratic terms affect all items
  - One latent style factor affecting all items; quadratic terms affecting the latent style factor



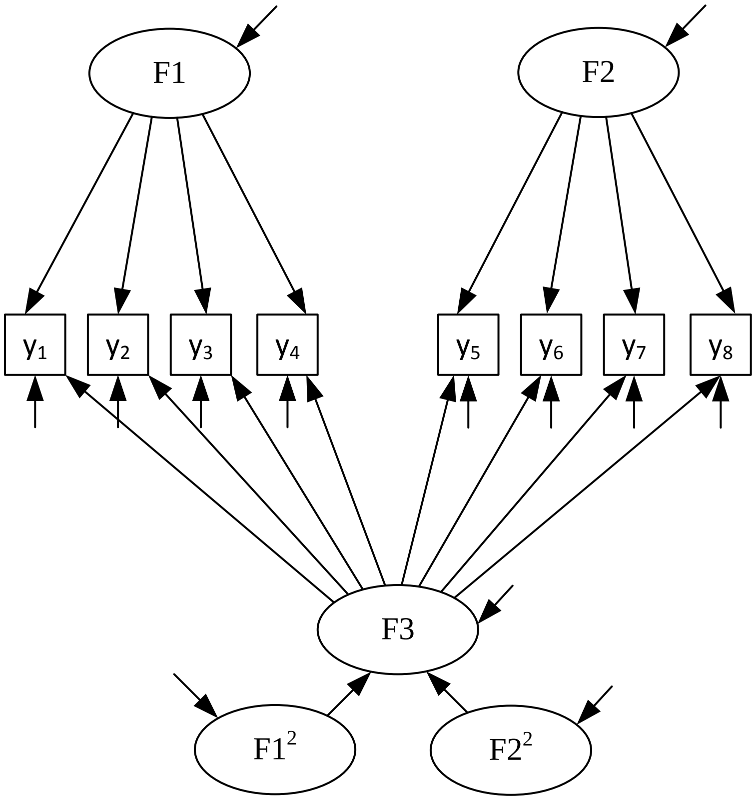
LMS approach (Moosbrugger et al 2009)

- Two quadratic latent variables are calculated, each relate to one item subset
- No correlation between the two latent quadratic terms was possible in Mplus => misspecification
- Other modeling approaches
  - Quadratic terms affect all items
  - One latent style factor affecting all items; quadratic terms affecting the latent style factor



LMS approach (Moosbrugger et al 2009)

- Two quadratic latent variables are calculated, each relate to one item subset
- No correlation between the two latent quadratic terms was possible in Mplus => misspecification
- Other modeling approaches
  - Quadratic terms only affect subset of items
  - One latent style factor affecting all items; quadratic terms affecting the latent style factor



Mplus syntax

```
DATA: FILE IS simulatedCond1.txt;

VARIABLE:
  NAMES ARE simnr y1-y8;
  USEVARIABLES ARE y1-y8;

ANALYSIS:
TYPE= RANDOM;
  ALGORITHM=INTEGRATION;
```

```
MODEL:
  f1@1;
  f1 by y1*;
  f1 by y2-y4;
  f1f1 | f1 XWITH f1;

  f2@1;
  f2 by y5*;
  f2 by y6-y8;
  f2f2 | f2 XWITH f2;

  y5-y8 ON f2f2;
  y1-y4 ON f1f1;

OUTPUT: TECH1 TECH8;
```

Unconstrained approach (Marsh et al)

- Product indicators
  - Subtract the intercepts from the observed values, multiply these centered variables

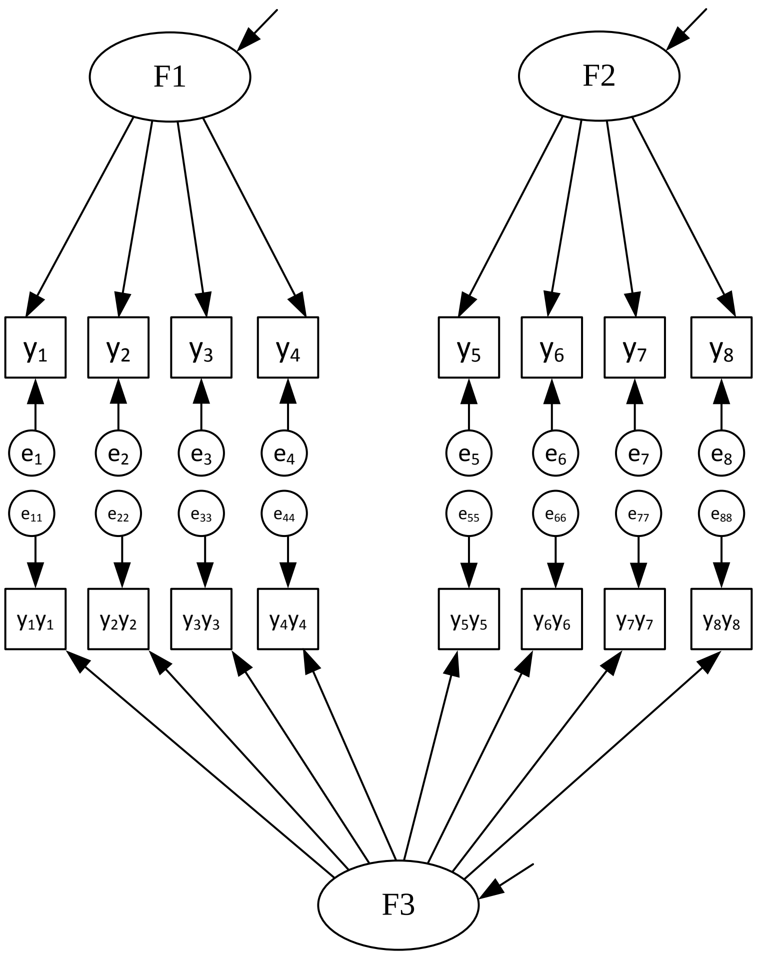
- **Message Mplus:**

THE MODEL ESTIMATION TERMINATED NORMALLY

- But….

THE STANDARD ERRORS OF THE MODEL PARAMETER ESTIMATES MAY NOT BE TRUSTWORTHY FOR SOME PARAMETERS DUE TO A NON-POSITIVE DEFINITE FIRST-ORDER DERIVATIVE PRODUCT MATRIX. THIS MAY BE DUE TO THE STARTING VALUES BUT MAY ALSO BE AN INDICATION OF MODEL NONIDENTIFICATION.THE CONDITION NUMBER IS -0.407D-15. PROBLEM INVOLVING PARAMETER 36 *[variance y01]*.

- Probably related to the ordinal nature of the product indicators (negative and positive values become very close)



Mplus syntax

```
DATA: FILE IS simulatedCond1_nolabels.txt;

DEFINE:
y01 = y1-2.985;
y02 = y2-3.013;
y03 = y3-2.998;
y04 = y4-2.966;
y05 = y5-2.977;
y06 = y6-3.019;
y07 = y7-3.032;
y08 = y8-3.042;
y1y1=y01*y01; y2y2=y02*y02;
y3y3=y03*y03; y4y4=y04*y04;
y5y5=y05*y05; y6y6=y06*y06;
y7y7=y07*y07; y8y8=y08*y08;

VARIABLE:
  NAMES ARE sim y1-y8;
  USEVARIABLES ARE y01-y08 y1y1 y2y2 y3y3 y4y4
y5y5 y6y6 y7y7 y8y8;

ANALYSIS:
  TYPE = GENERAL;
```

```
MODEL:
  f1@1;
  f1 by y01*;
  f1 by y02-y04;

  f2@1;
  f2 by y05*;
  f2 by y06-y08;

  f1 with f2;

  f3@1;
  f3 by y1y1* y2y2 y3y3 y4y4 y5y5 y6y6 y7y7 y8y8;
  f3 with f2@0 f1@0;

OUTPUT: MODINDICES(ALL) TECH1 TECH4;
```

Results (condition 1: F < ERS)

- Factor loadings
- Only depicted for 4 items: same pattern appears in the parameters relating to item 5-8

	item	Uncorrected CFA	LMS	Unconstrained
F	1	.552	.553	.552
	2	.577	.577	.577
	3	.574	.574	.574
	4	.503	.503	.503
ERS	1		.004	.548
	2		-.008	.566
	3		-.015	.534
	4		-.023	.625
BIC		113704.70 (d.f. = 25)	113769.73 (d.f. = 33)	237812.34 (d.f. = 49)
X^2		21.38^{ns}		135.54*

Results (condition 1: F < ERS)

Main findings

1. Factor loadings do not change after correcting for ERS
2. The ERS relationships cannot be detected by the LMS method
3. The model fit decreases when controlling for ERS

	item	Uncorrected CFA	LMS	Unconstrained
F	1	.552	.553	.552
	2	.577	.577	.577
	3	.574	.574	.574
	4	.503	.503	.503
ERS	1		.004	.548
	2		-.008	.566
	3		-.015	.534
	4		-.023	.625
BIC		113704.70 (d.f. = 25)	113769.73 (d.f. = 33)	237812.34 (d.f. = 49)
X^2		21.38^{ns}		135.54*

Results (condition 2: F = ERS)

Main findings

1. Factor loadings do not change after correcting for ERS
2. The ERS relationships cannot be detected by the LMS method
3. The ERS parameters are fairly similar to condition 1 (as simulated)
4. The model fit decreases when controlling for ERS

	item	Uncorrected CFA	LMS	Unconstrained
F	1	.718	.718	.718
	2	.860	.859	.860
	3	.786	.787	.787
	4	.776	.776	.776
ERS	1		-.010	.597
	2		.006	.551
	3		.006	.490
	4		-.002	.570
BIC		113553.28 (d.f. = 25)	113619.22 (d.f. = 33)	237840.13 (d.f. = 49)
X^2		10.62^{ns}		190.57***

Results (condition 3: F > ERS)

Main findings

1. Factor loadings do not change after correcting for ERS
2. The ERS relationships cannot be detected by the LMS method
3. Factor loadings differ from condition 2 (not simulated)
4. The ERS parameters are lower than condition 1 & 2 (as simulated)
5. The model fit decreases (increases) when correcting for ERS in unconstrained method (LMS)

	item	Uncorrected CFA	LMS	Unconstrained
F	1	.806	.806	.805
	2	.796	.796	.796
	3	.829	.829	.829
	4	.779	.779	.779
ERS	1		-.006	.444
	2		-.007	.392
	3		.000	.356
	4		-.023	.512
BIC		113704.70 (d.f. = 25)	113325.64 (d.f. = 33)	237875.37 (d.f. = 49)
X^2		21.38^{ns}		113.11^{ns}

ADDITIONAL ANALYSIS

SPVA dataset

SPVA data

Main findings

1. The unconstrained method detects quadratic effect, LMS to very small degree
2. LMS method leads to a improved model fit when correcting for ERS
3. The factor loadings remain fairly similar across methods (corrected or uncorrected for ERS)

	item	Uncorrected CFA	LMS	Unconstrained
F	1	.811	.795	.798
	2	.650	.636	.649
	3	.608	.622	.612
	4	.479	.498	.482
	5	.497	.496	.504
ERS	1		.009	-1.117
	2		-.031	-.831
	3		.072	-805
	4		.075	-.459
	5		-.076	-.602
BIC		98343.18 (d.f. = 31)	98225.05 (d.f = 41)	310278.32 (d.f. = 63)

Conclusions

1. The factor loadings related linearly to the factor are not affected by ERS

   but correction remains necessary for estimating the factor means, correlations among latent factors or group differences in factor means

2. PI approach might be useful in detecting ERS but leads to a worse fit in the model

3. PI approach is able to detect the quadratic factor-to-indicator effect simulated, but the estimates do not relate correctly to the simulated effect

4. LMS method does not detect the quadratic effect that is commonly detected in a multinomial response model

Conclusions

1. The factor loadings related linearly to the factor are not affected by ERS

   but correction remains necessary for estimating the factor means, correlations among latent factors or group differences in factor means

2. PI approach might be useful in detecting ERS but leads to a worse fit in the model

3. PI approach is able to detect the quadratic factor-to-indicator effect simulated, but the estimates do not relate correctly to the simulated effect

4. LMS method does not detect the quadratic effect that is commonly detected in a multinomial response model

LMS estimates factor-to-factor relationships that are quadratic, not factor-to-indicator relationships

ADDITIONAL ANALYSIS

Data Bauer (2005)

Additional analysis on data Bauer (2005)

Simulated dataset N=400

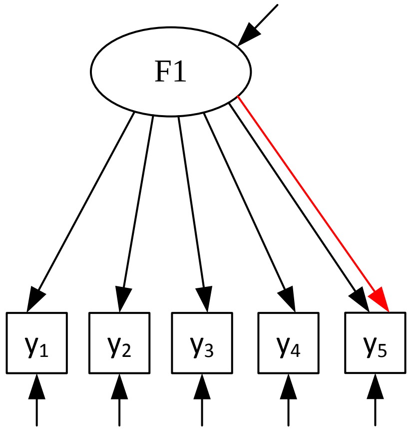
5 items, continuous, one factor

1 item relates linearly as well as quadratic

to one latent factor

Mplus syntax:

```
[x1@0];
f1 by x1@1;
f1 by x2-x5;
f1f1 | f1 XWITH f1;
x5 ON f1f1;
```



- Bauer estimated data in SAS, and estimated the quadratic relationship by a single parameter (in the multinomial approach, there are multiple category parameters)

Results on dataset Bauer (2005)

Main findings

- LMS provides a different estimate than unconstrained approach
- Unconstrained approach does not correct for biasing infleunce of quadratic effect on the linear parameter relating item 5 to the factor

	item	LG Uncorrected CFA	Mplus Uncorrected CFA	Mplus LMS	Mplus Unconstrained
linear	1	1	1	1	1
	2	.994	.964	.965	.965
	3	.990	.960	.961	.960
	4	.982	.953	.955	.953
	5	.889	.862	.973	.862
quadr	5			-.109	.296
BIC		4295.19	4369.53	4338.08	6064.98
X^2			101.30***		107.71***

Results multinomial model (LatentGold)

Mixed modeling:

- Categorize quadratic variable (into 6 categories, based on standard deviation 3, 2, and 1)
- Relate categorized variable to F1 nominally as well as linearly
- Category parameters show linear pattern (increase over categories)

Findings LG

- Model fit improves when correcting for quadratic effect
- Quadratic effect only appears in intercepts (not shown here)
- Biased factor loadings in LG approach

	item	LG Uncorrected CFA	LMS (simulated model)	LG Mixed modeling
linear	1	1	1	1
	2	.994	.965	.980
	3	.990	.961	.979
	4	.982	.955	.975
	5	.889	.973	.931
quadr	1 2 3 4 5 6		-.109	-8.887 -5.547 -1.606 2.736 5.667 7.637
BIC		4295.19	4338.08	4369.53
X^2				101.30***

Conclusions

1. LMS approach suitable to detect quadratic factor-to-indicator relationships

   IF the same latent factor affects the variable linearly as well as in a quadratic way

2. LMS is not well equipped to detect a quadratic effect of another latent factor on the same set of items (as is the case in ERS)

3. PI approach might be useful in correcting for ERS but leads to a worse fit in the model

4. The factor loadings in CFA are unaffected by ERS

   but correction is necessary for estimating the factor means, correlations among latent factors or group differences in factor means