Specification testing with Heywood cases: Can standard errors be trusted?

Stas Kolenikov

Department of Statistics
University of Missouri-Columbia

Joint work with Kenneth Bollen (UNC)
NSF support from SES-0617193 with funds from SSA

May 22, 2009

What is a Heywood case?

Heywood (1931) considered characterization of the correlation matrices:

$$\begin{pmatrix} 1 & r_1 r_2 & r_1 r_3 & \dots & r_1 r_n \\ r_1 r_2 & 1 & r_2 r_3 & \dots & r_2 r_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_1 r_n & r_2 r_n & r_3 r_n & \dots & 1 \end{pmatrix}, r_1 \geq r_2 \geq \ldots \geq r_n \geq 0$$

He showed that for this scheme

$$r_1^2 \leq 1 + \frac{1}{\frac{r_2^2}{1-r_2^2} + \frac{r_3^2}{1-r_3^2} + \dots \frac{r_n^2}{1-r_n^2}}$$

and hence r_1 can be greater than 1, which was an extension on the earlier results by Spearman and Garnett.

Causes of improper solutions

- Outliers (Bollen 1987).
- Non-convergence and underidentification (Van Driel 1978, Boomsma & Hoogland 2001).
- Empirical underidentification (Rindskopf 1984).
- Structurally misspecified models (Van Driel 1978, Dillon, Kumar & Mulani 1987, Sato 1987, Bollen 1989, Kolenikov & Bollen 2008).
- Sampling fluctuations (Van Driel 1978, Boomsma 1983, Anderson & Gerbing 1984).

Hypotheses on variances

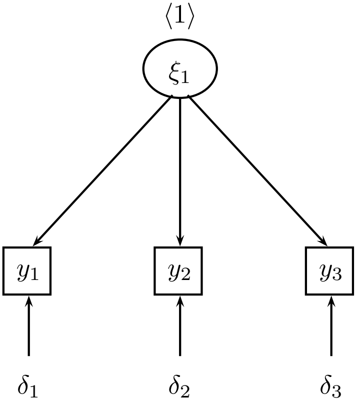
$H_+ : \sigma^2 > 0$	“Normal” situation
$H_0 : \sigma^2 = 0$	Perfect indicator (!?) or no factor present
$H_- : \sigma^2 < 0$	Misspecified model

	Sample estimate	
Population value	$\hat{\sigma}^2 > 0$	$\hat{\sigma}^2 < 0$
$\sigma^2 > 0$	☺	??
$\sigma^2 = 0$	☺	??
$\sigma^2 < 0$	Type I error	Misspecification detected!

Research question: testing specification, as expressed via

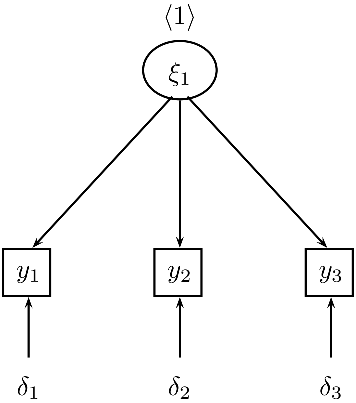
$$H_0 \cap H_+ \text{ against } H_-$$

Is that really a variance?



$$
\begin{aligned}
\sigma_{11} &= \lambda_1^2 + \theta_1 \\
\sigma_{12} &= \lambda_1 \lambda_2 \\
\sigma_{13} &= \lambda_1 \lambda_3 \\
\sigma_{22} &= \lambda_2^2 + \theta_2 \\
\sigma_{23} &= \lambda_2 \lambda_3 \\
\sigma_{33} &= \lambda_3^2 + \theta_3
\end{aligned}
$$

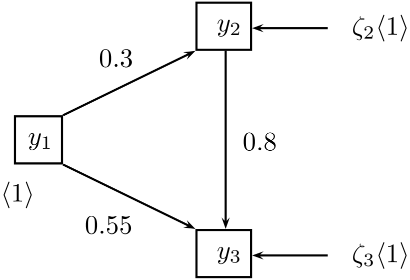
Is that really a variance?



$$\begin{aligned}
\hat\lambda_1 &= \sqrt{s_{12}s_{13}/s_{23}} \\
\hat\lambda_2 &= \sqrt{s_{12}s_{23}/s_{13}} \\
\hat\lambda_3 &= \sqrt{s_{13}s_{23}/s_{12}} \\
\hat\theta_1 &= s_{11} - s_{12}s_{13}/s_{23} \\
\hat\theta_2 &= s_{22} - s_{12}s_{23}/s_{13} \\
\hat\theta_3 &= s_{33} - s_{13}s_{23}/s_{12}
\end{aligned}$$

Population Heywood case

Data generating process



Jump to simulations

Population Heywood case

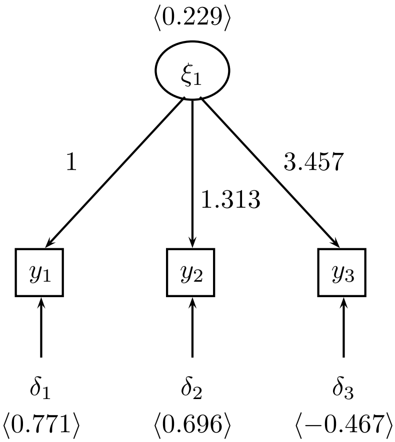
Covariance matrix

$$\begin{pmatrix} 1 & 0.3 & 0.79 \\ 0.3 & 1.09 & 1.037 \\ 0.79 & 1.037 & 2.264 \end{pmatrix}$$

Jump to simulations

Population Heywood case

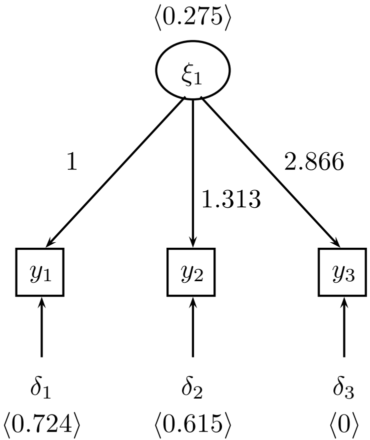
CFA model fit



Jump to simulations

Population Heywood case

CFA with restricted variance



Jump to simulations

Misspecified models

If the model is misspecified, doesn't everything just fall apart??

- Distributional misspecification: the distribution of the data is not normal, likelihood methods are not fully applicable (Satorra 1990, Satorra & Bentler 1994).
- Structural misspecification: the structure of the model, the number of latent variables, the relations between the variables in the model are not specified correctly (Yuan, Marshall & Bentler 2003, Yuan & Hayashi 2006)
- Heywood case: impossible value in the population; evidence of structural misspecification (or what??)

Misspecified models

Huber (1967) and White (1982):

- Point estimates are consistent for the minimizer of the population fit function:

$$\arg\min F(S, \Sigma(\theta)) \to \arg\min F(\Sigma, \Sigma(\theta)) = \theta_0,$$
$$F_{ML}(\Omega, \Sigma(\theta)) = \ln|\Sigma(\theta)| + \Sigma(\theta)^{-1}\Omega - \ln|\Omega| - \dim \Sigma(\theta)$$
$$F(S, \Sigma(\theta)) \stackrel{as.eq.}{=} (s - \sigma(\theta))' \, V_{NT} \, (s - \sigma(\theta)),$$
$$s = \mathrm{vech}\, S, \sigma(\theta) = \mathrm{vech}\, \Sigma(\theta)$$

- Uncertainty about these estimates is given by an asymptotic covariance matrix (sandwich estimator):

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \stackrel{d}{\longrightarrow} N\bigl(0, A^{-1} B A^{-T}\bigr),$$
$$A = \mathbb{E}\, \partial \psi(X, \theta_0), \quad B = \mathbb{E}\, \psi(X, \theta_0) \psi(X, \theta_0)^T,$$
$$\psi(X, \theta_0) = -\dot{\sigma}(\theta) V(s - \sigma(\theta))$$

Sandwich variance estimator

Empirical estimates of “bread” A and “meat” B:

$$\hat{A} = \frac{1}{n}\sum_{i=1}^{n} \partial\psi(x_i, \hat{\theta}),$$

$$\hat{B} = \frac{1}{n}\sum_{i=1}^{n} \psi(x_i, \hat{\theta})\psi(x_i, \hat{\theta})^T$$

- Easily doable with analytic or numeric derivatives
- Byproduct of optimization problem
- Makes fewer assumptions, robust to broader range of model assumption violations
- Studied in and out in regression and econometrics literature
- Easily extendable to clustered/complex survey data

Sandwich variance estimator

Special cases:

- Eicker (1967) and White (1980), linear regression with heteroskedastic errors:

$$\hat{\beta} = (X'X)^{-1}X'y, \quad \hat{v}(\hat{\beta}) = (X'X)^{-1}(X'ee'X)(X'X)^{-1}$$

- Browne (1974): least squares estimates for SEM
- White (1982): somewhat milder regularity conditions that are easier to check in practice; application to some common econometric models.
- Arminger & Schoenberg (1989): an econometrics paper in *Psychometrika*
- Satorra & Bentler (1994): model based sandwich estimator with explicit expression for matrix B (the meat of the sandwich) as a function of the model parameters and the fourth order moments of data
- Yuan & Hayashi (2006): comparison of empirical sandwich and the bootstrap standard errors for SEM

Tests of $H_0 \cap H_+$ vs. H_-

Likelihood ratio type tests

- Overall fit
- χ^2-difference tests (from $\theta_k = 0$)
- Signed root tests for simple χ^2 difference ΔT and scaled Satorra & Bentler (2001) difference ΔT_{sc}

$$r(\theta_0) = \mathrm{sign}[\hat{\theta} - \theta_0]\sqrt{\Delta T_{sc}}$$

Wald type tests, using

- Information matrix standard errors
- Satorra-Bentler standard errors
- Huber empirical sandwich standard errors
- Bootstrap standard errors (empirical, Bollen-Stine)

Can standard errors be trusted?

Distributional specification	Structural specification: Correct	Structural specification: Incorrect
Correct	IM, S-B, ES, EB, BSB	IM, ES, EB
Incorrect	S-B, ES, EB, BSB	ES, EB

Analytic standard errors:

- **IM** observed or expected information matrix
- **S-B** Satorra-Bentler standard errors
- **ES** empirical (Huber) sandwich

Resampling standard errors:

- **EB** empirical bootstrap
- **BSB** Bollen-Stine rotating bootstrap

Simulation study

Saturated model with 3 variables and 6 parameters: $\chi^2_0 \equiv 0$, no way to test the model fit... unless you hit a Heywood case!

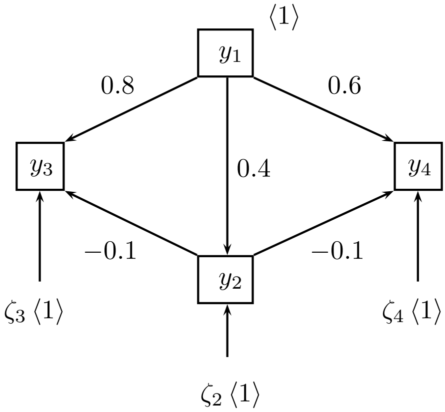
Jump to the Heywood case in population example

Main results:

- HUGE biases of the Heywood case in small samples.
- Non-normal data: information matrix standard errors are biased leading to undercoverage of CIs

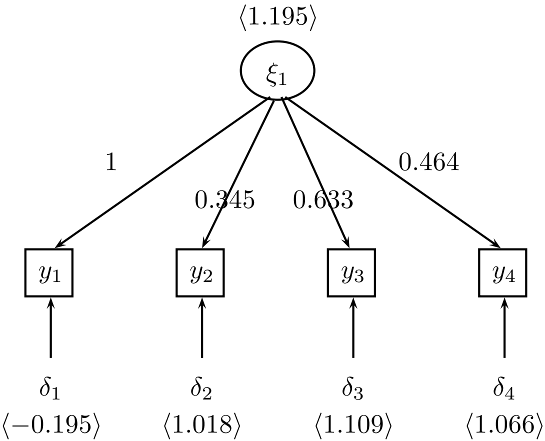
Simulation study

Larger model with 4 variables, 8 parameters and 2 d.f.
Data generating process



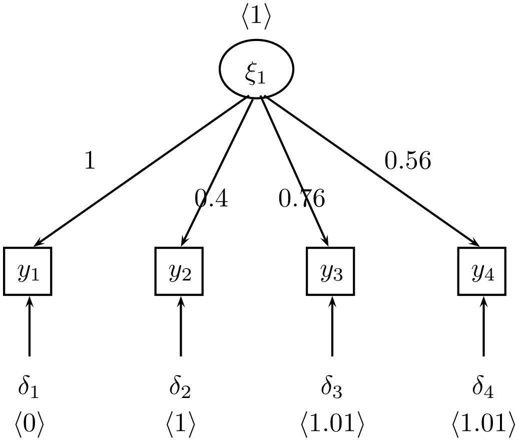
Simulation study

Larger model with 4 variables, 8 parameters and 2 d.f. CFA model fit



Simulation study

Larger model with 4 variables, 8 parameters and 2 d.f. CFA with restricted variance



Simulation study

Larger model with 4 variables, 8 parameters and 2 d.f.
Main results:

- HUGE biases of the Heywood case in small samples.
- Information matrix standard errors are biased leading to undercoverage of CIs based on them when the data are non-normal.
- Satorra-Bentler standard errors are biased leading to undercoverage *even when the data are normal!*
- Huber empirical sandwich standard errors are asymptotically accurate.
- Empirical bootstrap standard errors are even more accurate in smaller samples, and better centered to the true variability.
- Signed root of Satorra-Bentler scaled difference is the most accurate test in terms of the size, and has the greatest power among accurate tests!

Conclusions and recommendations

- Heywood cases are possible in population if the model is grossly misspecified.
- Testing Heywood cases might be the only way to test specification in exactly identified models.
- Empirical sandwich and empirical bootstrap standard errors are the only ones asymptotically accurate under double misspecification.
- Signed root of scaled difference test is a viable alternative to unwieldy likelihood ratio tests when testing at the boundary

References I

Anderson, J. C. & Gerbing, D. (1984), 'The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis', *Psychometrika* **49**, 155–173.

Arminger, G. & Schoenberg, R. J. (1989), 'Pseudo maximum likelihood estimation and a test for misspecification in mean and covariance structure models', *Psychometrika* **54**, 409–426.

Bollen, K. A. (1987), 'Outliers and improper solutions: A confirmatory factor analysis example', *Sociological Methods and Research* **15**, 375–384.

Bollen, K. A. (1989), *Structural Equations with Latent Variables*, Wiley, New York.

Boomsma, A. (1983), *On the Robustness of LISREL (Maximum Likelihood Estimation) Against Small Sample Size and Nonnormality*, Sociometric Research Foundation, Amsterdam, the Netherlands.

References II

Boomsma, A. & Hoogland, J. J. (2001), 'The robustness of LISREL modeling revisited', Structural Equation Modeling: Present and Future.

Browne, M. W. (1974), 'Generalized least squares estimators in the analysis of covariances structures', *South African Statistical Journal* **8**, 1–24.

Dillon, W. R., Kumar, A. & Mulani, N. (1987), 'Offending estimates in covariance structure analysis: Comments on the causes and solutions to Heywood cases', *Psychological Bulletin* **101**, 126–135.

Eicker, F. (1967), Limit theorems for regressions with unequal and dependent errors, *in* 'Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability', Vol. 1, University of California Press, Berkeley, pp. 59–82.

Heywood, H. B. (1931), 'On finite sequences of real numbers', *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **134**(824), 486–501.

References III

Huber, P. (1967), The behavior of the maximum likelihood estimates under nonstandard conditions, *in* 'Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability', Vol. 1, University of California Press, Berkeley, pp. 221–233.

Kolenikov, S. & Bollen, K. A. (2008), Testing negative error variances: Is a Heywood case a symptom of misspecification? under review in *Sociological Methods and Research*.

Rindskopf, D. (1984), 'Structural equation models: Empirical identification, Heywood cases, and related problems', *Sociological Methods and Research* **13**, 109–119.

Sato, M. (1987), 'Pragmatic treatment of improper solutions in factor analysis', *Annals of the Institute of Statistics and Mathematics, part B* **39**, 443–455.

Satorra, A. (1990), 'Robustness issues in structural equation modeling: A review of recent developments', *Quality and Quantity* **24**, 367–386.

Satorra, A. & Bentler, P. (2001), 'A scaled difference chi-square test statistic for moment structure analysis', *Psychometrika* **66**(4), 507–514.

References IV

Satorra, A. & Bentler, P. M. (1994), Corrections to test statistics and standard errors in covariance structure analysis, *in* A. von Eye & C. C. Clogg, eds, 'Latent variables analysis', Sage, Thousands Oaks, CA, pp. 399–419.

Van Driel, O. P. (1978), 'On various causes of improper solutions in maximum likelihood factor analysis', *Psychometrika* **43**, 225–43.

White, H. (1980), 'A heteroskedasticity-consistent covariance-matrix estimator and a direct test for heteroskedasticity', *Econometrica* **48**(4), 817–838.

White, H. (1982), 'Maximum likelihood estimation of misspecified models', *Econometrica* **50**(1), 1–26.

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.

Yuan, K. H., Marshall, L. L. & Bentler, P. M. (2003), 'Assessing the effect of model misspecifications on parameter estimates in structural equation models', *Sociological Methodology* **33**(1), 241–265.