Limited-Information Testing for Structural Models with Categorical Data

Scott Monroe and Li Cai

IMPS, 2013

UCLA

Outline

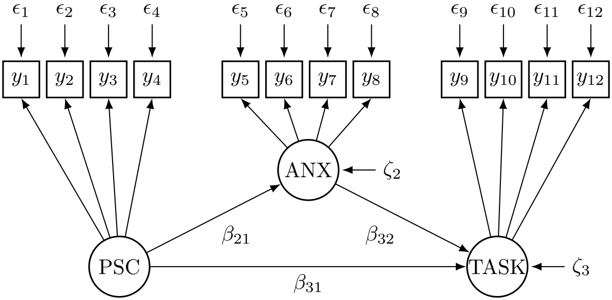
A Motivating Example: PISA Student Questionnaire

Example PISA (2003) Items Measuring Self-Related Cognition in Mathematics

- *How much do you disagree or agree with the following statements?*
  - I learn mathematics quickly.
  - I get very nervous doing mathematics problems.
- *How confident do you feel about having to do the following calculations?*
  - Using a <train timetable>, how long it would take to get from Zedville to Zedtown?

A Proposed Ordinal Structural Model

Latent Mediation Model for PISA Questionnaire Data



- ***PSC***: Positive self-concept as a mathematics student
- ***ANX***: Mathematics anxiety
- ***TASK***: Task-specific confidence

Estimation for Ordinal Structural Models

This research considers the *multistage* estimator, which estimates:

1. thresholds by ML
2. polychoric correlations by ML
    - stages 1 and 2 yield a sample polychoric correlation matrix
3. structural parameters by some form of least squares

Goodness-of-Fit Testing for Ordinal Structural Models

First type:
statistic based on minimized fit-function value

- Let F be the minimum fit function value from estimation
- Then, $T = (N - 1)F$ is used to construct a test statistic
- Typically, T is adjusted to approximate a chi-square variate using moment-matching (e.g., Satorra and Bentler, 1994)
  - define T_U and T_D as mean- and variance-adjusted stats based on ULS and DWLS, respectively

Second type:
statistic based on contingency table residuals (Maydeu-Olivares, 2001)

- theoretical appeal of accounting for all levels of uncertainty
- Maydeu-Olivares (2001) derived 3 test statistics:
  1. distributional
  2. structural
  3. overall
- like T_U and T_D, all 3 statistics formed by matching moments

Limited-Information Test Statistics and M_2

Maydeu-Olivares and Joe (2005, 2006) proposed M_2

- quadratic form based on first- and second-order marginal residuals
- *limited-information* statistic
- M_2^*, a version of M_2 for polytomous responses (Joe and Maydeu-Olivares, 2010, Cai and Hansen, 2012)
- chi-square distributed

Research Focus

M_2 has been successfully applied to many IRT models, estimated by ML.

But, M_2 is not limited to IRT or ML (Maydeu-Olivares and Joe, 2006).

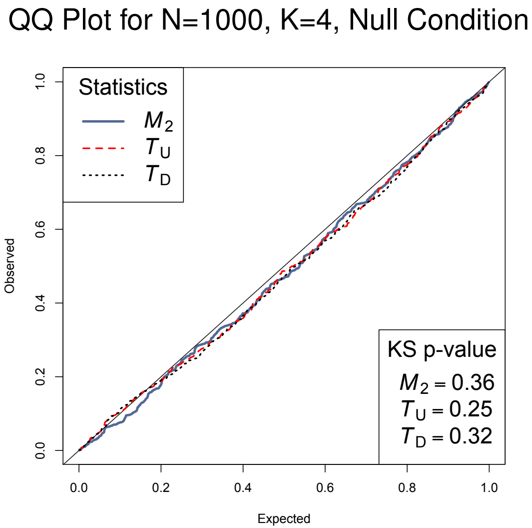
The current research uses M_2 and M_2^* as an overall test for ordinal structural models, estimated by the multistage estimator.

Simulation Study

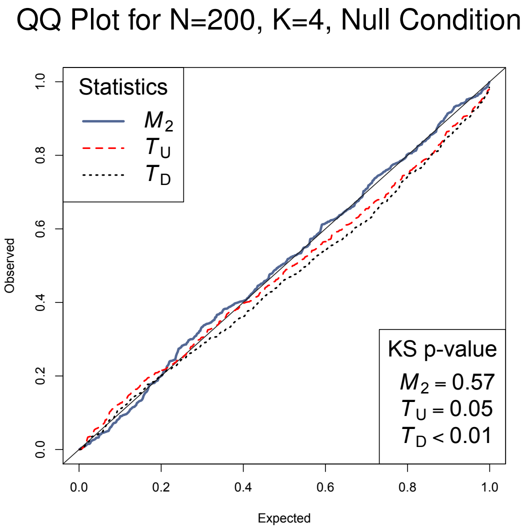
- Purpose:
  1. show M_2 is chi-squared
  2. compare M_2 to T_U and T_D in terms of calibration and power

- Conditions:
  - 500 replications attempted
  - model identical to PISA example (latent mediation)
  - $N = 100, 200, 500, 1000$
  - $K = 2$ or 4 categories per item
  - model misspecification via Tucker, Koopman, and Linn (TKL, 1969)

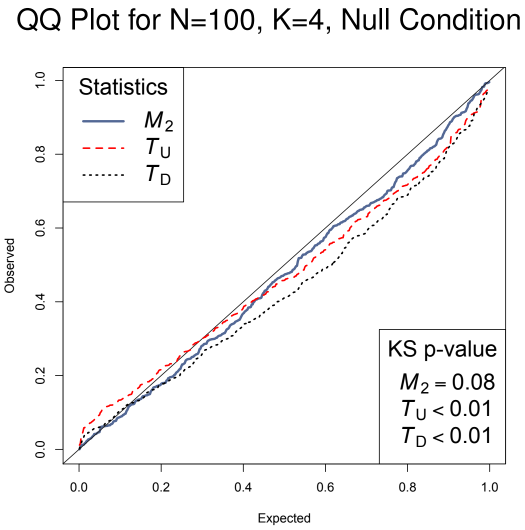
Calibration of Test Statistics



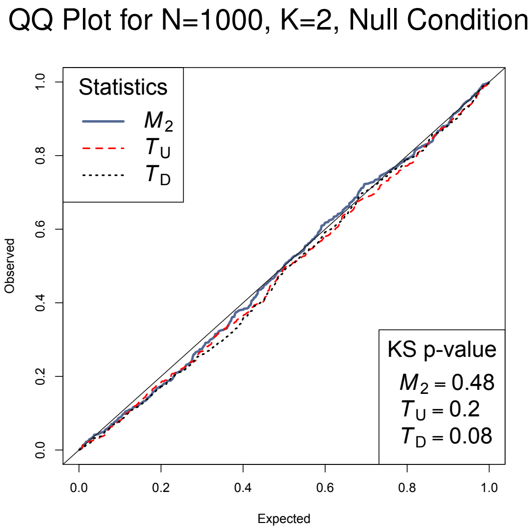
Calibration of Test Statistics



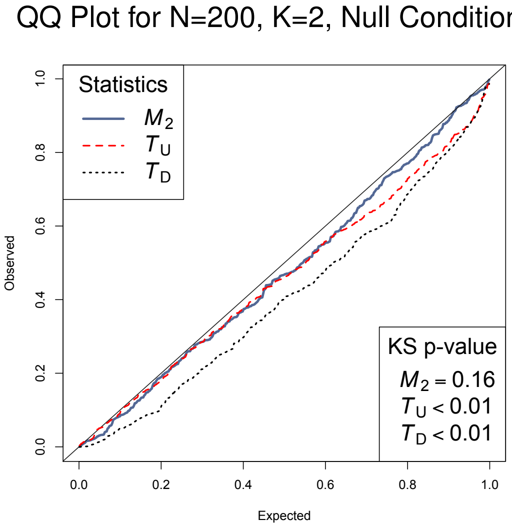
Calibration of Test Statistics



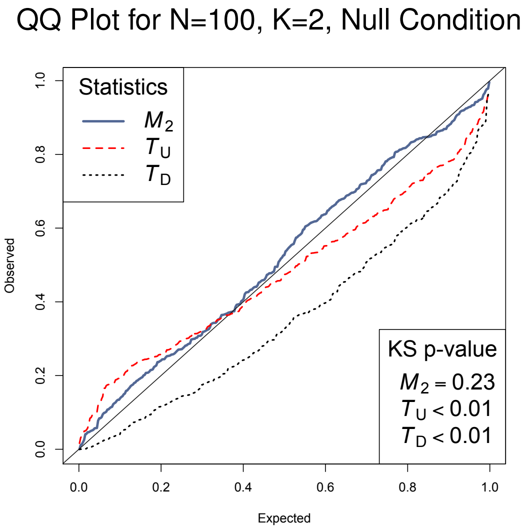
Calibration of Test Statistics



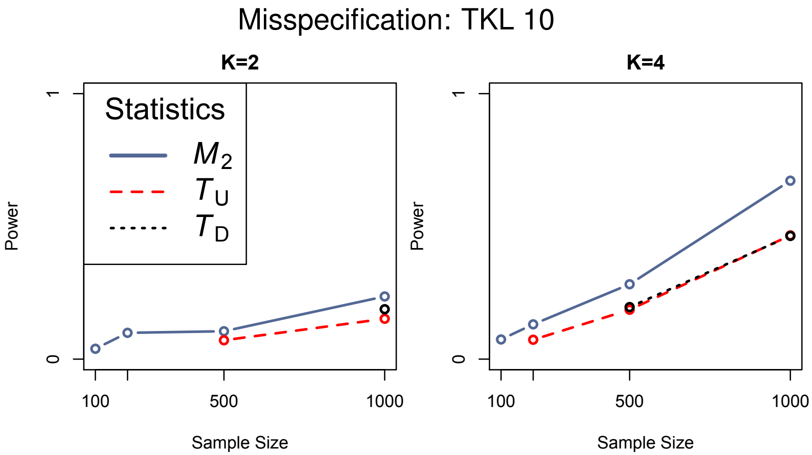
Calibration of Test Statistics



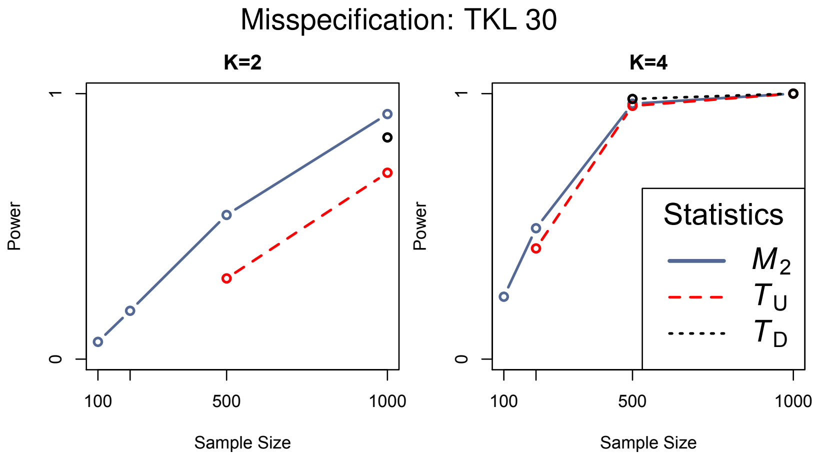
Calibration of Test Statistics



Power of Test Statistics at $\alpha = .05$



Power of Test Statistics at $\alpha = .05$



An Aside: RMSEA for Discretized Latent Variable

For TKL10, the *population* RMSEA is .033

Mean (SD) M_2-based RMSEA for TKL10

	Sample Size			
K	100	200	500	1000
2	.017 (.023)	.016 (.018)	.011 (.011)	.011 (.008)
4	.027 (.028)	.022 (.022)	.022 (.014)	.025 (.010)

An Aside: RMSEA for Discretized Latent Variable

For TKL30, the *population* RMSEA is .070

Mean (SD) M_2-based RMSEA for TKL30

	Sample Size			
K	100	200	500	1000
2	.021 (.023)	.023 (.017)	.026 (.011)	.027 (.006)
4	.045 (.032)	.046 (.023)	.050 (.011)	.051 (.008)

Empirical Application

Results for PISA data example (US sample, $N = 5,086$)

Stat	Value	*df*	*p*	TLI	RMSEA	90% CI
T_U	330.16	30^{**}	$< .001$	0.995	0.044	(0.040, 0.048)
T_D	571.50	33^{**}	$< .001$	0.995	0.057	(0.053, 0.061)
M_2	108.62	27	$< .001$	0.997	0.024	(0.020, 0.029)

note: ** indicates an approximation to *df*

Conclusion

M_2 can be applied to structural equation models when the data are categorical.

Advantages of M_2:

- better calibration than T_U & T_D, particularly with small samples
- more powerful

Disadvantages of M_2:

- computationally demanding
- not as versatile as traditional stats

Questions:

- how do M_2-based fit indices perform?
- does M_2 have power against distributional misspecifications?

Key References I

Cai, L., & Hansen, M. (2012). Limited-information goodness-of-fit testing of hierarchical item factor models. *British Journal of Mathematical and Statistical Psychology*, *66*(2), 245-276.

Maydeu-Olivares, A. (2006). Limited information estimation and testing of discretized multivariate normal structural models. *Psychometrika*, *71*(1), 57-77.

Maydeu-Olivares, A., & Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, *71*, 713–732.

Satorra, A., & Bentler, P. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications to developmental research* (p. 399-419). Newbury Park, CA: Sage Publishing.

Acknowledgements

This research is supported by grants from the Institute of Education Sciences (R305B080016 and R305D100039) and the National Institute on Drug Abuse (R01DA026943 and R01DA030466).

UCLA