Alignment with Bayesian Region of Measurement Equivalence (ABROME) Approach for Multiple Groups Comparisons

Yichi Zhang, Hok Chio (Mark) Lai

University of Southern California

Measurement and Multilevel Modeling Lab

USC University of Southern California

Agenda

- Background of Measurement Invariance (MI) Research
- Alignment with Bayesian Region of Measurement Equivalence (ABROME) Framework
- ABROME for Multiple Groups
- Illustrative Example

Background of Measurement Invariance Research

- Existing methods for testing MI
  - Likelihood ratio test
  - Goodness-of-fit indices (e.g., RMSEA)
  - Effect size indices (e.g., d_{MACS} by Nye & Drasgow, 2011)
- Limitation of existing MI literature
  - Difficult to interpret
  - Not directly supportive of the null hypothesis
  - Iteratively searching for biased items

Background of Measurement Invariance Research

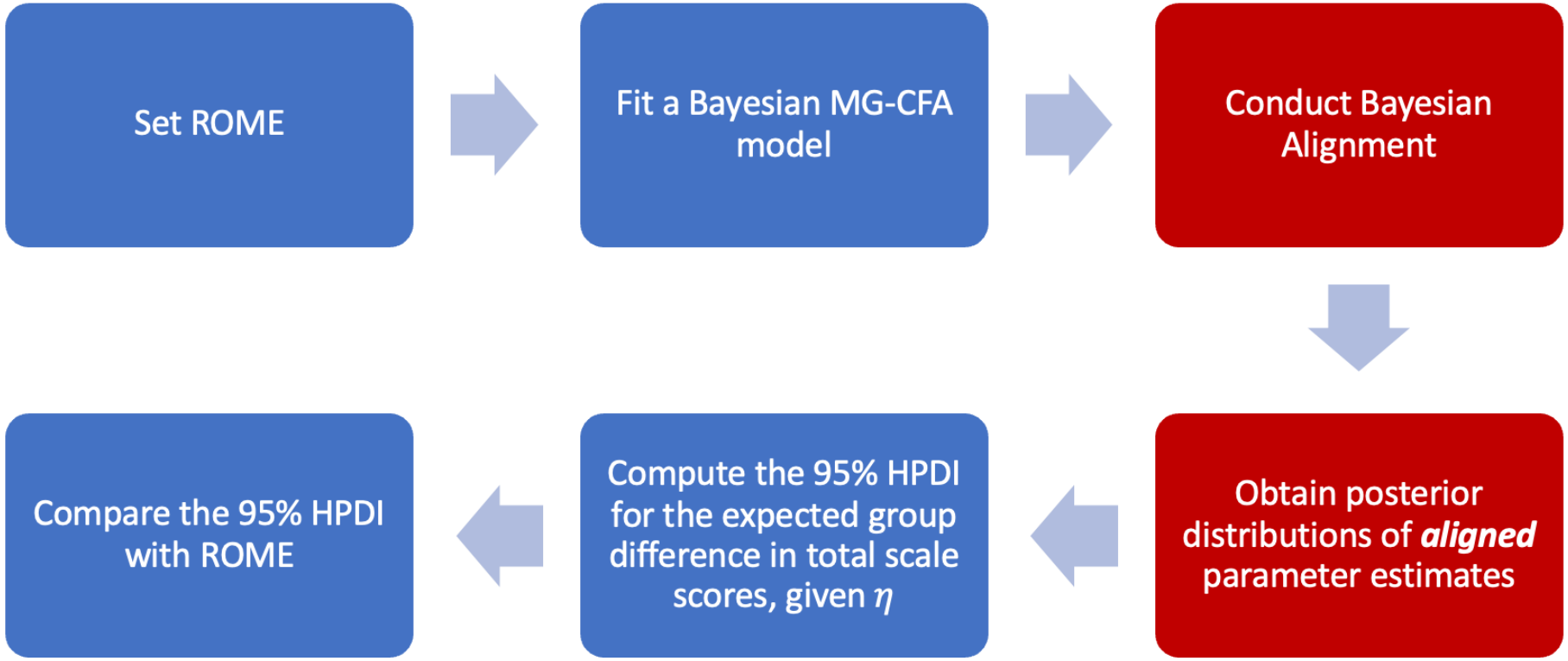
- Equivalence testing (Yuan & Chan, 2016)
  - Test whether degree of noninvariance < prespecified threshold of "negligible" (ε_0)
- Bayesian Region of Measurement Equivalence (ROME) framework (Zhang et al., 2022)
  - A prespecified tolerable range of expected group difference in total scale scores (*ROME*).
  - The most probable group difference due to noninvariance/bias (*highest posterior density interval; HPDI*)

Alignment with Bayesian ROME Framework

- The Bayesian ROME framework
  - Requires prior knowledge of noninvariant items
    - Conventional methods of MI testing
    - Alignment Optimization
- Alignment (Asparouhov & Muthén, 2014)
  - Minimizes the amount of total measurement noninvariance
  - Automates the MI analysis without requiring exact MI

Items	Aligned Loadings		Aligned Intercepts	
	Male	Female	Male	Female
Item1	0.734	0.721	2.445	2.425
Item 2	0.720	0.696	2.307	2.206
Item 3	0.819	0.817	2.349	2.339
Item 4	0.837	0.844	2.457	2.542
Item 5	0.814	0.850	2.494	2.567

ABROME Framework



ABROME for Multiple Groups

- Applied researchers are often interested in whether a scale is practically invariant across three or more groups (e.g., race/ethnicity)
- Bayesian ROME for Multiple Groups
  - Test overall noninvariance across all groups
  - Conduct pairwise group comparisons if practical noninvariance is founded

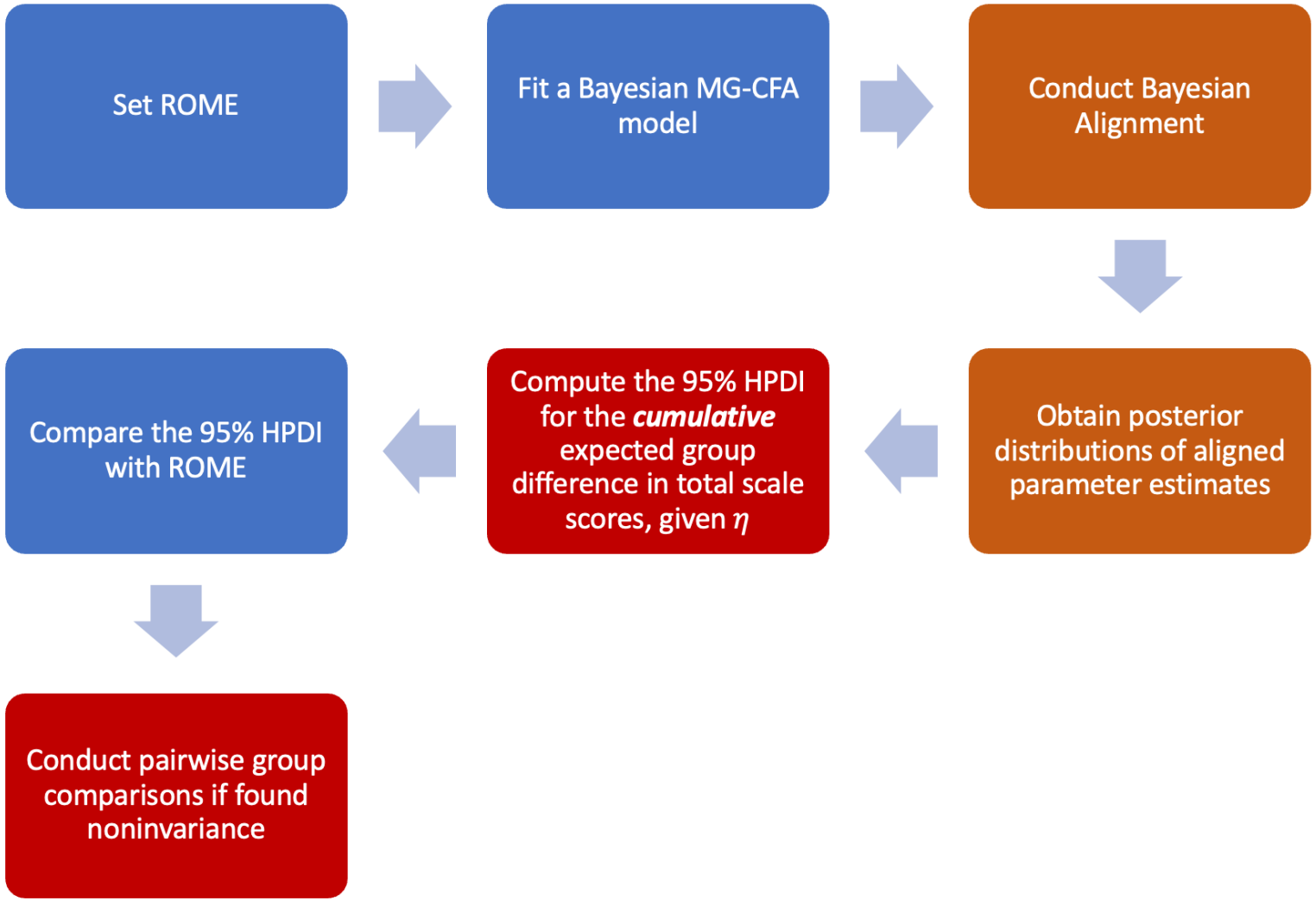
ABROME for Multiple Groups

- Computation of the cumulative expected group differences in scale scores (Lai et al., 2023)

$$E(D_T|\eta) = \sum_{k=1}^{m} E(T_k - \bar{T}_k|\eta)^2$$

- Specifying ROME
  - For two group comparisons, $[-0.1s_p, 0.1s_p]$ (Cohen, 1988)
  - For multiple group comparisons, $[0, 0.05\ s_p]$ (Zhang et al., 2022)
  - s_p is the pooled standard deviation across all groups

ABROME For Multiple Groups



Illustrative Example

- Goal: Investigate practical invariance of a five-item mathematics-specific self-efficacy scale across racial groups from a nationally representative sample (Educational Longitudinal Study of 2002)
- Randomly chose 2,000 grade 10 students
  - 60.35% White
  - 9.45% Asian
  - 14.65% Black
  - 10.2% Hispanic

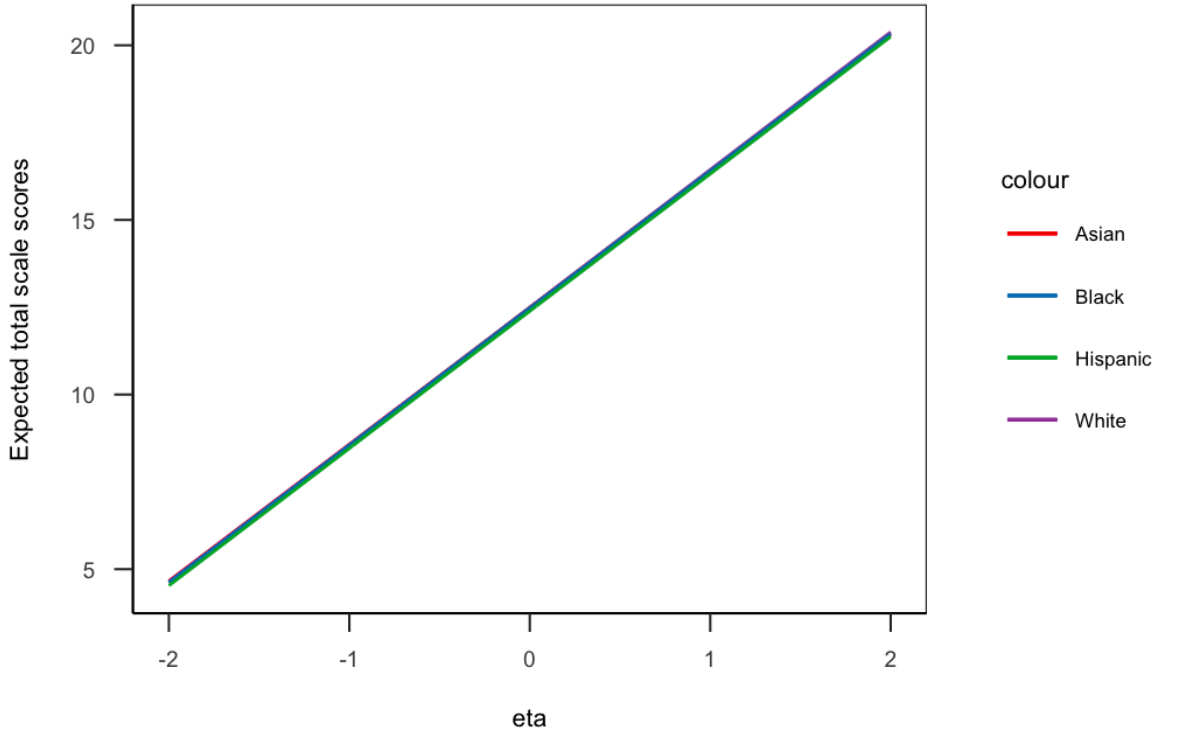
Illustrative Example

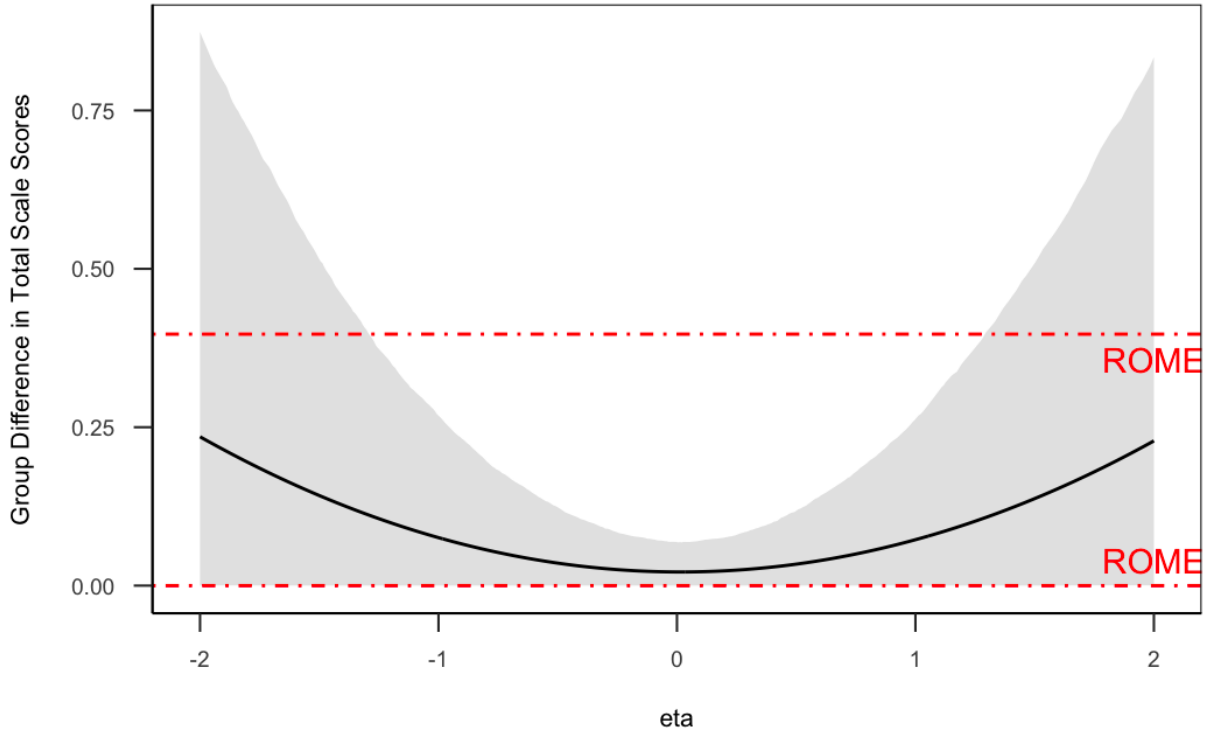
Mathematics Self-Efficacy Scale

Item Labels	Description	Factor Loadings
BYS89A	Can do excellent job on math tests	.829
BYS89B	Can understand difficult math texts	.836
BYS89L	Can understand difficult math class	.876
BYS89R	Can do excellent job on math assignments	.876
BYS89U	Can master math class skills	.872
Percent of Variance Explained		78.9

Illustrative Example - Overall

- ROME was set as [0, 0.397]
- Students from different racial backgrounds perform similarly
- The practical significance of noninvariance remains inconclusive

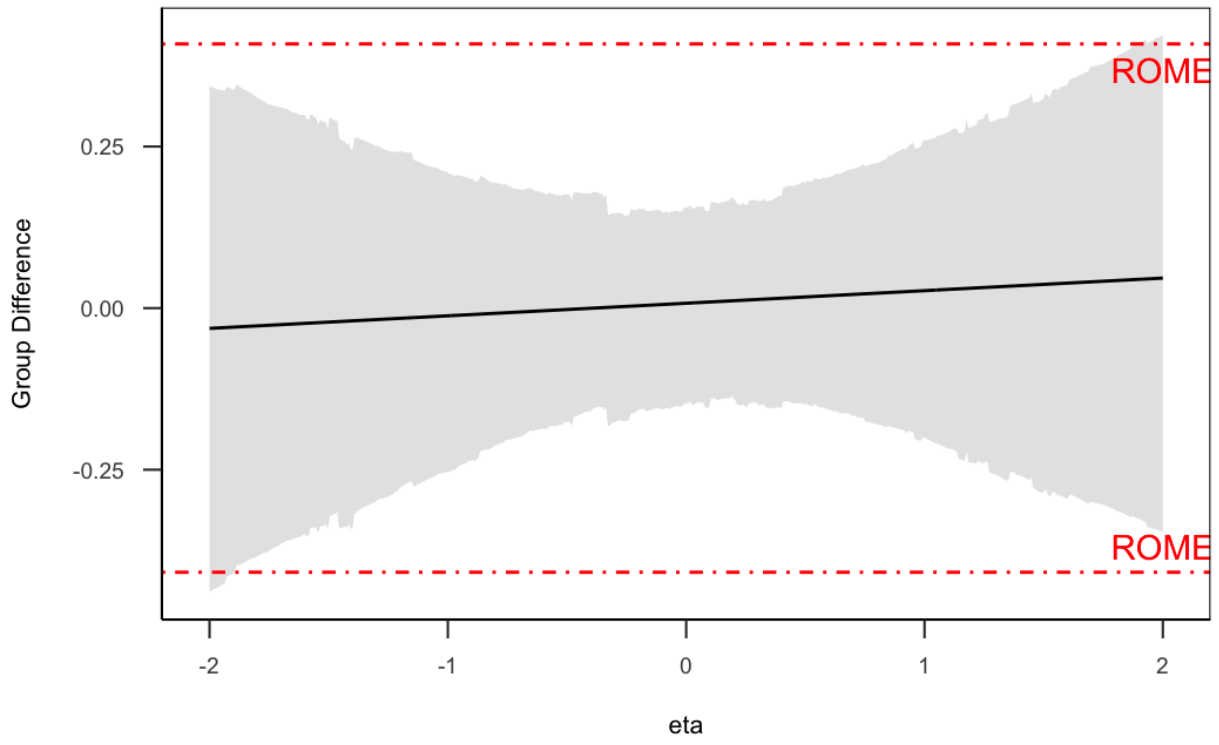




Illustrative Example – Pairwise Group Comparisons

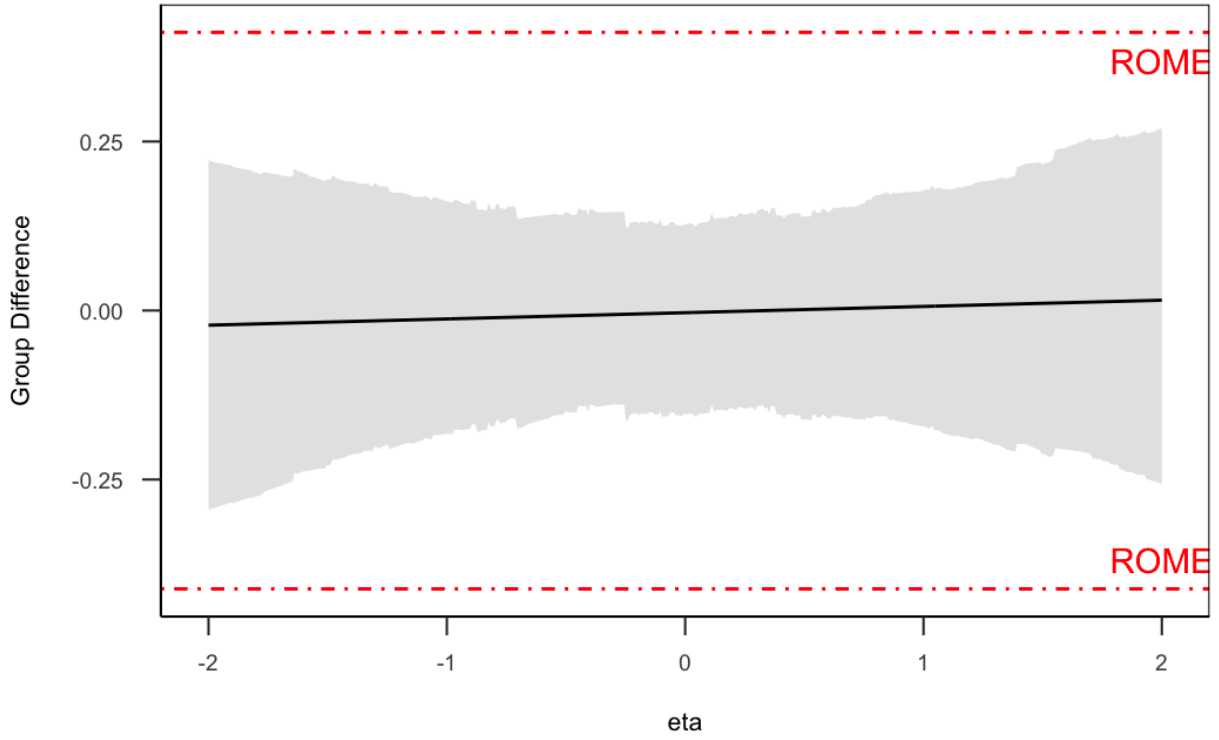
White vs Asian

ROME [-0.409, 0.409]



White vs Black

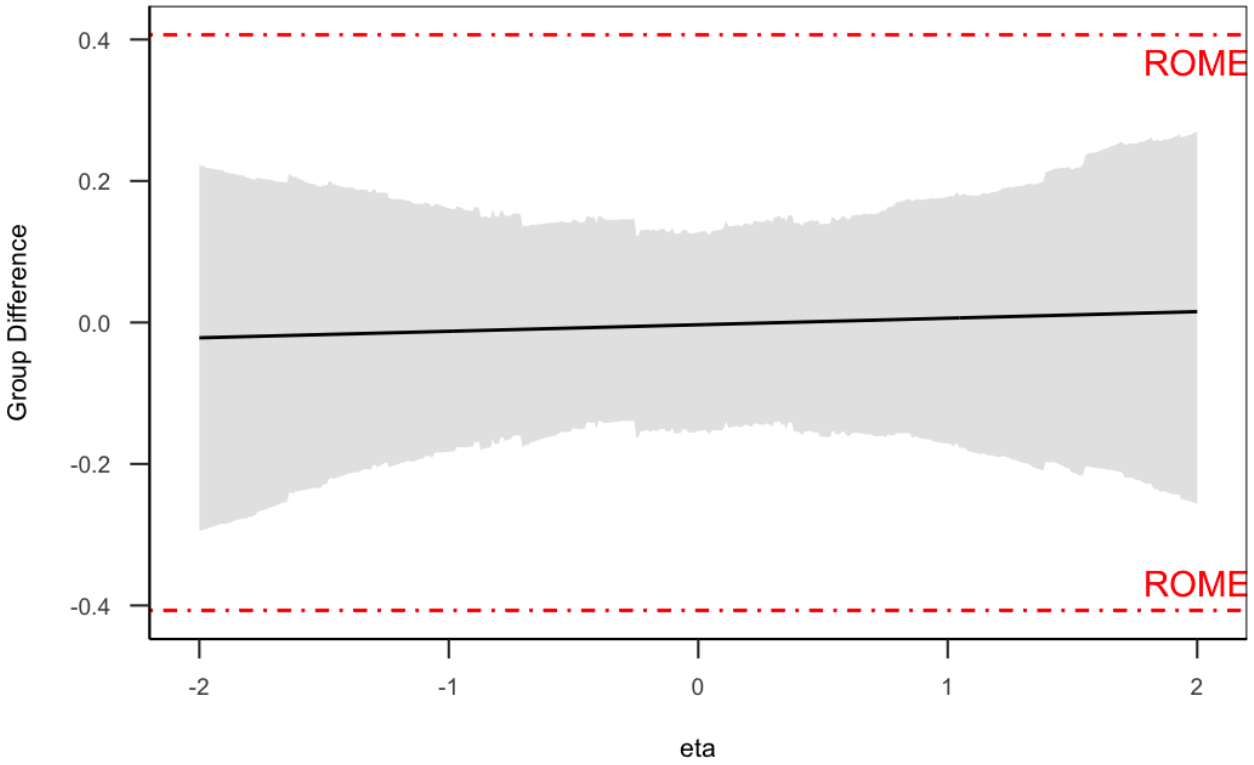
ROME [-0.412, 0.412]



Illustrative Example – Pairwise Group Comparisons

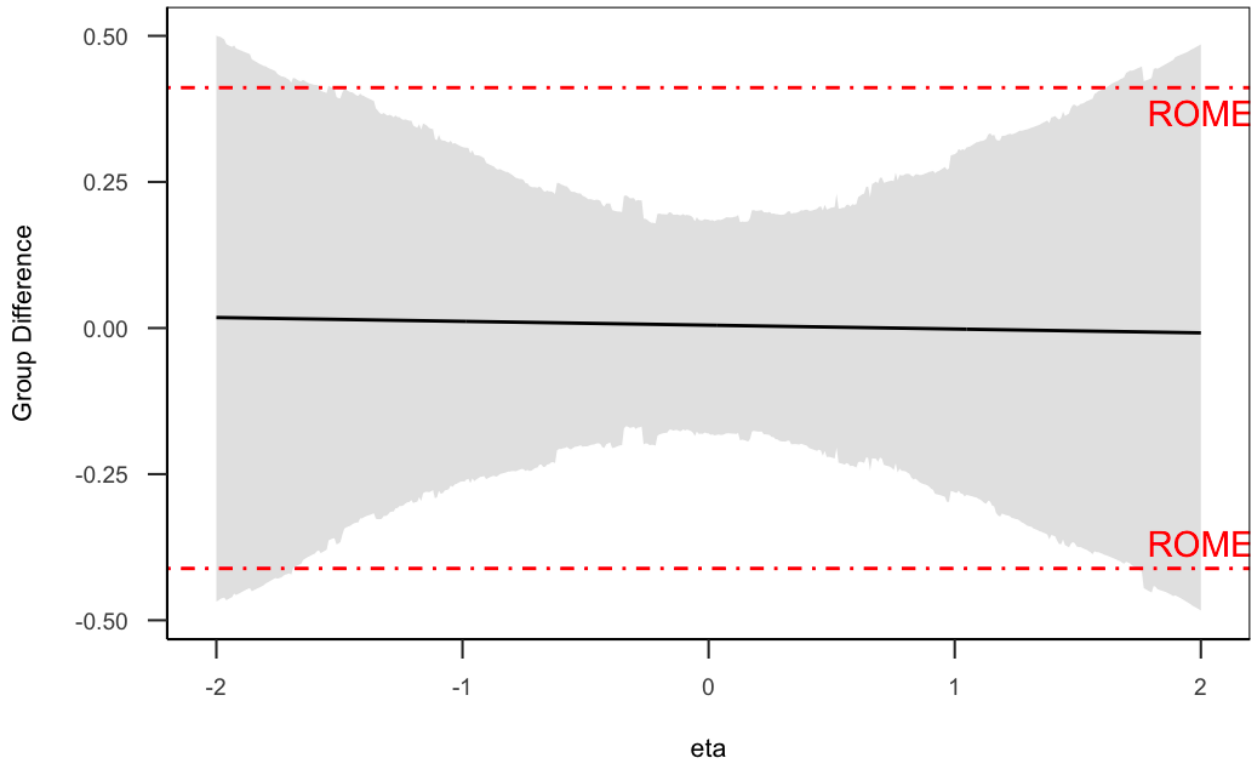
White vs Hispanic

ROME [-0.407, 0.407]



Asian vs Black

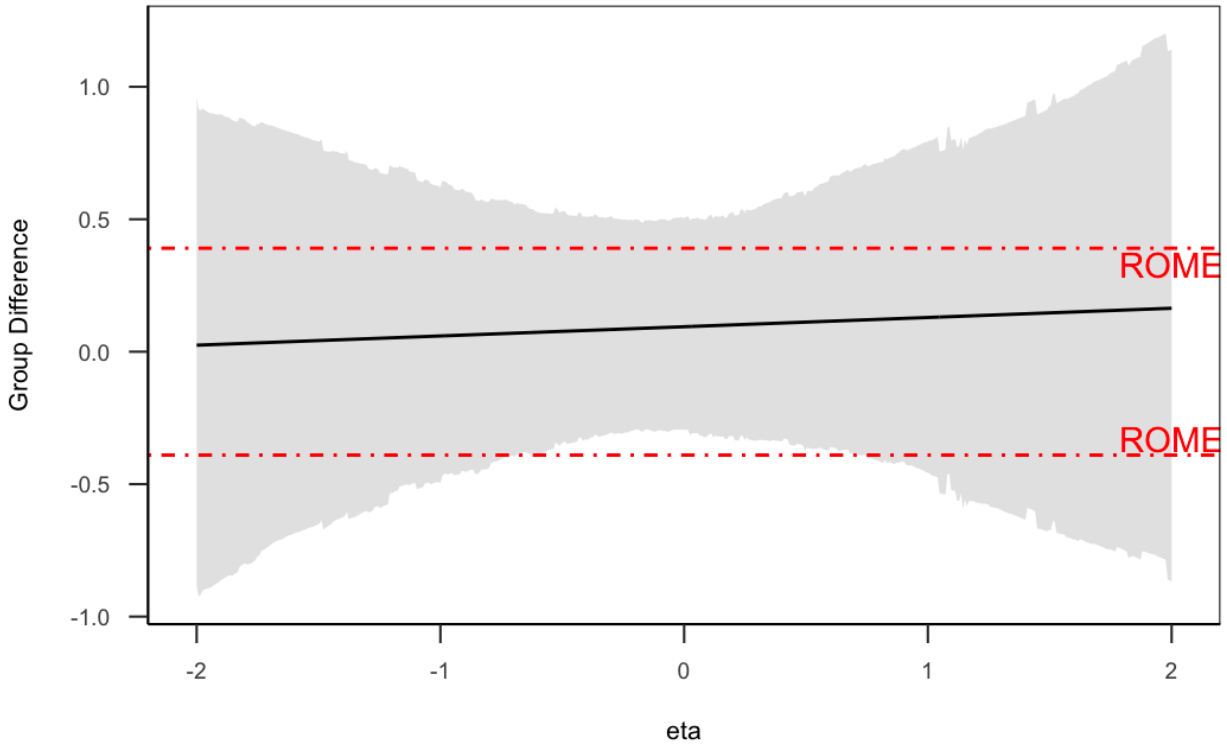
ROME [-0.411, 0.411]



Illustrative Example – Pairwise Group Comparisons

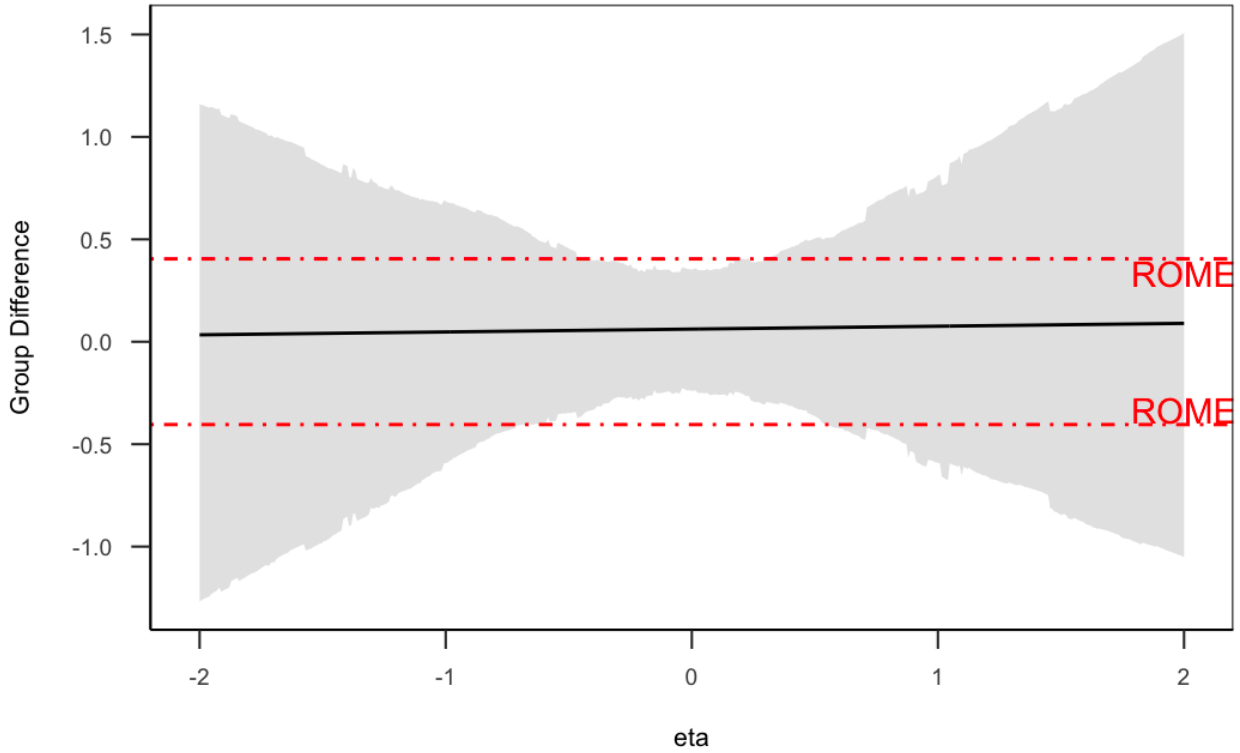
Asian vs Hispanic

ROME [-0.391, 0.391]



Black vs Hispanic

ROME [-0.404, 0.404]



Results Summary

- Whether the mathematics self-efficacy scale (math SE) is practically invariant across all racial groups (White, Asian, Black, Hispanic) remains inconclusive
- By conducting pairwise group comparisons using the Bayesian ROME framework
  - Math SE is practically invariant between
    - White and Asian
    - White and Black
    - White and Hispanic
  - Further investigation is needed for
    - Asian vs Black
    - Asian vs Hispanic
    - Black vs Hispanic

Limitation and Future Directions

- The ABROME with multiple groups framework could be time consuming when the number of groups is large
  - Consider using Bayesian ROME with Multilevel Confirmatory Factor Analysis
- The 95% HPDI could be outside ROME if uncertainty is high
  - Increase sample size
  - Use anchor item

Contributions

- The proposed Bayesian ROME with multiple groups framework
  - Enables researchers to support practical invariance of the scale of interest across multiple groups
  - Quantifies the degree of noninvariance on an easily understandable metric (observed total scale scores)
  - Provides more useful information
    - Who are mostly biased by the scale
    - Which pair of groups have significant bias

References

Asparouhov, T., & Muthén, B. (2014). Multiple-Group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(4), 495–508. https://doi.org/10.1080/10705511.2014.919210

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Routledge.

U.S. Department of Education, National Center for Education Statistics. (2004). Education longitudinal study of 2002: Base year data file user's manual, by Steven J. Ingels, Daniel J. Pratt, James E. Rogers, Peter H. Siegel, and Ellen S. Stutts. Washington, DC.

Yuan, K.-H., & Chan, W. (2016). Measurement invariance via multigroup SEM: Issues and solutions with chi-square-difference tests. *Psychological Methods*, 21(3), 405–426. https://doi.org/10.1037/met0000080

Zhang, Y., Lai, M. H. C., & Palardy, G. J. (2022). A Bayesian region of measurement equivalence (ROME) approach for establishing measurement invariance. *Psychological Methods.* Advance online publication. https://doi.org/10.1037/met0000455

Zhang, Y., Lai, M. H. C. (2022). Bayesian region of measurement equivalence (ROME) approach with Alignment [Manuscript in preparation].

Thank you!

Feel free to contact me via email (yzhang97@usc.edu)