A Dirichlet Model for Interval Responses

Matthias Kloft

FU Berlin, Steffi Pohl Lab - 7th September 2022

1 - Introduction

„I like being around other people“

Why Use Interval Responses?

Motivating Example:

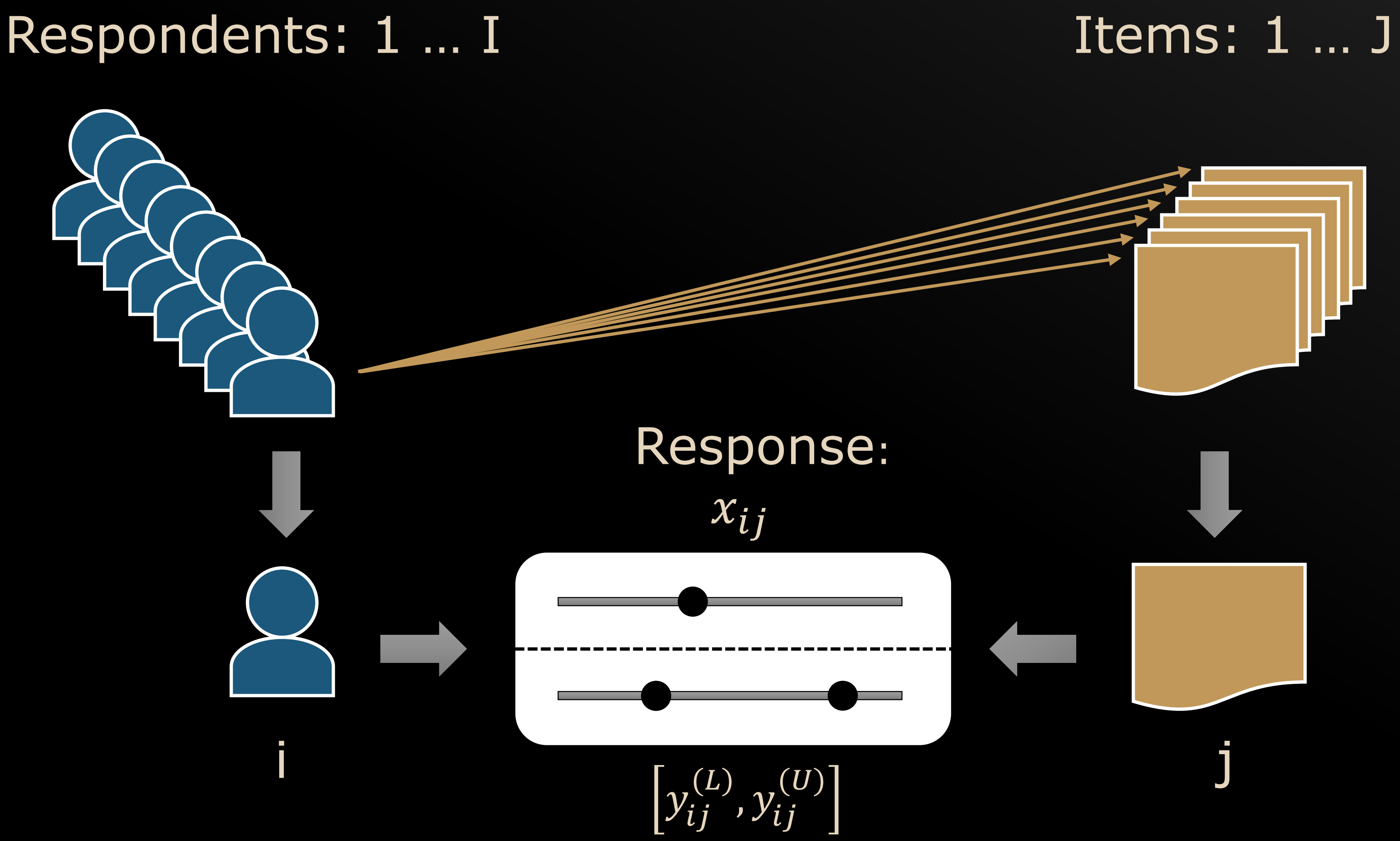
- Whole Trait Theory (Fleeson, 2001)
  - Trait: Distribution of states

➢ Accounting for variability

➢ Range of valid values

2 – IRT Models

Testing Scenario

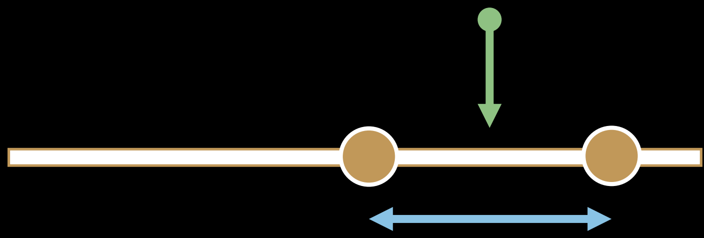


Interval Response

Manifest Response:

Interval Location (Midpoint):

$$\frac{y_{ij}^{(L)} + y_{ij}^{(U)}}{2}$$



Interval Width: $y_{ij}^{(U)} - y_{ij}^{(L)}$

Interval Response

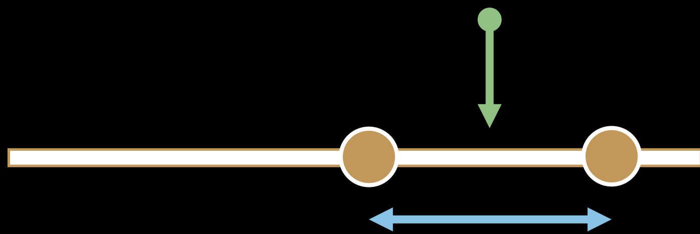
Manifest Response:

Latent Space:

Interval Location (Midpoint):

$$\frac{y_{ij}^{(L)} + y_{ij}^{(U)}}{2}$$

Location Dimension



Interval Width: $y_{ij}^{(U)} - y_{ij}^{(L)}$

Expansion Dimension

Compositional Data

- Components must sum to one: simplex

Restrictions

No support for zero-components

- Single Response:

$$0 < x_{ij} < 1$$

- Interval Response:

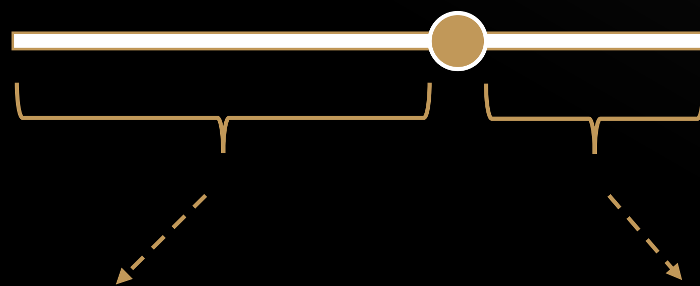
$$0 < y_{ij}^{(L)} < y_{ij}^{(U)} < 1$$

BETA RESPONSE MODEL (BRM)

Noel & Dauvier (2007)

$$x_{ij} \sim \mathrm{Beta}\left(a_{ij}, d_{ij}\right);$$

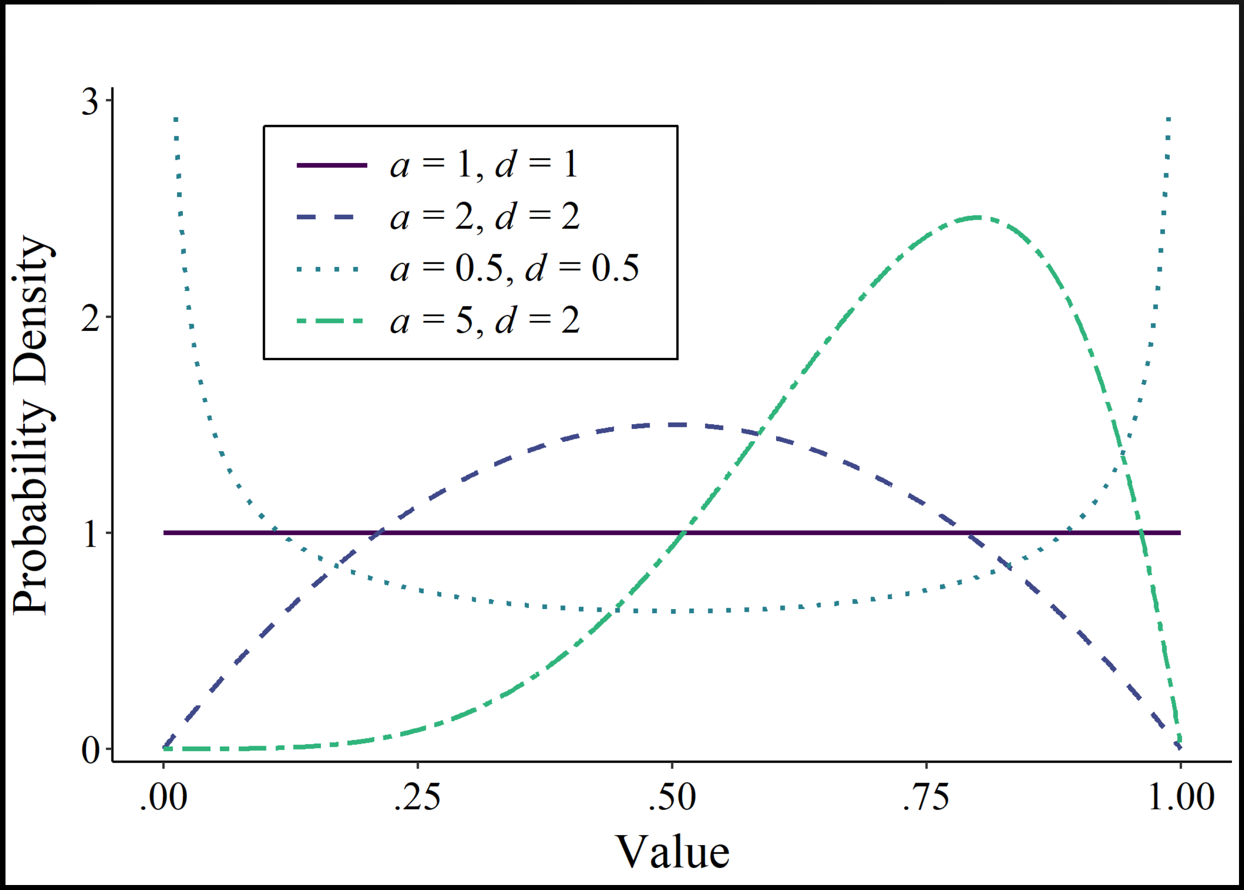
$$E\left(x_{ij}\right) = \frac{a_{ij}}{a_{ij}+d_{ij}}$$



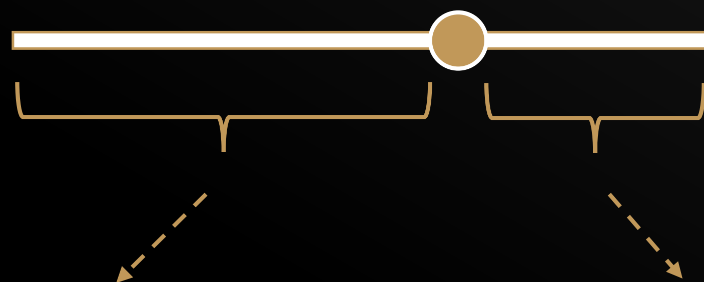
$$a_{ij} = \exp\left[\alpha\left(\theta_i - \delta_j\right) + \tau_j\right]$$

$$d_{ij} = \exp\left[-\alpha\left(\theta_i - \delta_j\right) + \tau_j\right]$$

BETA DISTRIBUTION



Beta Response Model (BRM)
Parameters: Ability / Difficulty



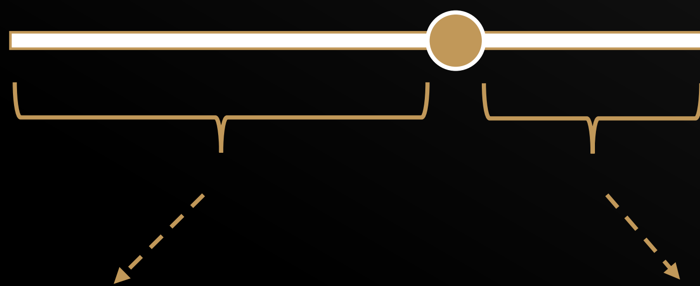
$$a_{ij} = \exp\left[\alpha\underbrace{\left(\theta_i - \delta_j\right)} + \tau_j\right] \qquad d_{ij} = \exp\left[-\alpha\underbrace{\left(\theta_i - \delta_j\right)} + \tau_j\right]$$

θ_i: Person ability

δ_j: Item difficulty

➢ Classic interpretation

Beta Response Model (BRM)
Parameters: Scaling



$$a_{ij} = \exp\left[\alpha\left(\theta_i - \delta_j\right) + \tau_j\right] \qquad d_{ij} = \exp\left[-\alpha\left(\theta_i - \delta_j\right) + \tau_j\right]$$

$\pm\alpha > 0$: Scaling

- Continuous model:
  Not a discrimination parameter!!

Beta Response Model (BRM)

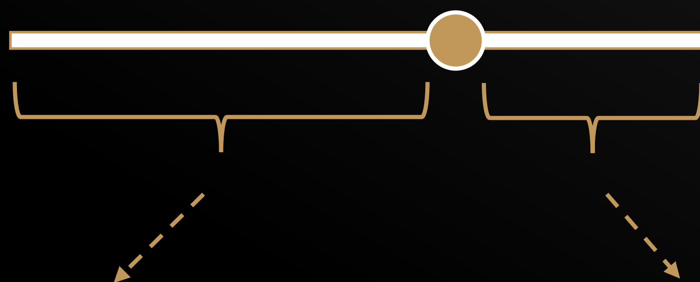
Parameters: Precision

$$a_{ij} = \exp\left[\alpha\left(\theta_i - \delta_j\right) + \underbrace{\tau_j}\right] \qquad d_{ij} = \exp\left[-\alpha\left(\theta_i - \delta_j\right) + \underbrace{\tau_j}\right]$$

$\tau_j > 0$: Item precision (both additive!)

- ➢ Steeper density curves

BETA RESPONSE MODEL (BRM)
Parameters: Exponential Link

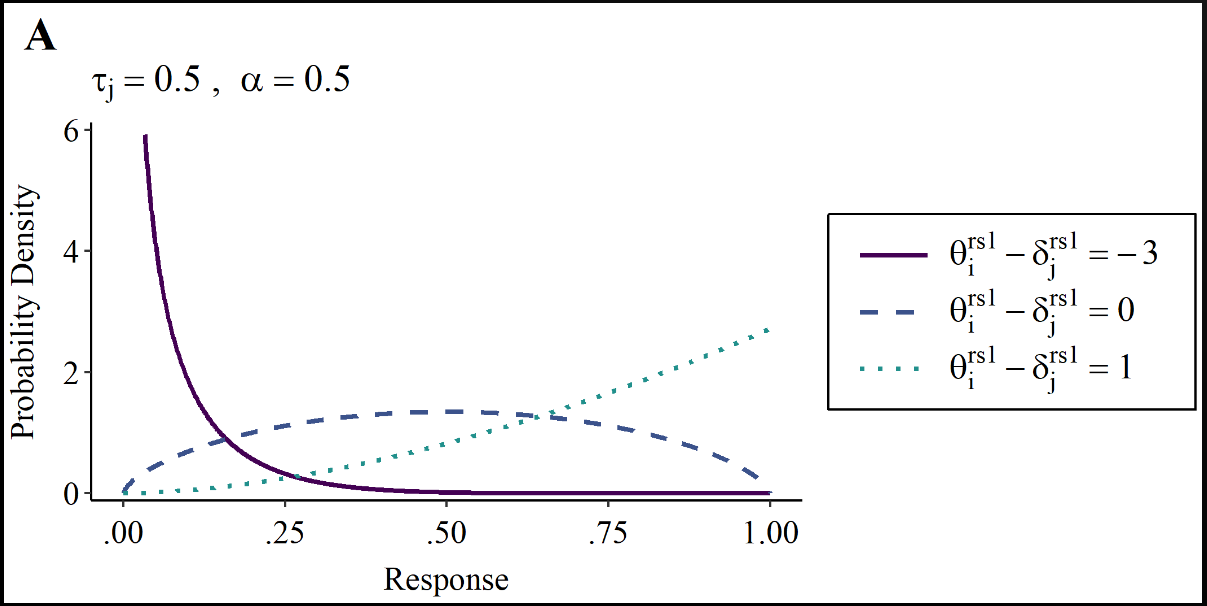


$$a_{ij} = \exp\left[\alpha\left(\theta_i - \delta_j\right) + \tau_j\right] \qquad d_{ij} = \exp\left[-\alpha\left(\theta_i - \delta_j\right) + \tau_j\right]$$

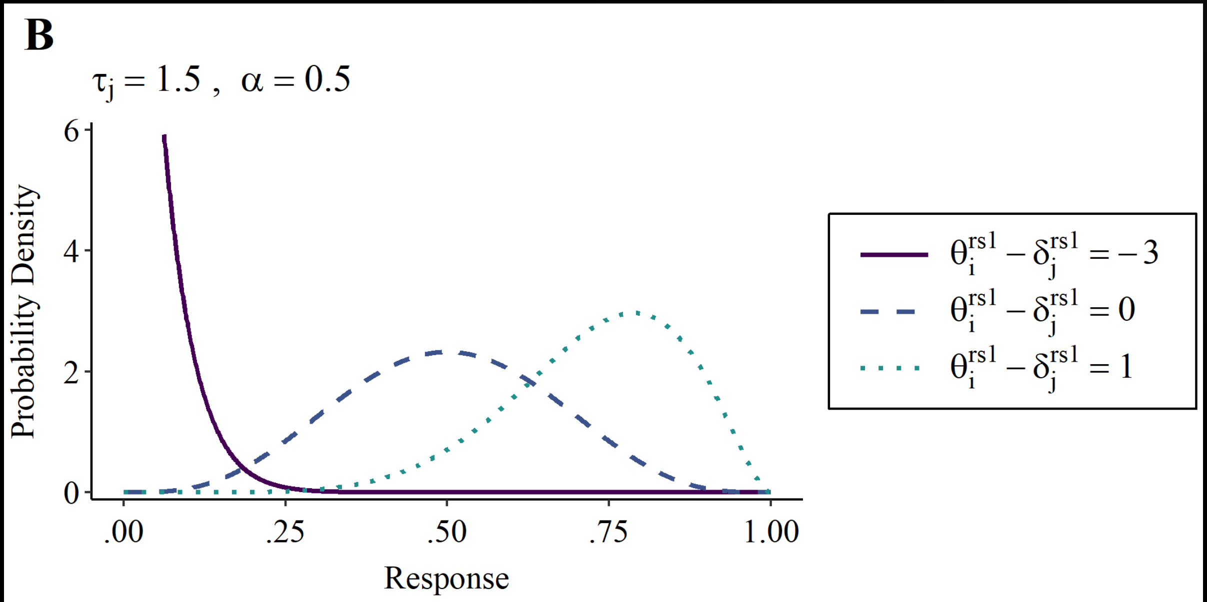
Example:

- $\theta_i - \delta_j = 0; \qquad \tau_j = 0$
- ➢ $\exp(0) = 1$
- ➢ $\mathrm{Beta}(1, 1)$: uniform

BRM: Examples

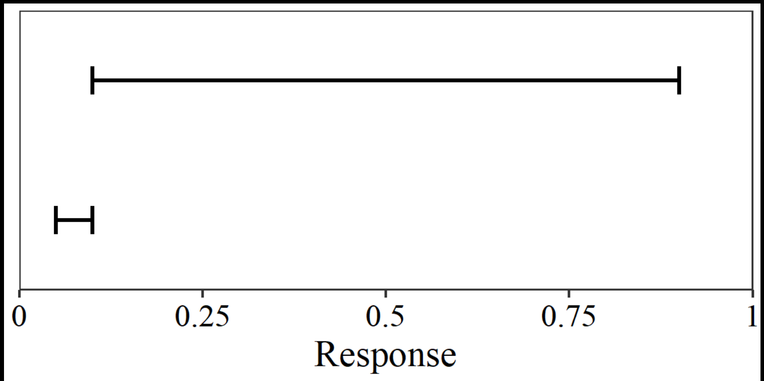


$\tau_j = 0.5$

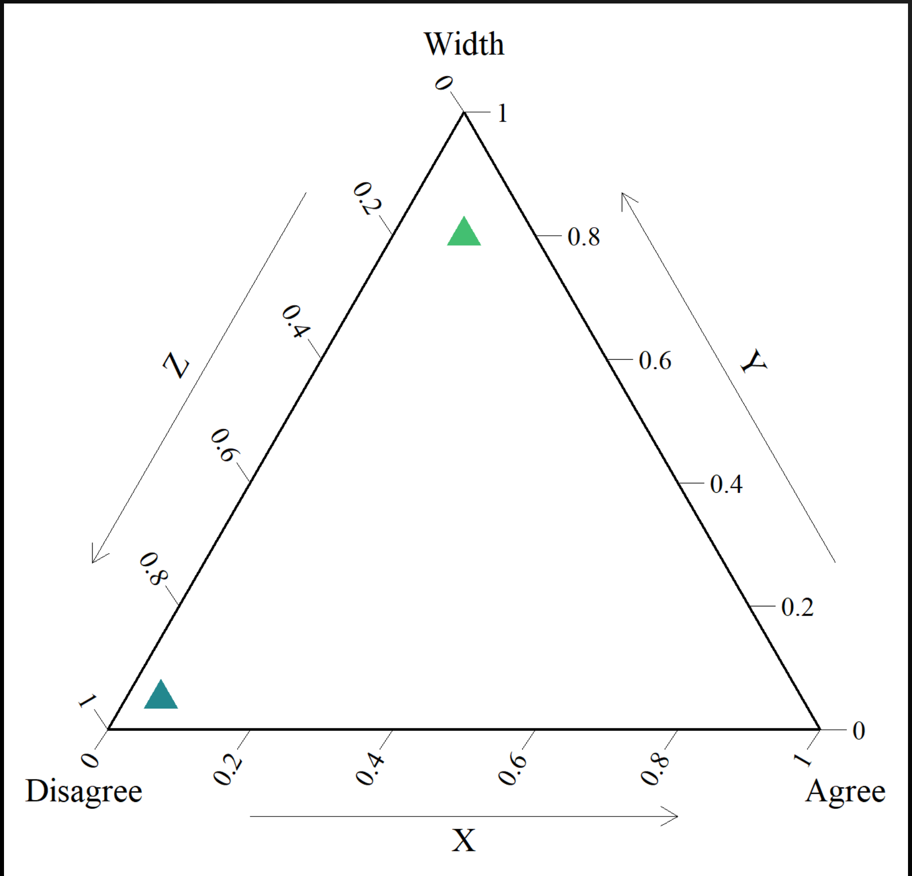


$\tau_j = 1.5$

Response Intervals - Ternary Space

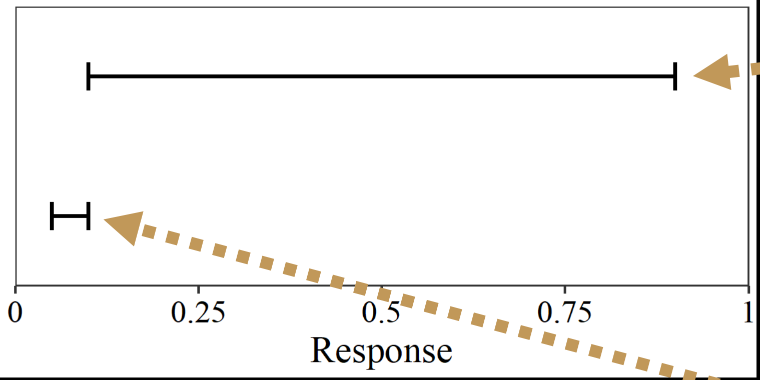


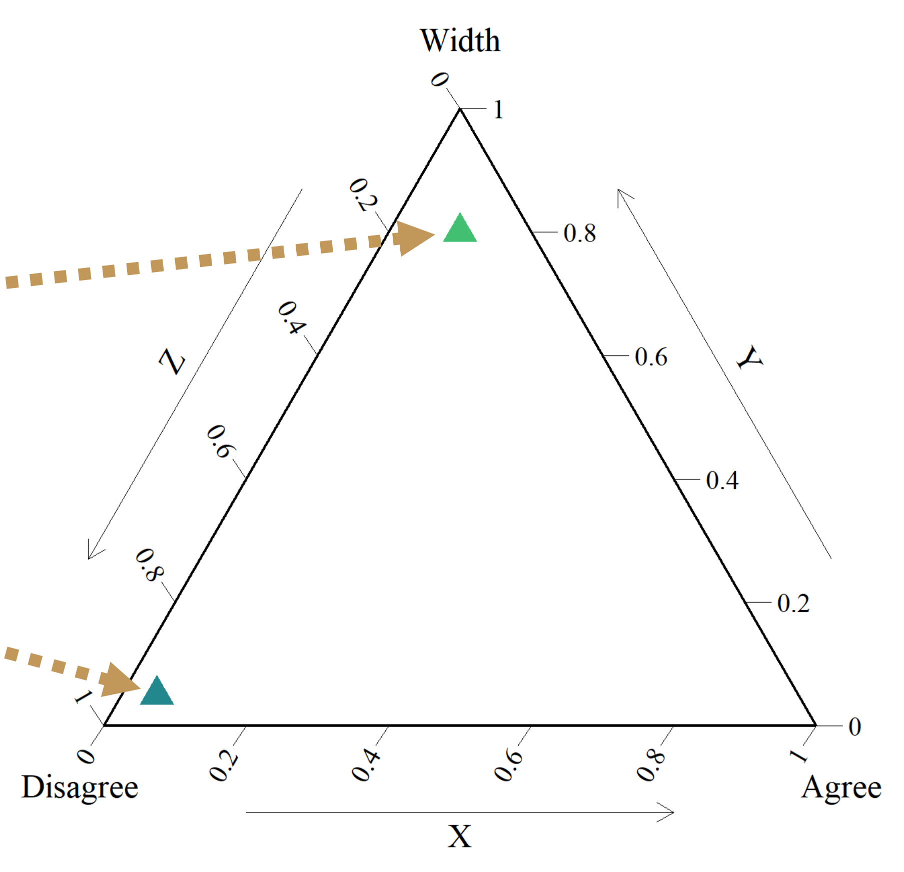
Response intervals



Location in ternary space

RESPONSE INTERVALS - TERNARY SPACE



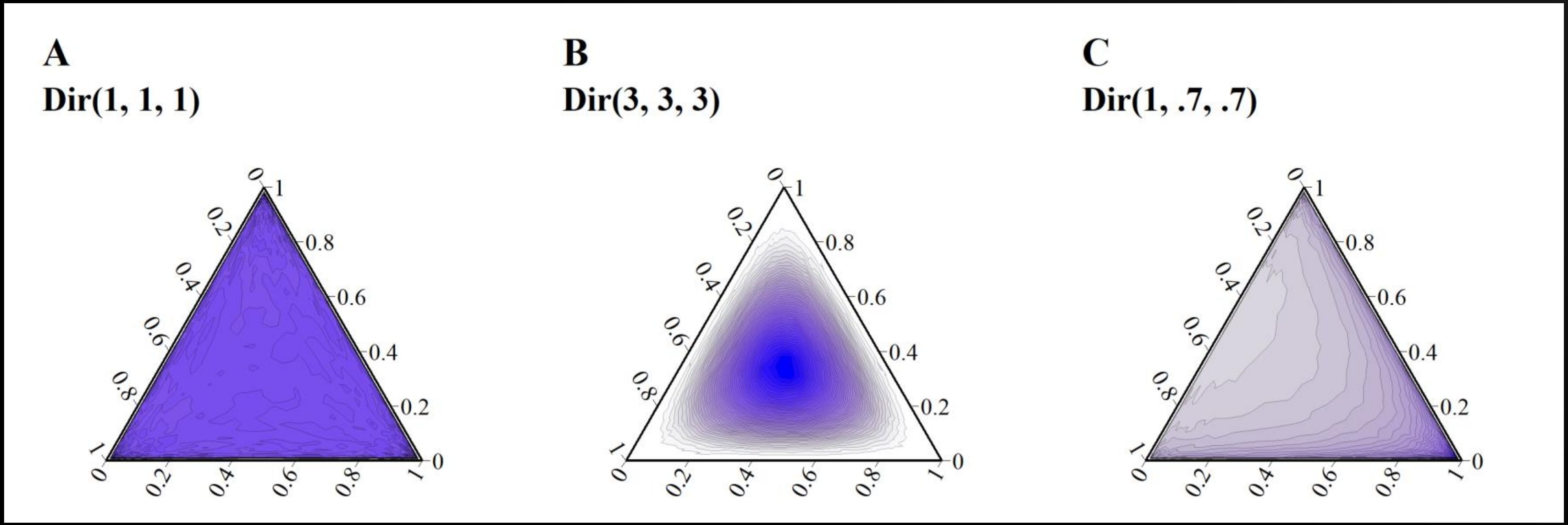


Response intervals

Location in ternary space

Dirichlet Distribution

The beta distribution generalizes to the Dirichlet distribution.

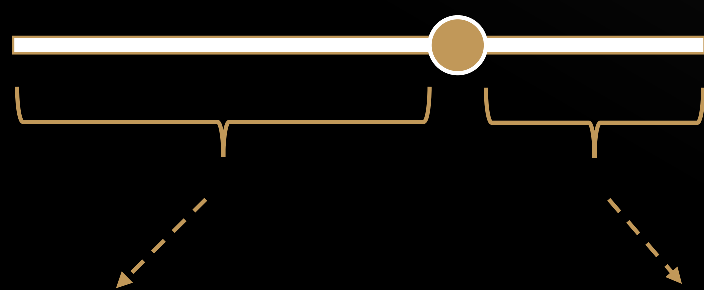


Uniform Unimodal Multimodal

Beta Response Model (BRM)
Noel & Dauvier (2007)

$x_{ij} \sim \text{Beta}\left(a_{ij}, d_{ij}\right);$

$$E\left(x_{ij}\right) = \frac{a_{ij}}{a_{ij}+d_{ij}}$$



$$a_{ij} = \exp\left[\alpha\left(\theta_i - \delta_j\right) + \tau_j\right] \qquad d_{ij} = \exp\left[-\alpha\left(\theta_i - \delta_j\right) + \tau_j\right]$$

Dirichlet Dual Response Model (DDRM)

Latent Parameterization

$$\boldsymbol{y}_{ij} \sim \text{Dirichlet}(a_{ij}, e_{ij}, d_{ij});$$

$$E\left(\boldsymbol{y}_{ij}\right) = \frac{a_{ij}}{a_{ij}+e_{ij}+d_{ij}}, \frac{e_{ij}}{a_{ij}+e_{ij}+d_{ij}}, \frac{d_{ij}}{a_{ij}+e_{ij}+d_{ij}}$$

$$a_{ij} = \exp[\alpha_\lambda(\theta_i - \delta_j) + \tau_j] \qquad d_{ij} = \exp[-\alpha_\lambda(\theta_i - \delta_j) + \tau_j]$$

$$e_{ij} = \exp[\alpha_\epsilon(\eta_i + \gamma_j) + \tau_j]$$

Dirichlet Dual Response Model (DDRM)

Parameters: Precision

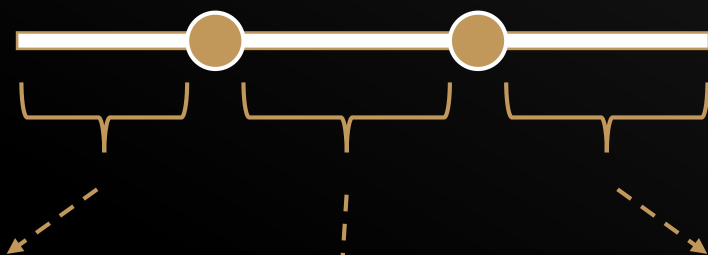
$$a_{ij} = \exp\left[\alpha_\lambda(\theta_i - \delta_j) + \tau_j\right] \qquad d_{ij} = \exp\left[-\alpha_\lambda(\theta_i - \delta_j) + \tau_j\right]$$

$$e_{ij} = \exp\left[\alpha_\epsilon(\eta_i + \gamma_j) + \tau_j\right]$$

- Location dimension: equivalent to the BRM
- Expansion dimension: controls the interval width
- Scaling $\alpha_\lambda/\alpha_\epsilon$ per dimension
- Precision τ_j across both dimensions

Dirichlet Dual Response Model (DDRM)

Parameters: Expansion Dimension



$$a_{ij} = \exp\left[\alpha_\lambda\left(\theta_i - \delta_j\right) + \tau_j\right] \quad d_{ij} = \exp\left[-\alpha_\lambda\left(\theta_i - \delta_j\right) + \tau_j\right]$$

$$e_{ij} = \exp\left[\alpha_\epsilon\left(\eta_i + \gamma_j\right) + \tau_j\right]$$

η_i: Person expansion (preference for wider intervals)
γ_j: Item expansion (strength to evoke wider intervals)

➢ Higher values = wider response intervals

Dirichlet Dual Response Model (DDRM)

Exponential Link

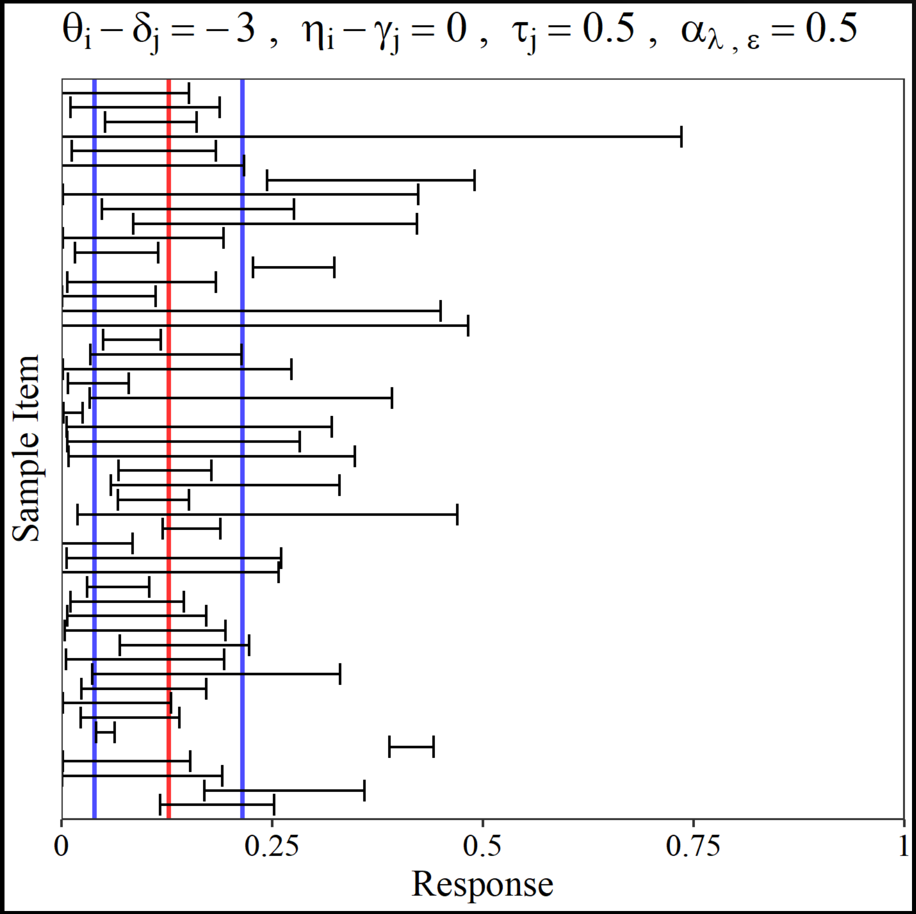
$$a_{ij} = \exp\left[\alpha_\lambda(\theta_i - \delta_j) + \tau_j\right] \qquad d_{ij} = \exp\left[-\alpha_\lambda(\theta_i - \delta_j) + \tau_j\right]$$

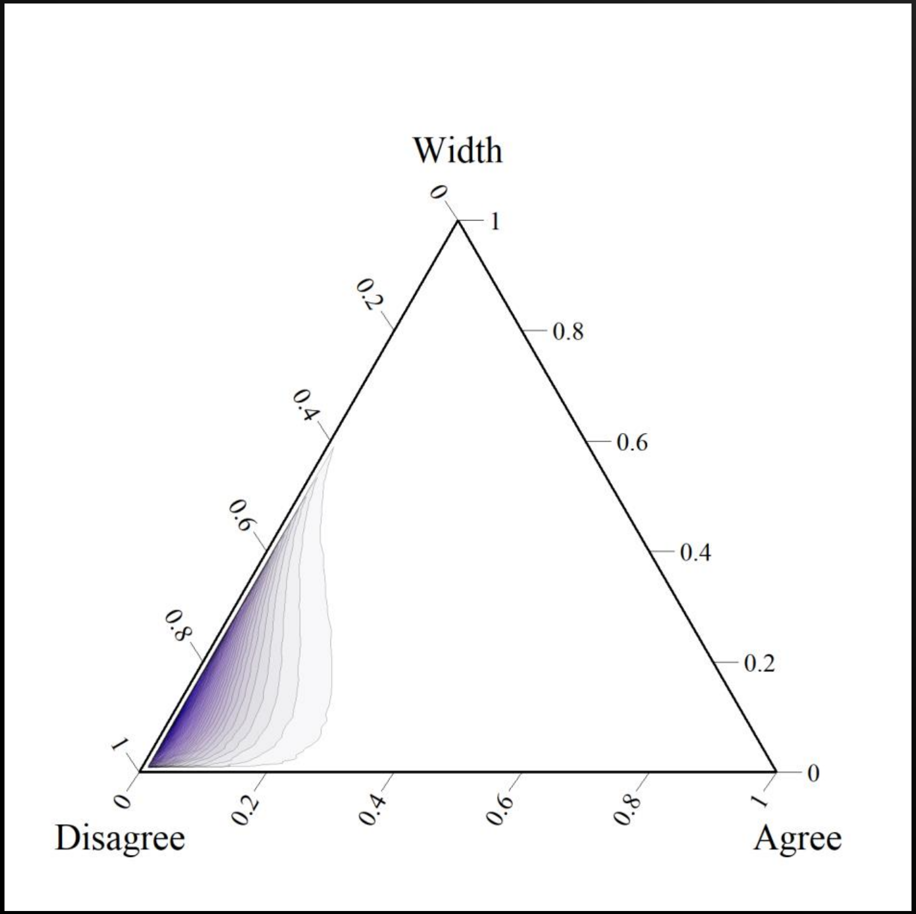
$$e_{ij} = \exp\left[\alpha_\epsilon(\eta_i + \gamma_j) + \tau_j\right]$$

Example:

- $\theta_i - \delta_j = 0;$ $\qquad \eta_i + \gamma_j = 0;$ $\qquad \tau_j = 0$
- ➢ $\exp(0) = 1$
- ➢ $Dirichlet(1,1,1)$: uniform distribution over the simplex

DDRM Examples



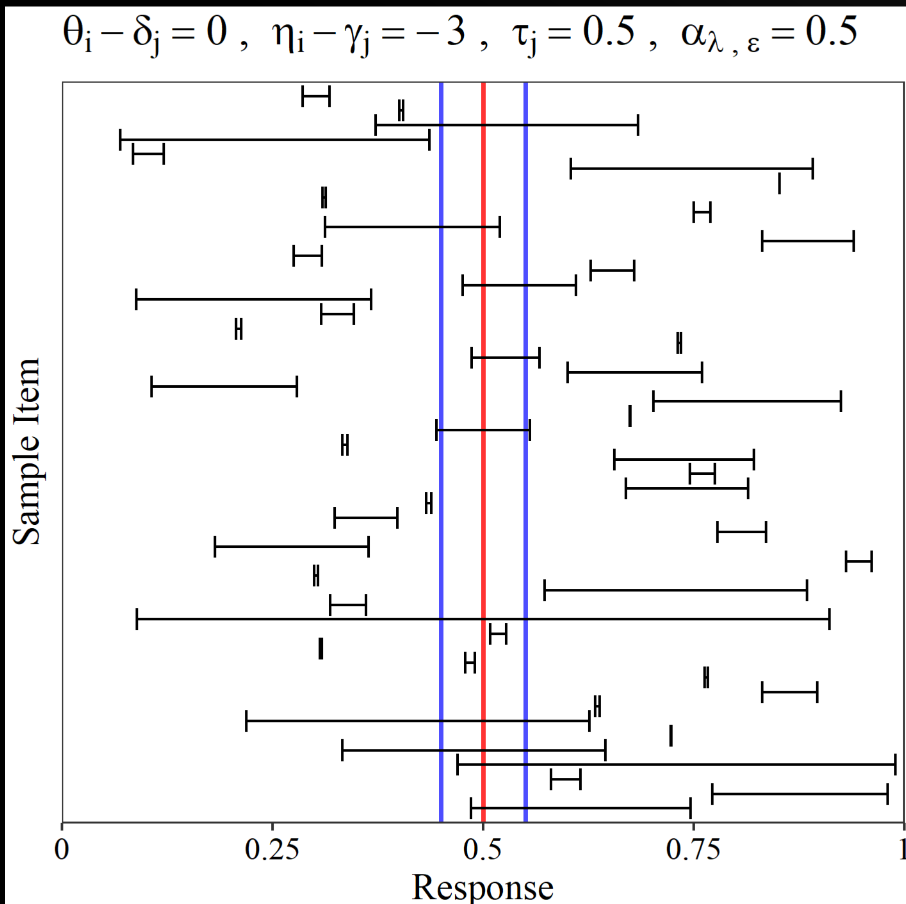


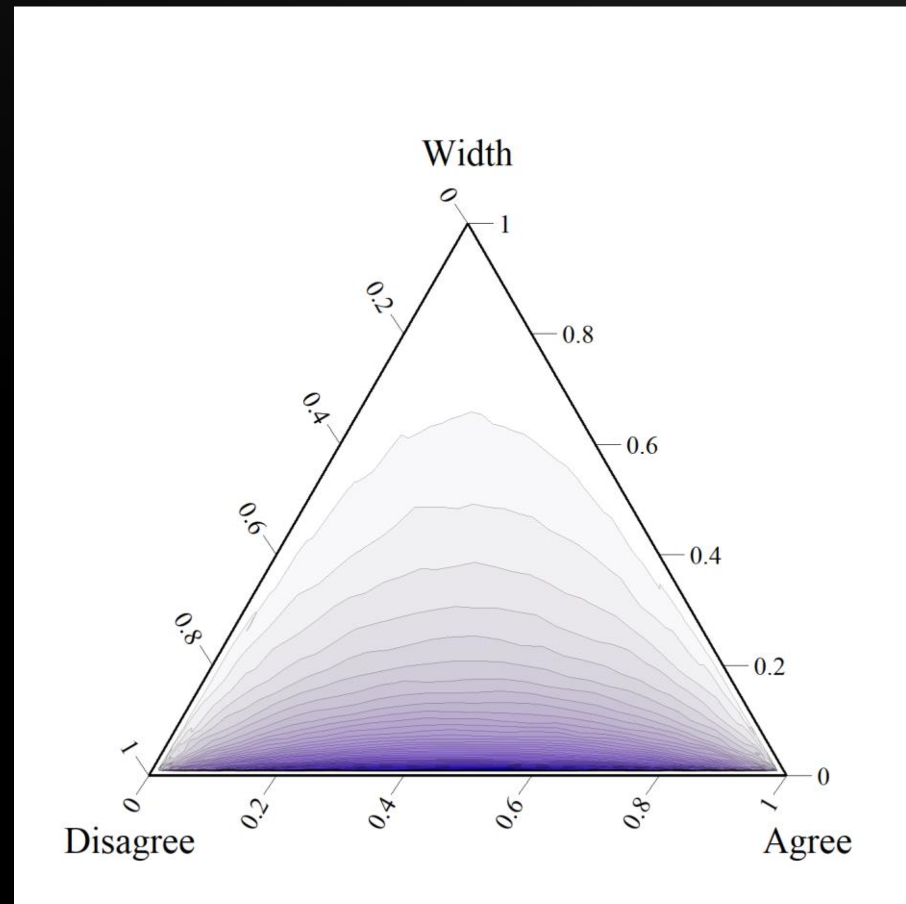
50 randomly drawn intervals

Response distribution density

- Red vertical line: expected interval location (midpoint)
- Blue vertical lines: expected lower and upper bound

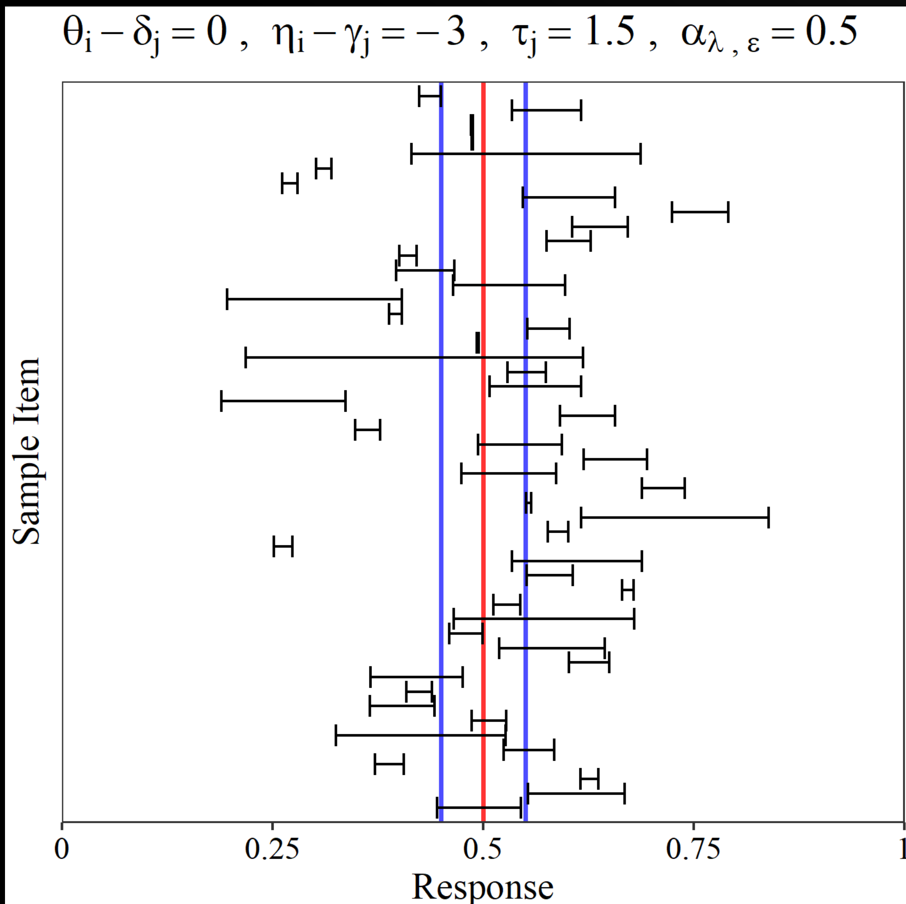
DDRM Examples
Comparison: Precision

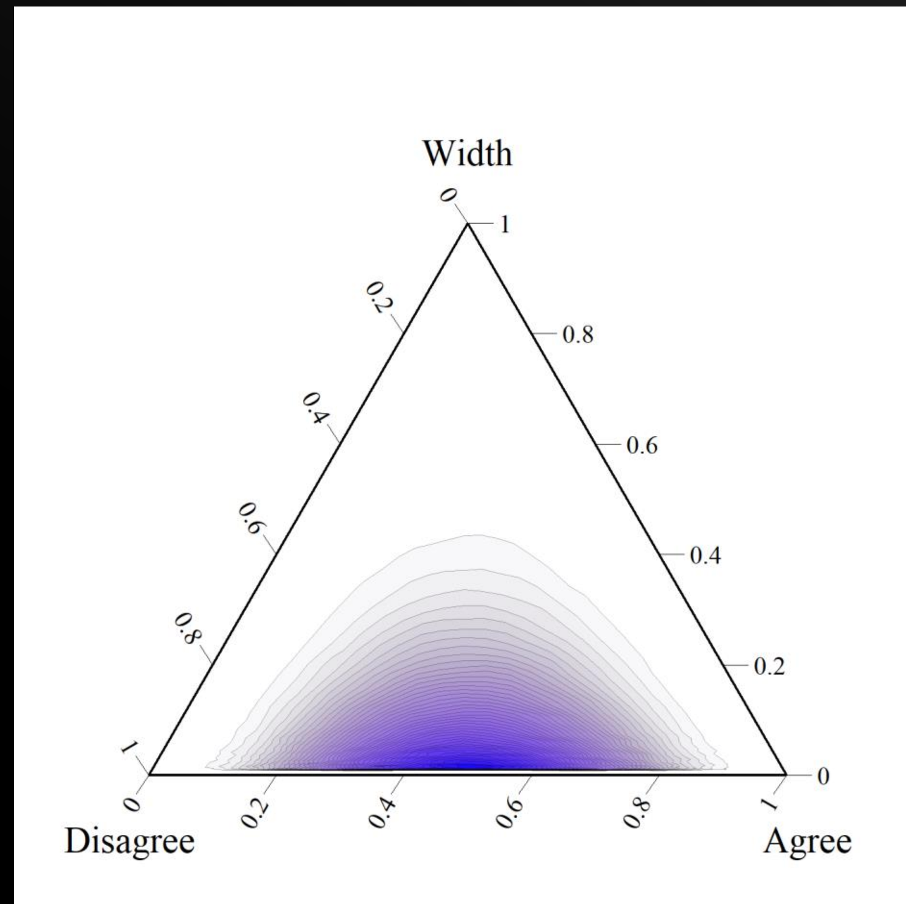




$\tau_j = 0.5$

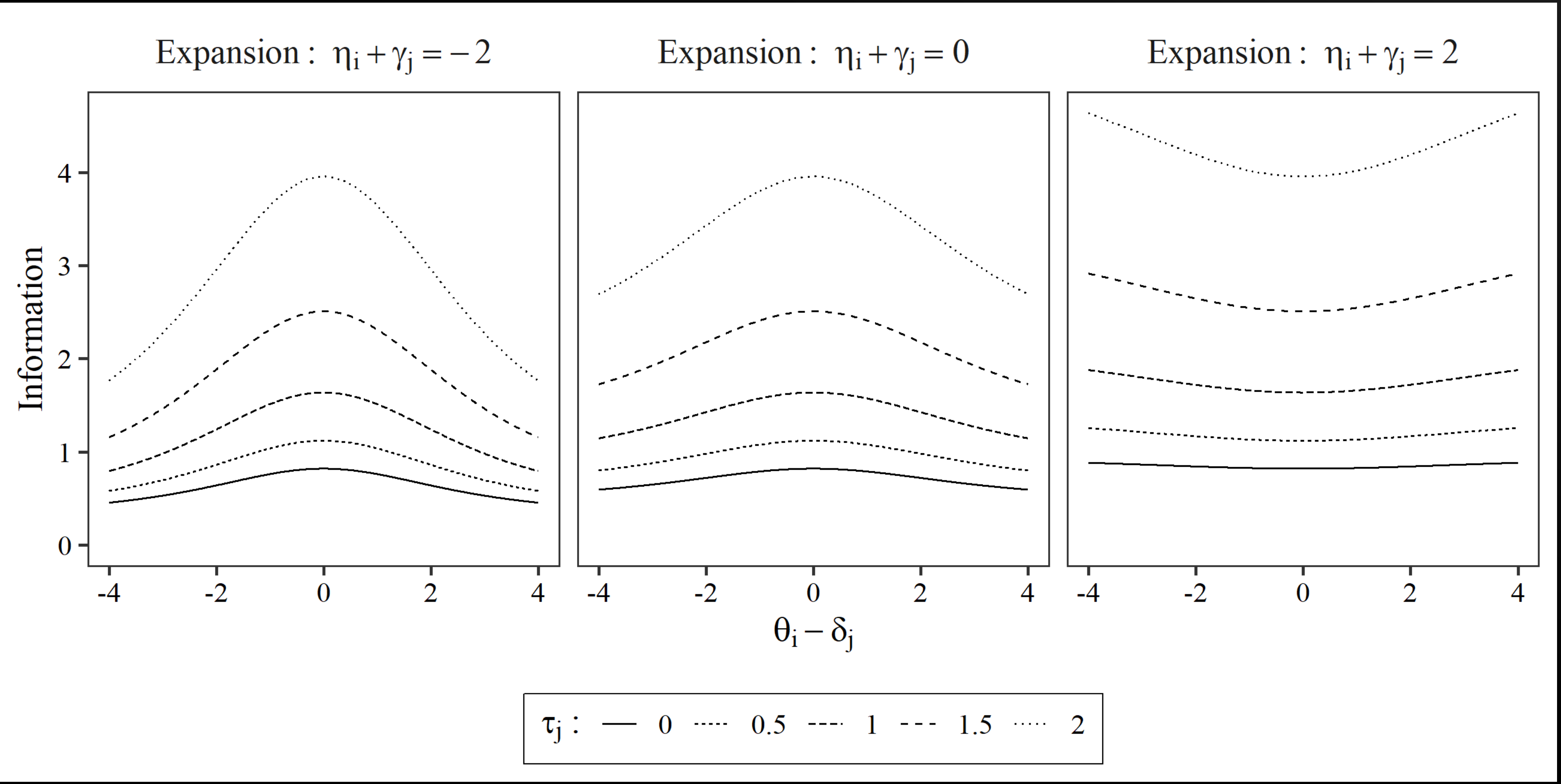
DDRM Examples
Comparison: Precision



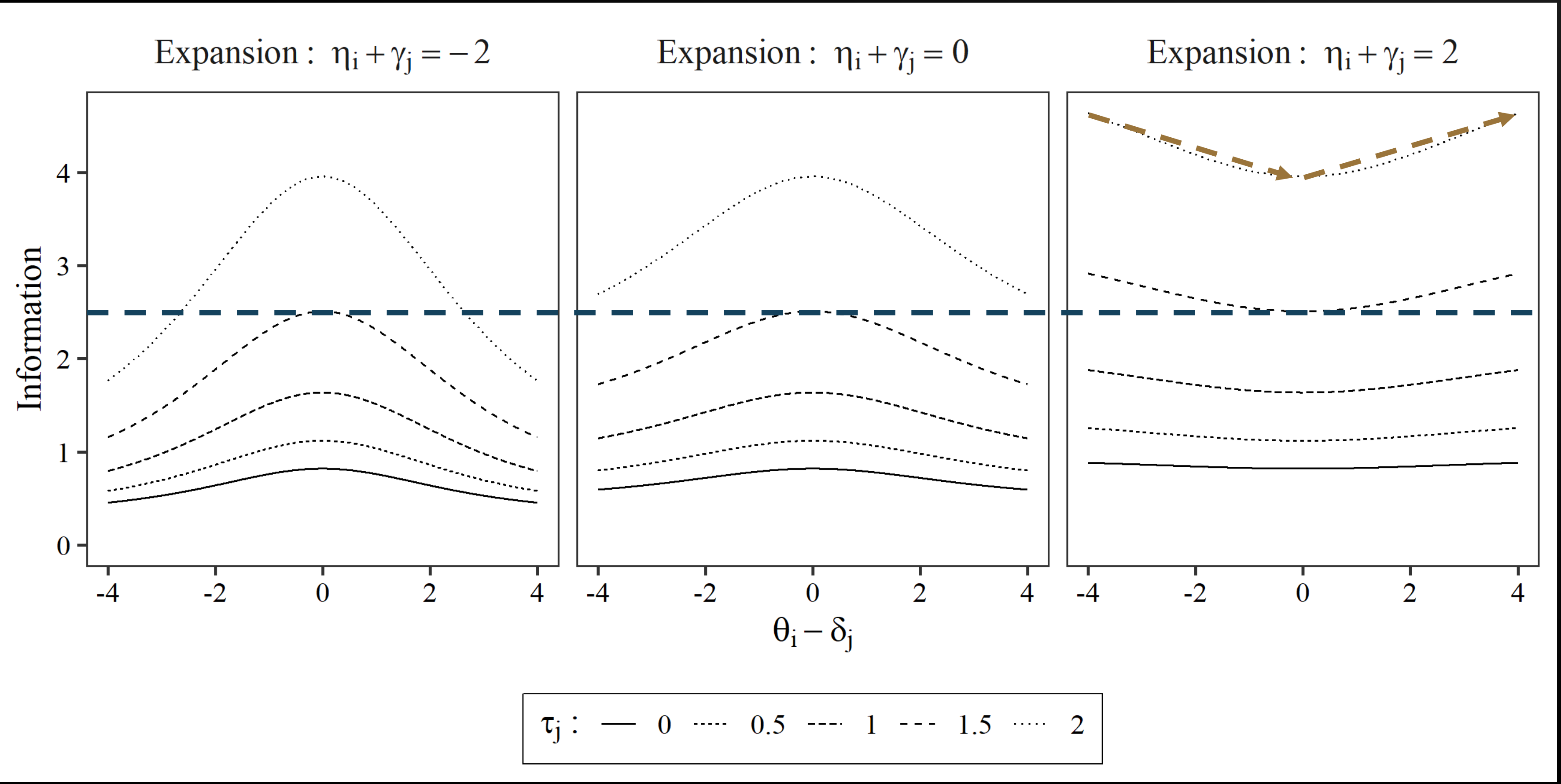


$\tau_j = 1.5$

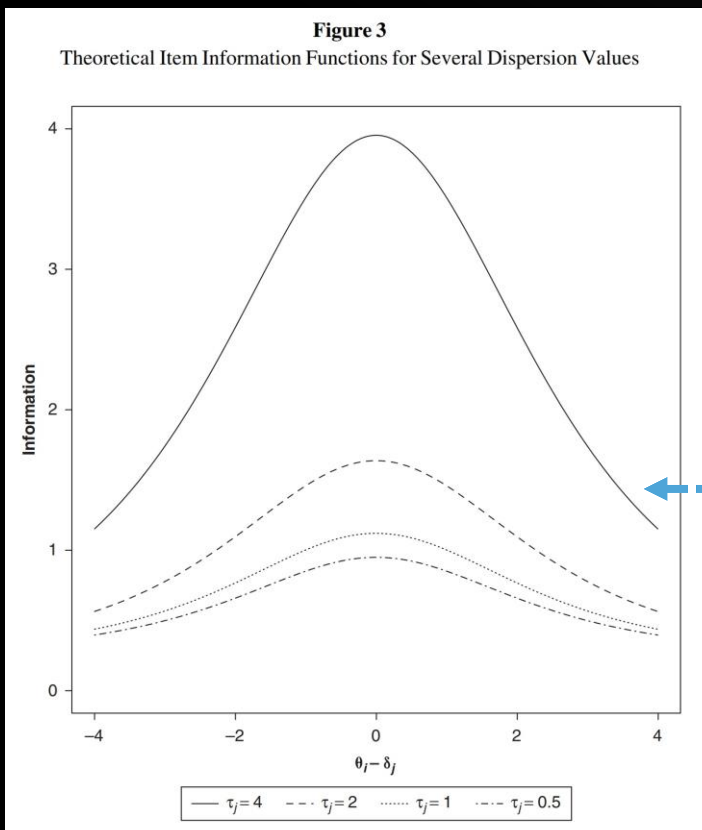
Item Information
Location Dimension



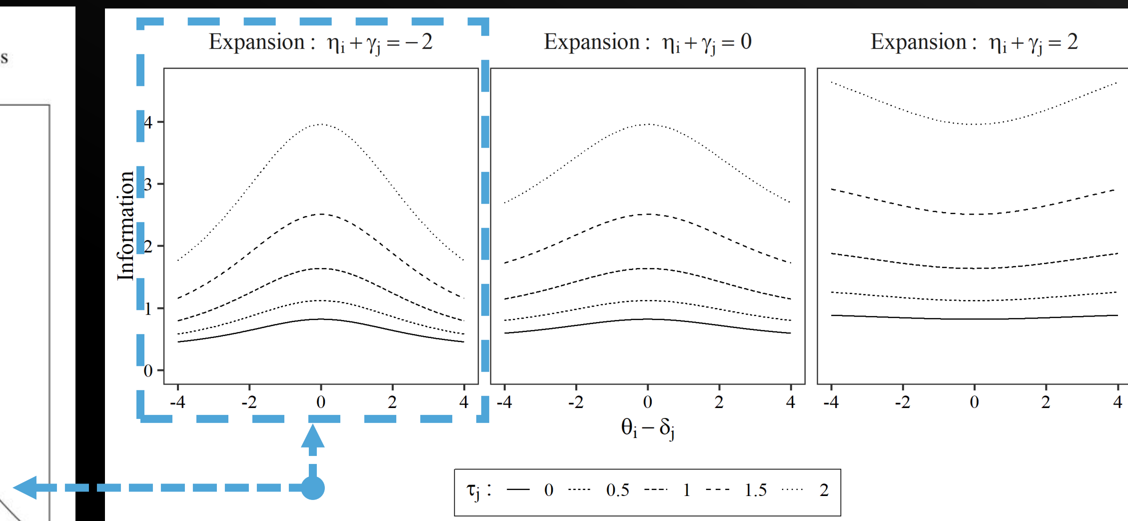
Item Information
Location Dimension



ITEM INFORMATION
Comparison with Beta Response Model

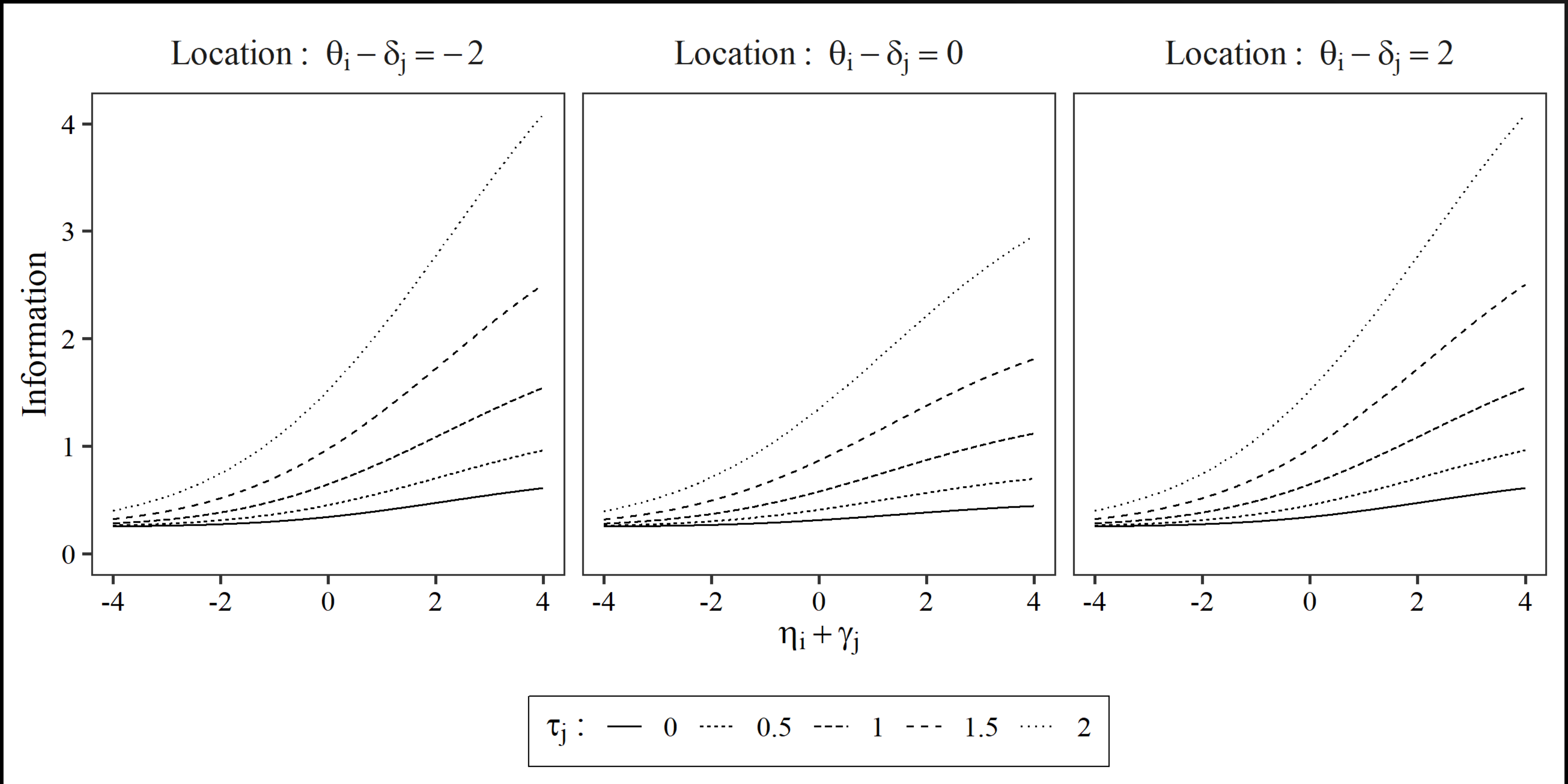


BRM (Noel & Dauvier, 2007)



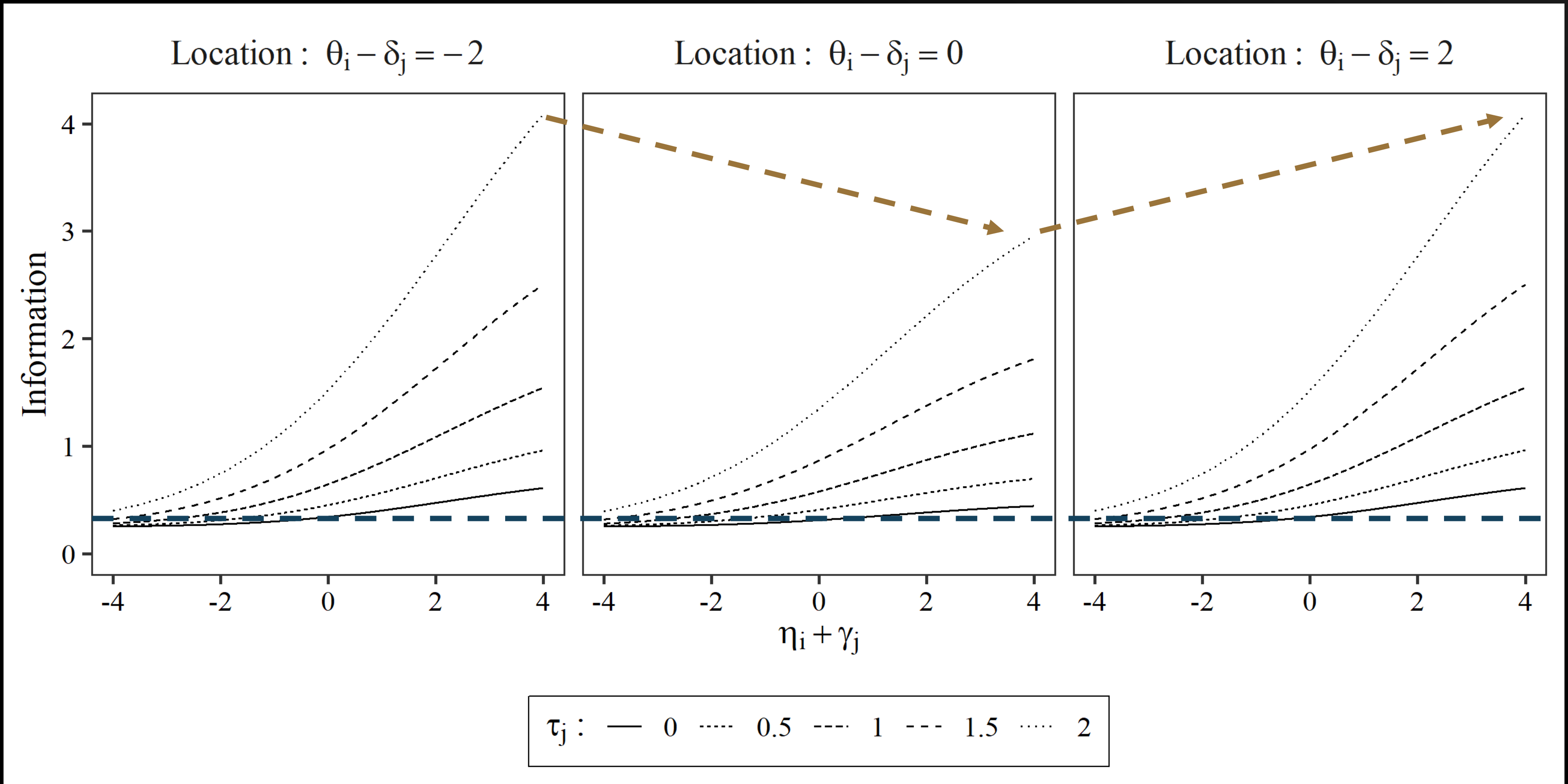
DDRM

Item Information
Expansion Dimension



ITEM INFORMATION

Expansion Dimension



Item Information
Conclusion

High sensitivity when:

- Location dimension = low / high (away from zero)
- Expansion dimension: high

➢ More information when response needs to be pushed towards the bounds of the response scale

3 - Simulation

Setup

Numbers of

- Persons: 100, 250, 500
- Items: 10, 15, 20, 30
- Replications per condition: 200

Person Parameters:

- $\theta_i, \eta_i \sim N(0,1)$

Item Parameters:

- $\delta_j, \gamma_j \sim \text{sequence } [\text{-}2, 2] \text{ by } 4/n_{items}$
- $\tau_j \sim U(0,2)$

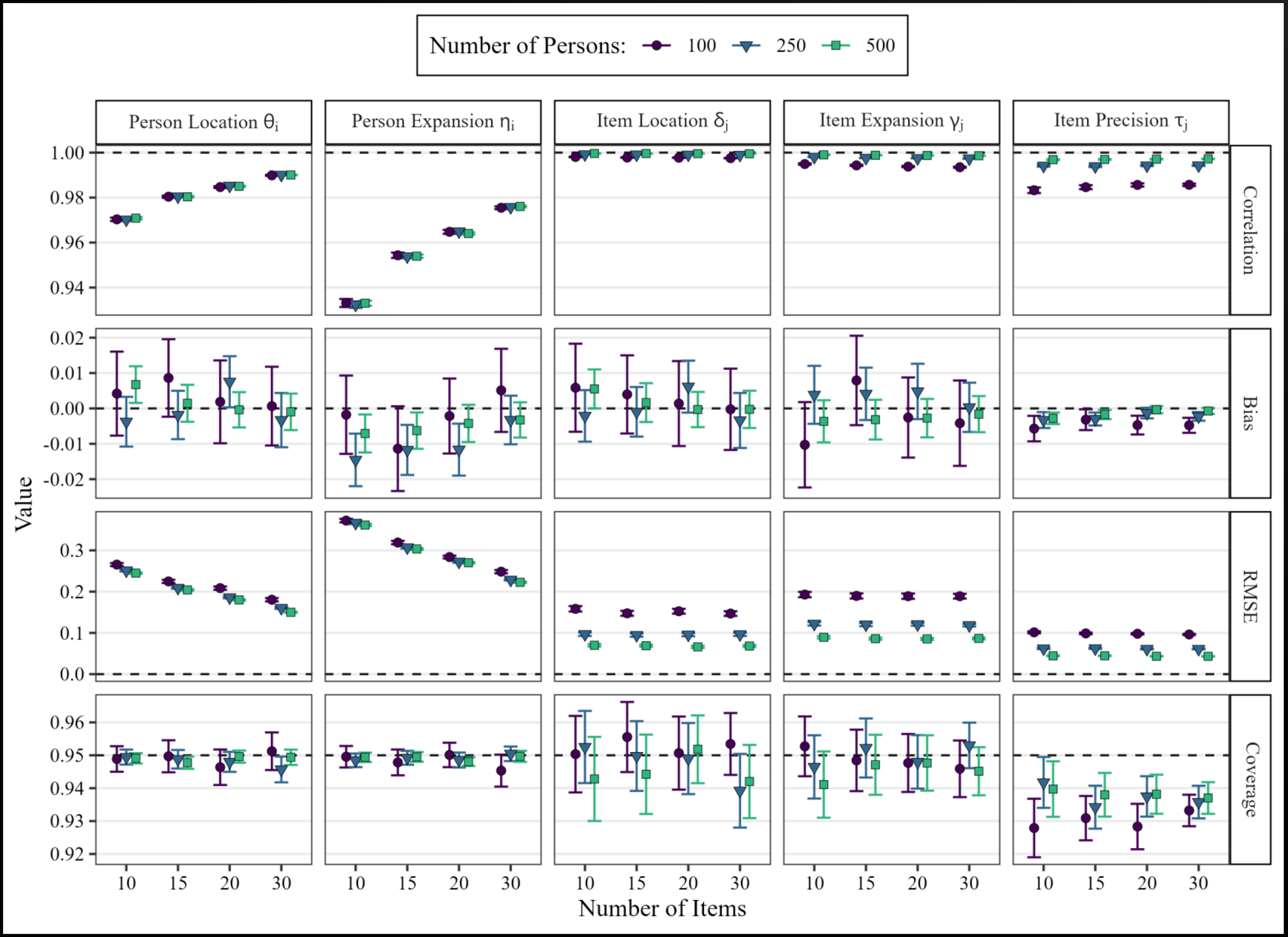
Scaling Parameters:

- $\alpha_\lambda, \alpha_\epsilon = 0.5$

Fit Measures

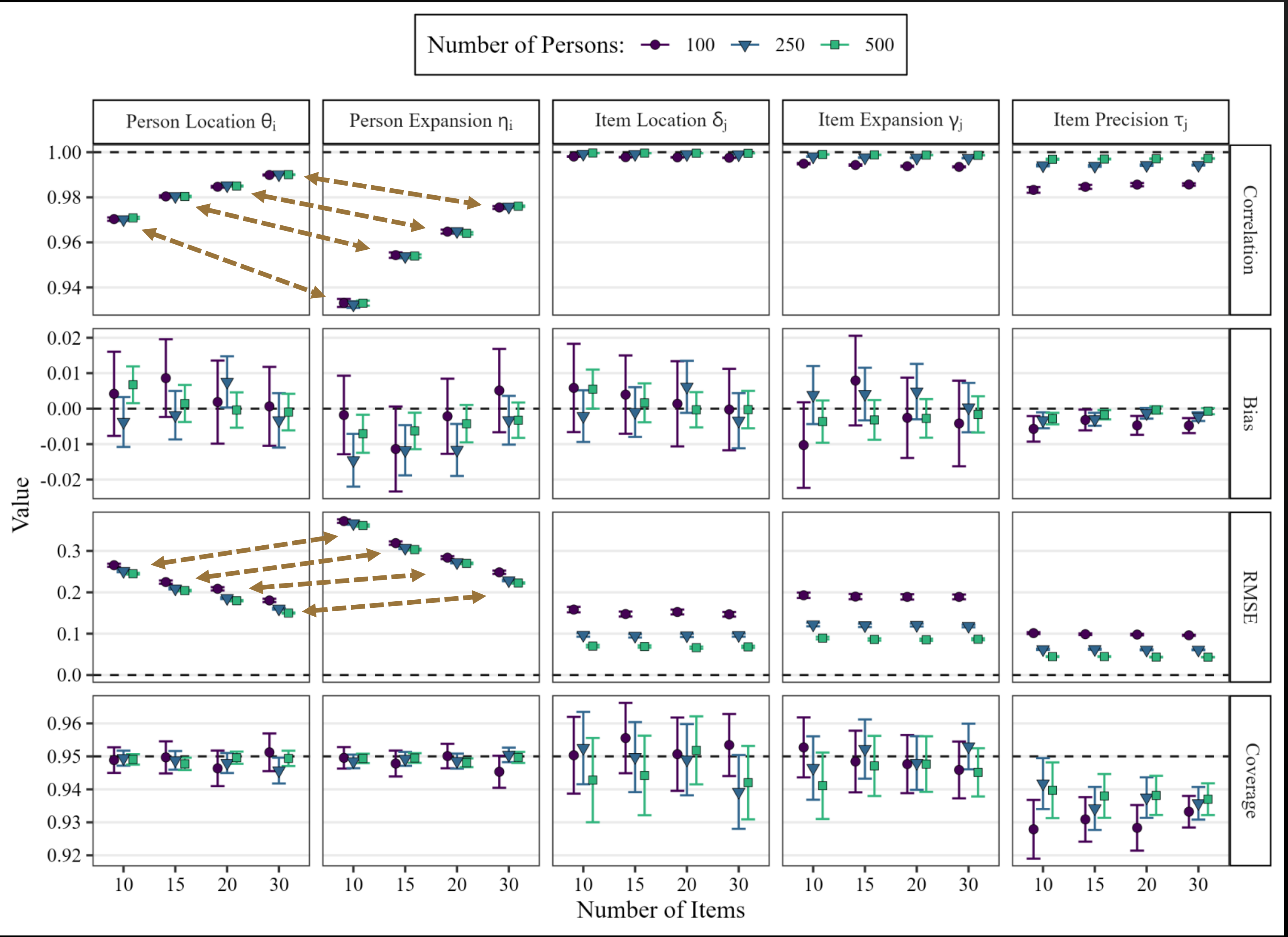
- Correlation: true vs. estimated
- Mean Signed Difference (Bias)
- Root Mean Squared Error (RMSE)
- Coverage: 90%CIs

Results



Number of Persons: 100 250 500
Person Location θ_i
Person Expansion η_i
Item Location δ_j
Item Expansion γ_j
Item Precision τ_j
Correlation
Bias
RMSE
Coverage
Value
Number of Items
1.00
0.98
0.96
0.94
0.02
0.01
0.00
-0.01
-0.02
0.3
0.2
0.1
0.0
0.96
0.95
0.94
0.93
0.92
10
15
20
30

RESULTS



Number of Persons: 100 250 500
Person Location θ_i
Person Expansion η_i
Item Location δ_j
Item Expansion γ_j
Item Precision τ_j
Correlation
Bias
RMSE
Coverage
Value
Number of Items
10 15 20 30

Recommendations

- Use more than 200 persons
- Use more than 15 items

4 – Empirical Example

Empirical Example

Methods

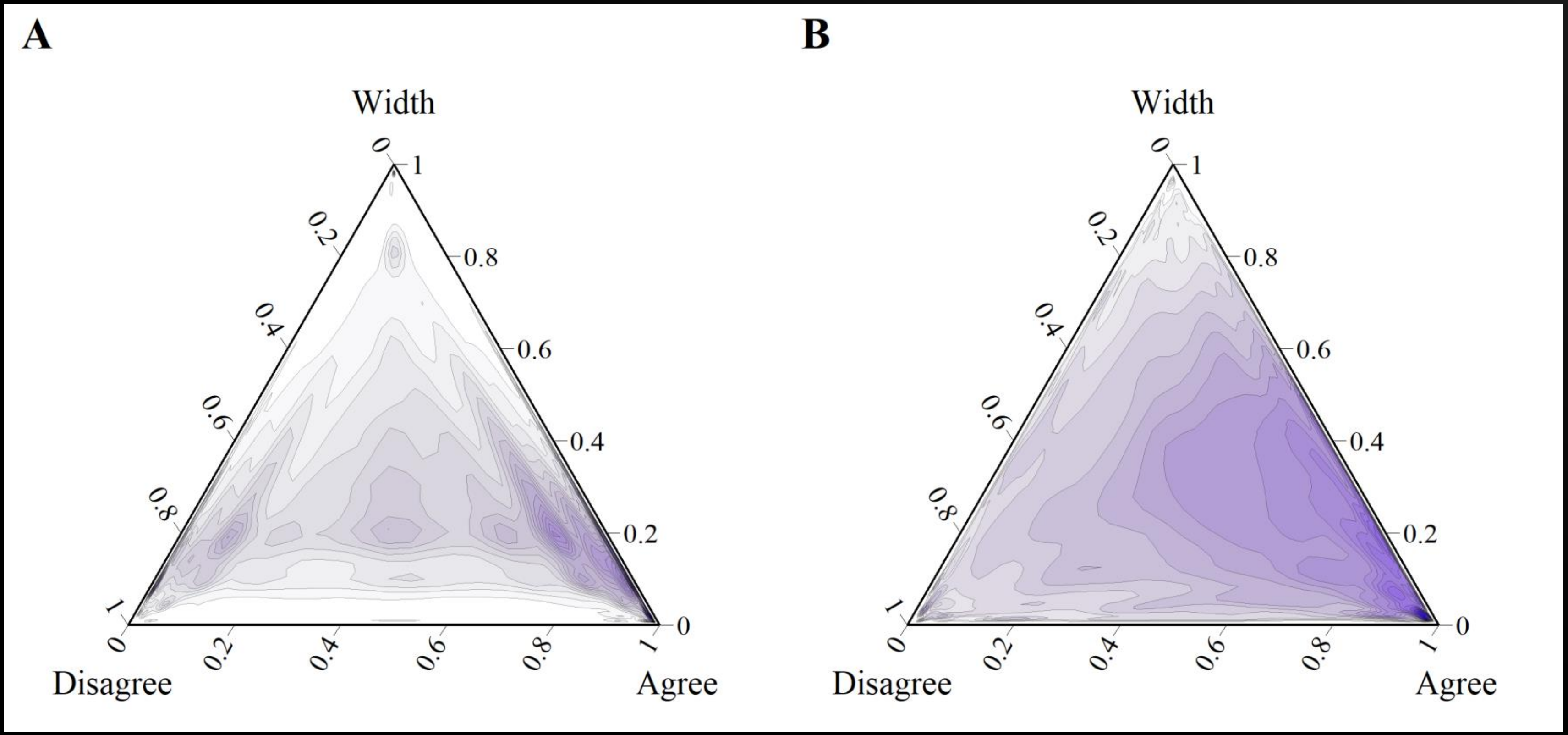
Two Extraversion scales:

- IPIP: 36 items (Interval Responses)
- BFI-2: 12 items (Single Responses)

Sample: n = 222 (f: 140 , m: 80, d: 2)

POSTERIOR PREDICTIVE CHECKS

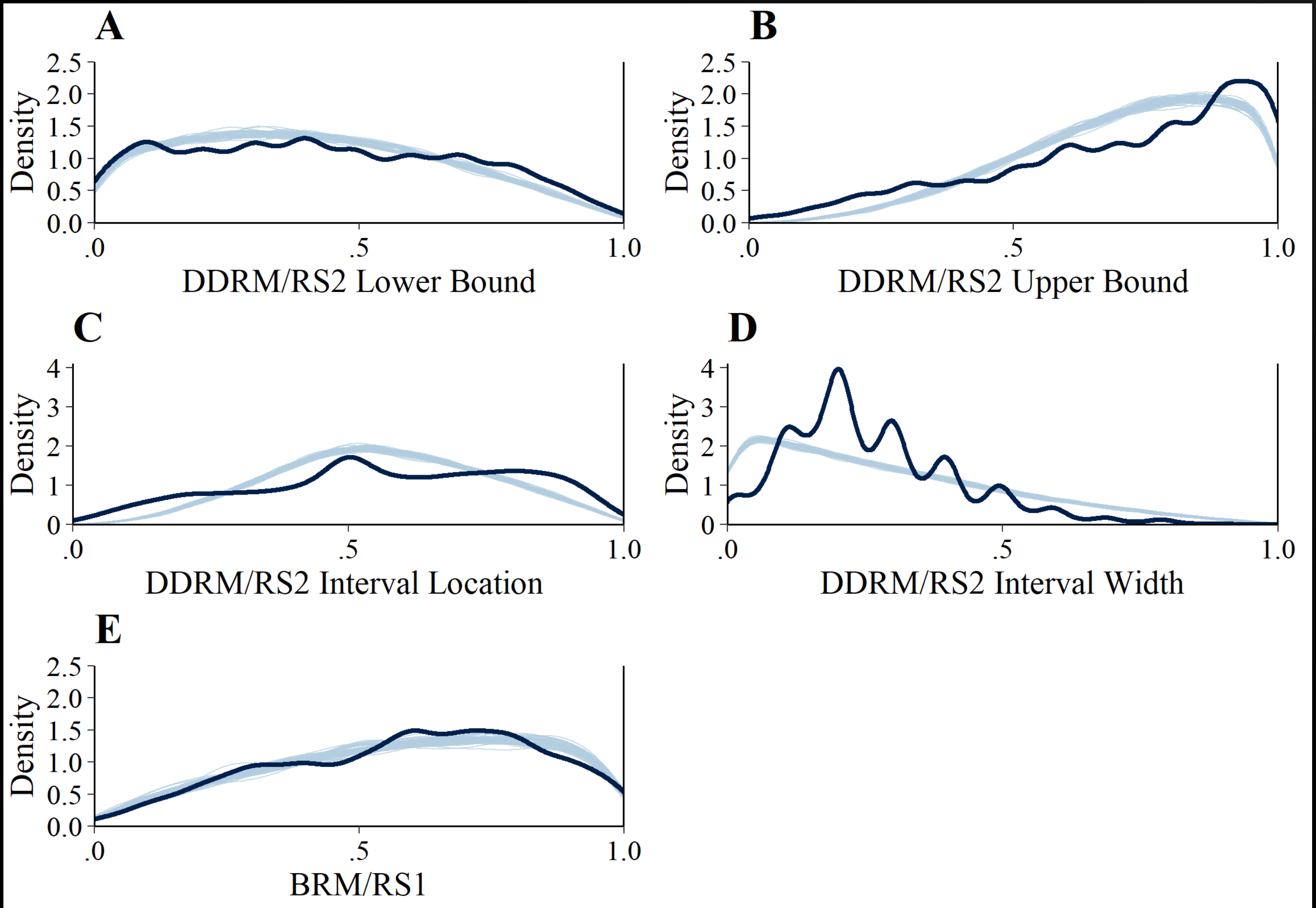
Ternary



Empirical

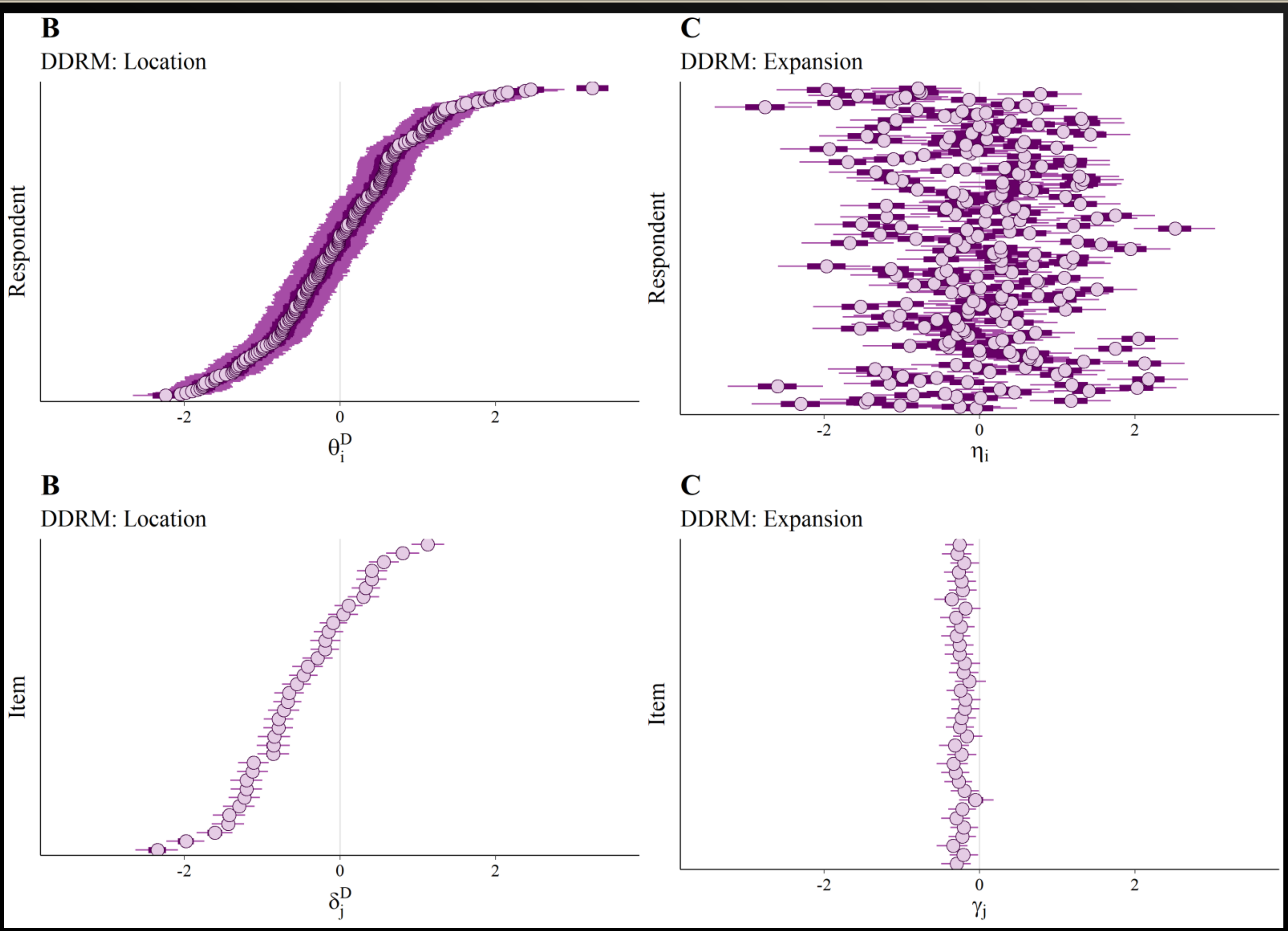
Replicated

POSTERIOR PREDICTIVE CHECKS
Binary Marginal Densities



Dark lines: empirical;   Lightlines: replicated

Parameter Estimates



Top: Person

Bottom: Item

5 - Why Do We Need A Model?

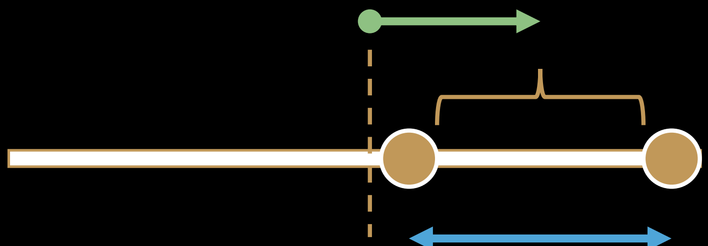
Boundedness

Scale-Inherent Correlation

- ADSM: Absolute Distance from Scale Midpoint

ADSM:

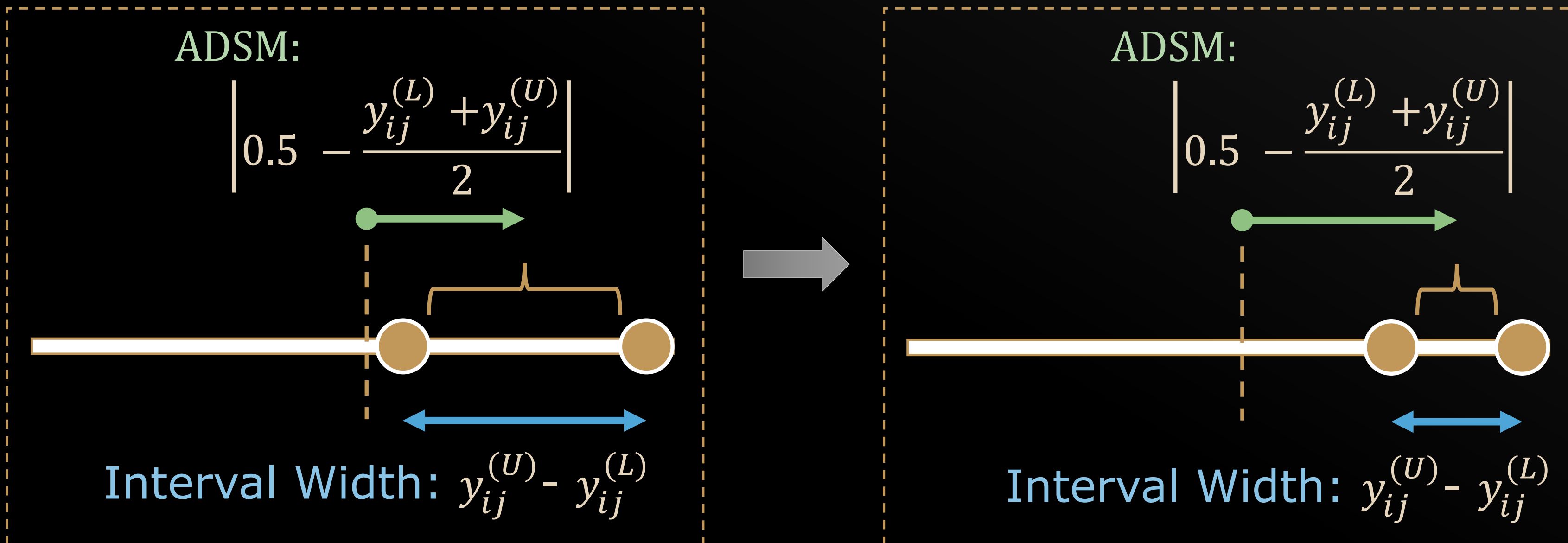
$$\left| 0.5 - \frac{y_{ij}^{(L)} + y_{ij}^{(U)}}{2} \right|$$



Interval Width: $y_{ij}^{(U)} - y_{ij}^{(L)}$

Boundedness
Scale-Inherent Correlation

- ADSM: Absolute Distance from Scale Midpoint



- Negative correlation between ADMS and Interval Width

Boundedness

Comparison of Correlations

Manifest correlation: person mean scores

- ADSM
- Interval Width

Latent variable correlation: person parameters

- Absolute Location ($|\theta|$; remember: $M(\theta) = 0$)
- Expansion (η)

BOUNDEDNESS
Comparison of Correlations

	Mean scores	Model parameters
Empirical:	$r = -.57$	$r = -.19$
Simulation:	$r = -.74$	$r = -.02$
True:		$(r = -.002)$

➢ The model accounts for the scale-inherent correlation

Convergent Validity: Response Formats

Mean Scores vs. Estimates

Manifest correlation: person mean scores

- Single response
- Interval location (midpoint)

Latent variable correlation: person parameters

- Person location θ_{BRM}
- Person location θ_{DDRM}

Convergent Validity: Response Formats

Correlations in the Empirical Study:

Mean scores: $r = .81$

Model Parameters: $r = .87$

➢ Latent model improves convergence

Take Home Points

- High **convergent validity** of response formats
- Model **accounts** for **boundedness**
- **Additional information**: expansion dimension
  - Validity? What does it measure?
- Useful tool for analysis of interval responses

Future Research

- Application to rating- and forecasting data
  - ➢ Cultural consensus models
  - ➢ Are certain respondents more accurate?
- Test-retest reliability
- Discriminant validity:
  - ➢ useful information vs. response biases

THANKS TO:

- Prof. Dr. Daniel W. Heck

- Prof. Dr. Andreas Voss

- Dr. Raphael Hartmann

Contact: kloft@uni-marburg.de

Slides: https://github.com/matthiaskloft/

References

- Danner, D., Rammstedt, B., Bluemke, M., Lechner, C., Berres, S., Knopf, T., Soto, C. J., & John, O. P. (2019). Das Big Five Inventar 2: Validierung eines Persönlichkeitsinventars zur Erfassung von 5 Persönlichkeitsdomänen und 15 Facetten. *Diagnostica*, 1–12. https://doi.org/10.1026/0012-1924/a000218
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*(6), 1011–1027. APA PsycArticles. https://doi.org/10.1037/0022-3514.80.6.1011
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, *7*(1), 7–28.
- Noel, Y. (2014). A beta unfolding model for continuous bounded responses. *Psychometrika*, *79*(4), 647–674. https://doi.org/10.1007/s11336-013-9361-1
- Noel, Y., & Dauvier, B. (2007). A beta item response model for continuous bounded responses. *Applied Psychological Measurement*, *31*(1), 47–73. https://doi.org/10.1177/0146621605287691
- Samejima, F. (1973). Homogeneous case of the continuous response model. *Psychometrika*, *38*(2), 203–219. https://doi.org/10.1007/BF02291114
- Soto, C. J., & John, O. P. (2017). The Next Big Five Inventory (BFI-2): Developing and Assessing a Hierarchical Model With 15 Facets to Enhance Bandwidth, Fidelity, and Predictive Power. *Journal of Personality & Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096