A Mixture Framework for Scaling with Anchoring Items

Juraj Medzihorsky

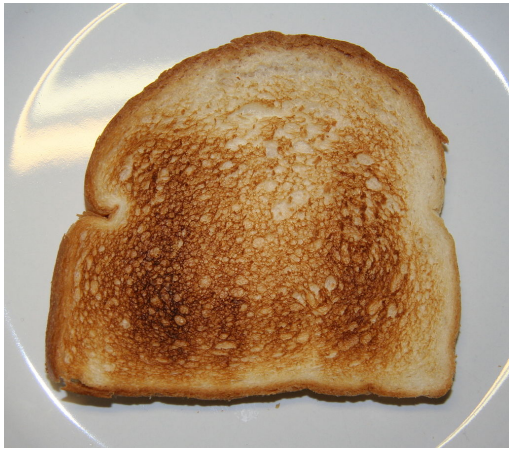
6 April 2019

The Problem

How crunchy do you like your toast?

1. Very crunchy
2. Crunchy
3. Neither crunchy nor soft
4. Soft
5. Very soft

The Solution



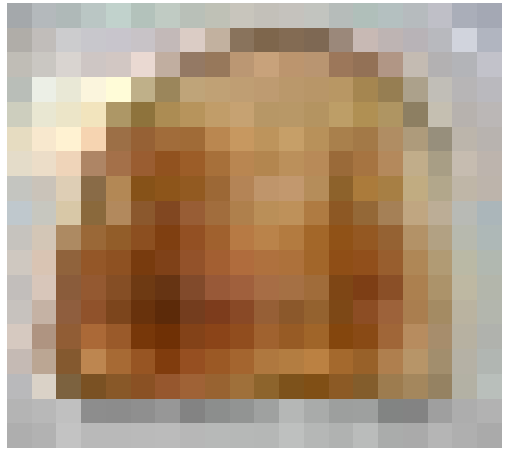
Source: Wikimedia Commons

How crunchy is this toast?

1. Very crunchy
2. Crunchy
3. Neither crunchy nor soft
4. Soft
5. Very soft

Source: Wikimedia Commons

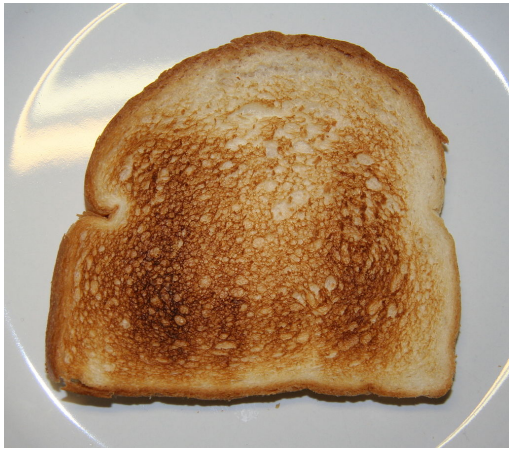
But did it work?



Adapted from Wikimedia Commons

How crunchy is this toast?

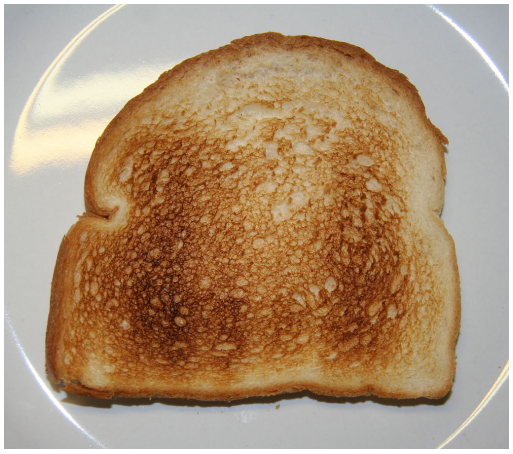
1. Very crunchy
2. Crunchy
3. Neither crunchy nor soft
4. Soft
5. Very soft



Source: Wikimedia Commons

How [redacted] is this toast?

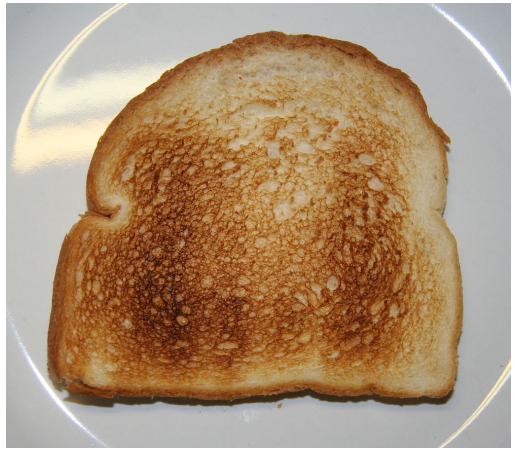
1. Very [redacted]
2. [redacted]
3. Neither [redacted] nor [redacted]
4. [redacted]
5. Very [redacted]



Source: Wikimedia Commons

How crunchy is this toast?

1. Very crunchy
2. Crunchy
3. Neither crunchy nor soft
4. Soft
5. Very soft



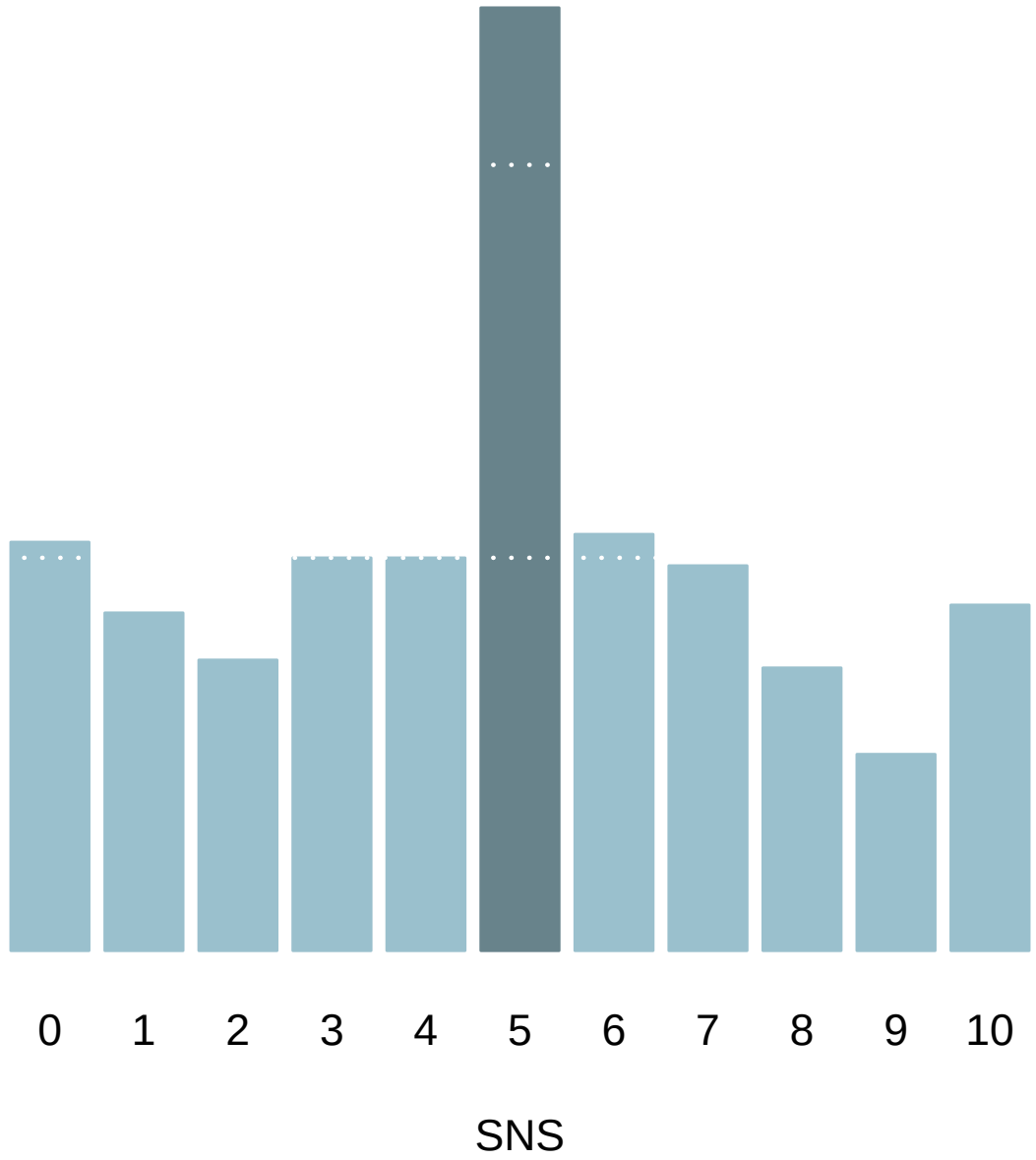
Source: Wikimedia Commons

How ~~crunchy~~ **hot** is this toast?

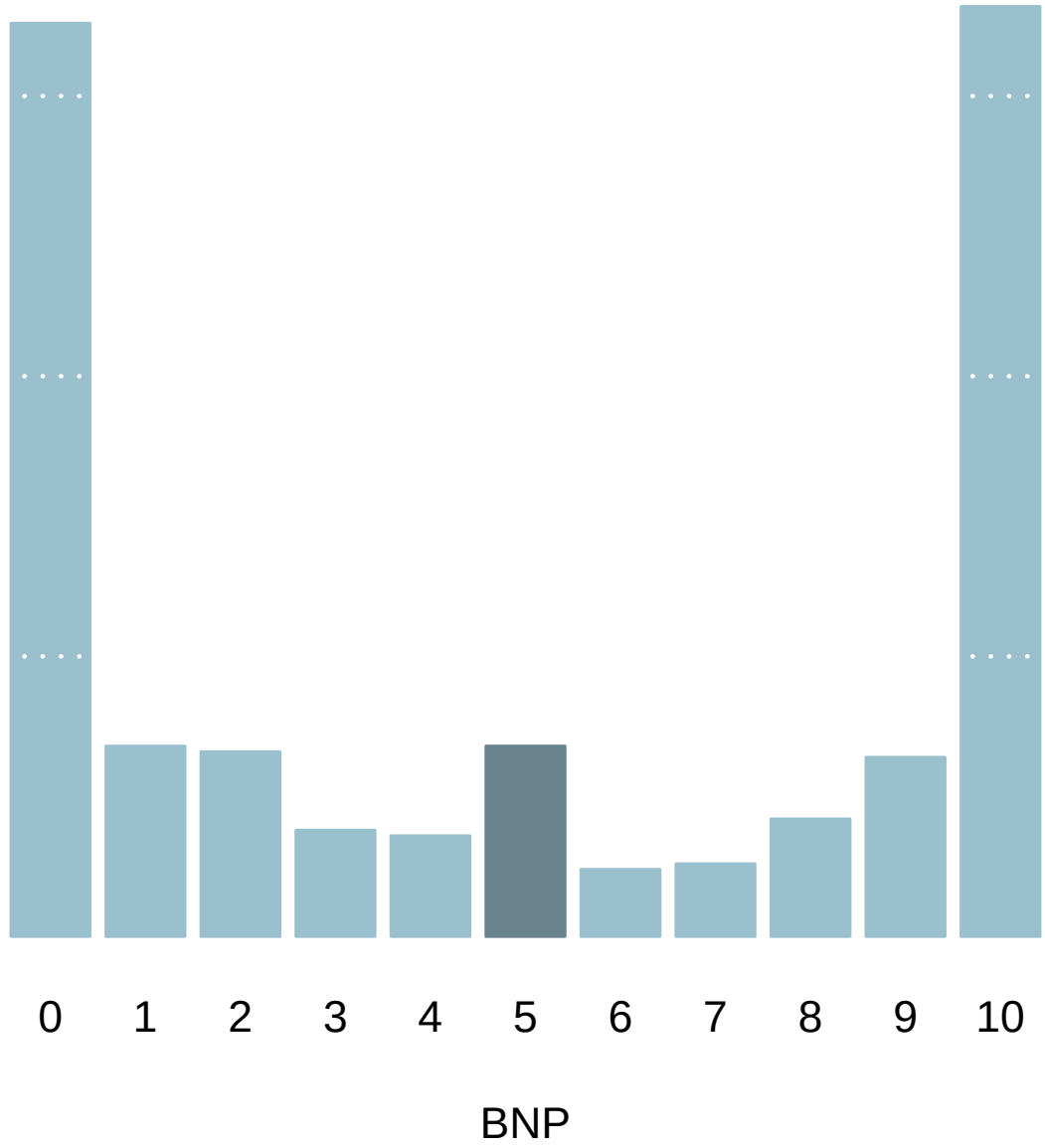
1. Very ~~crunchy~~ **hot**
2. ~~Crunchy~~ **Hot**
3. Neither ~~crunchy~~ **hot** nor ~~soft~~ **cold**
4. ~~Soft~~ **Cold**
5. Very ~~soft~~ **cold**

Source: telegraph.co.uk

EES VS L/R Batteries

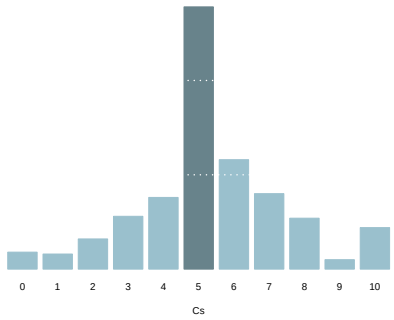


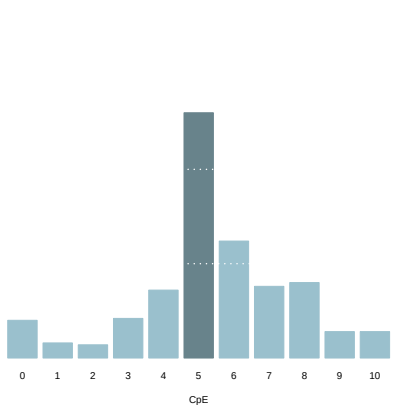
EES VS L/R Batteries

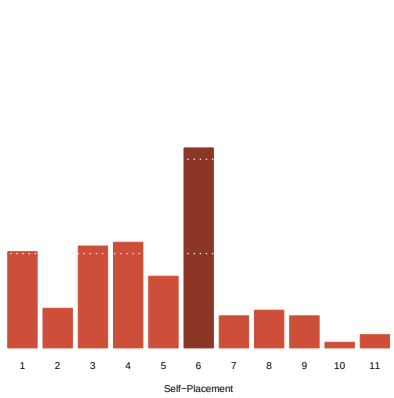


EES VS L/R Batteries

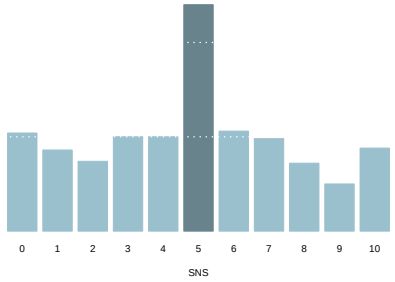
Spain 2014

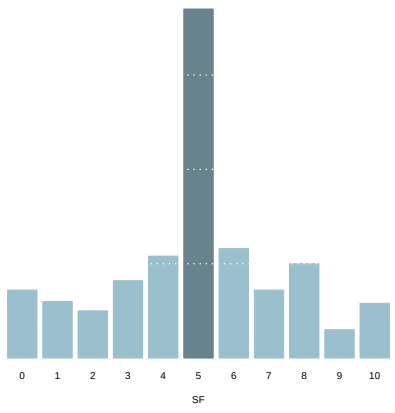


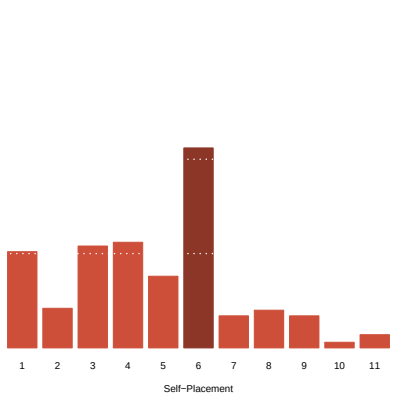




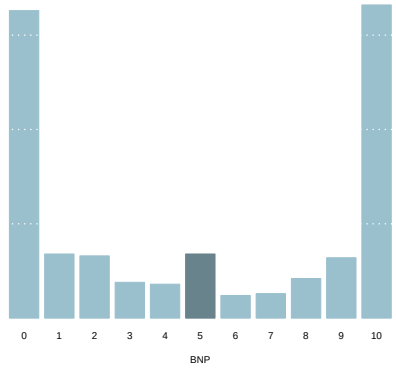
Slovakia 2009

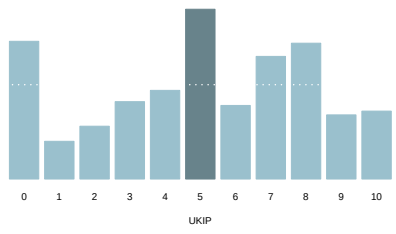


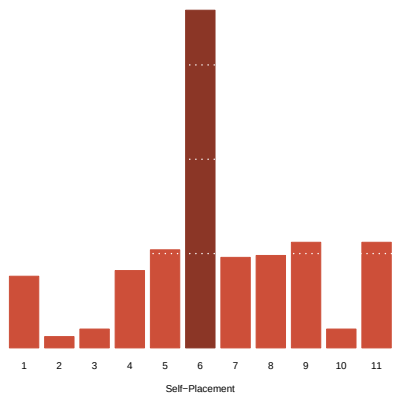




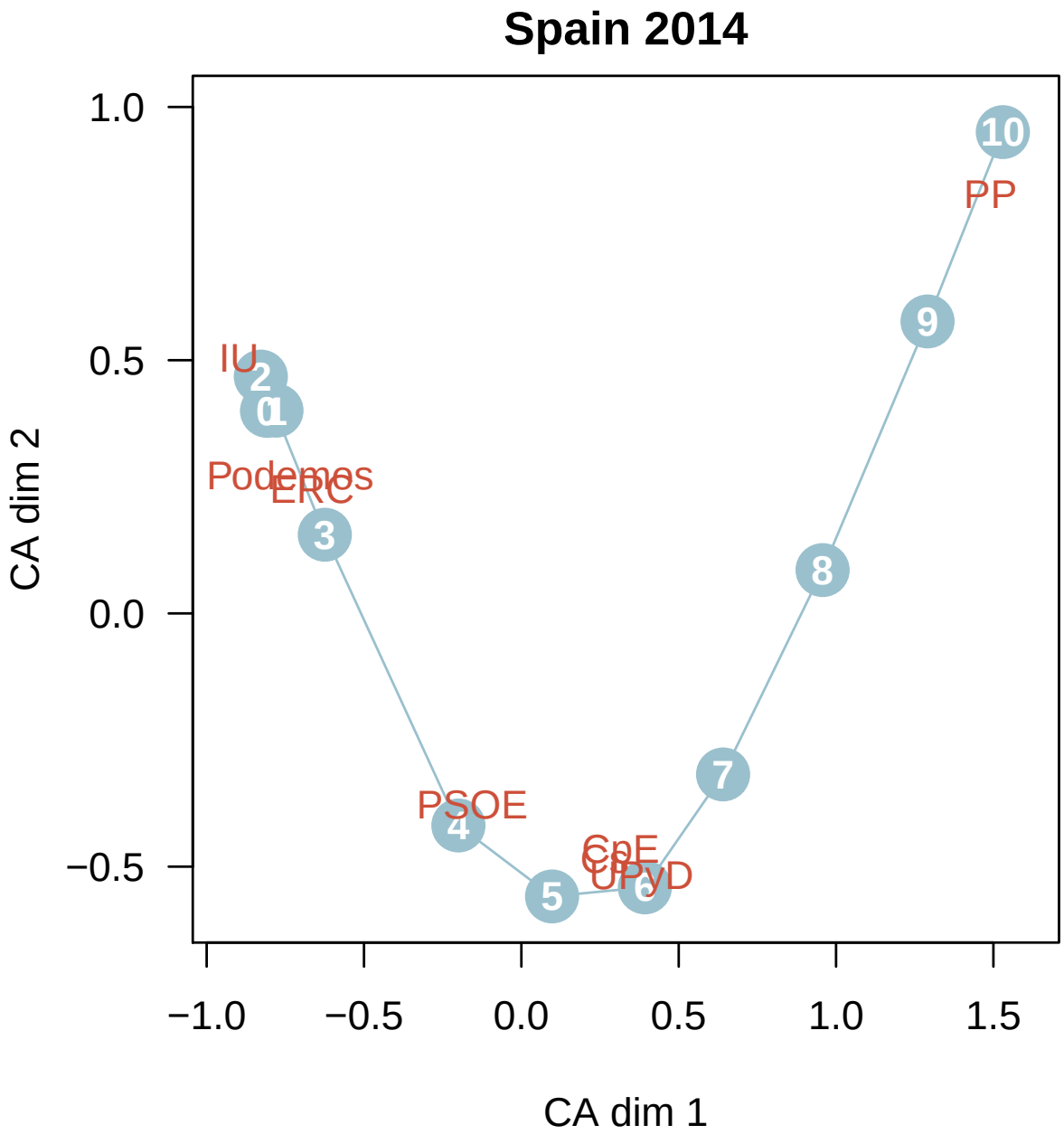
UK 2009



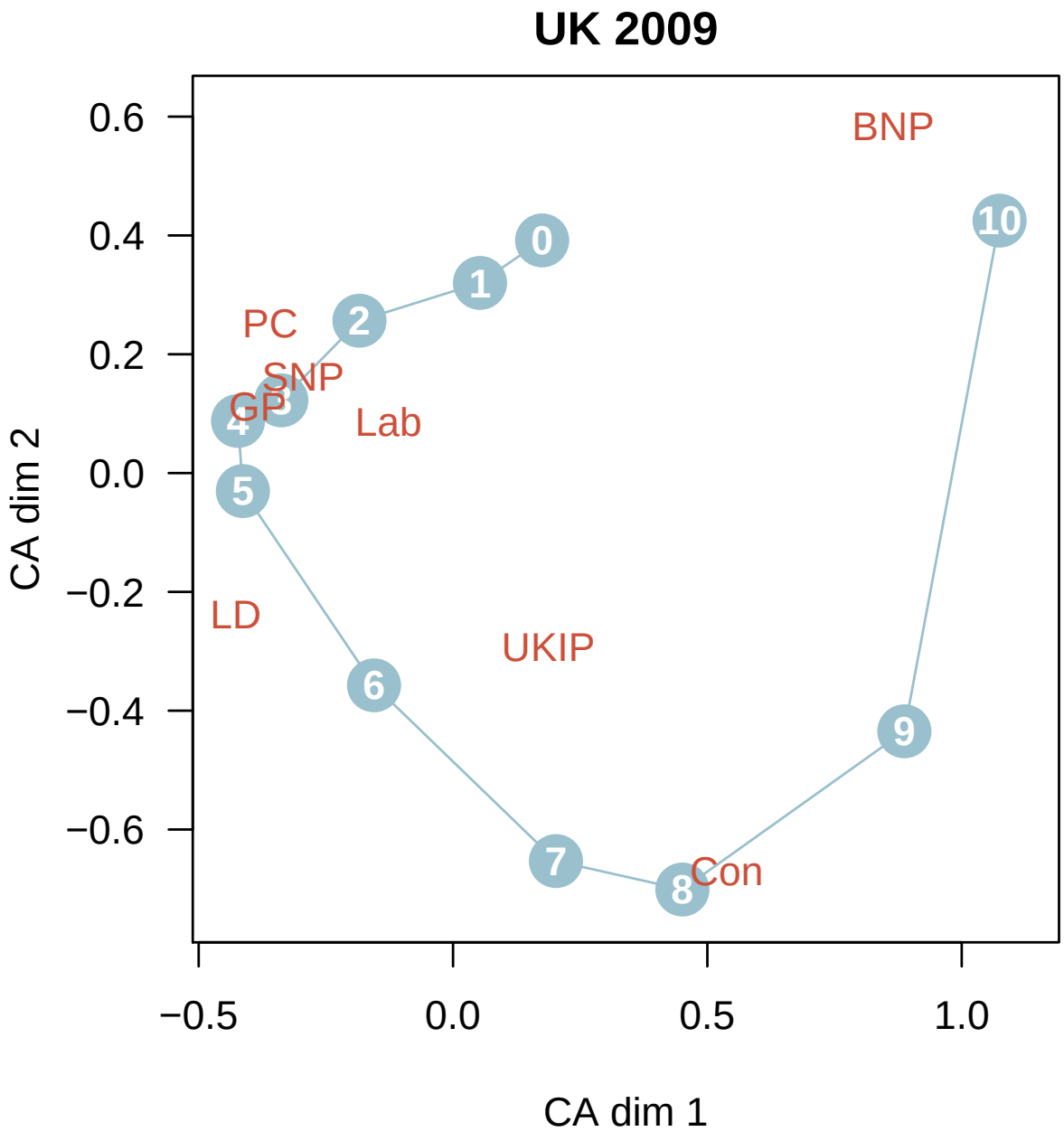




CA: EES VS L/R Batteries



CA: EES VS L/R Batteries



Parametric Scaling

Aldrich-McKelvey Scaling

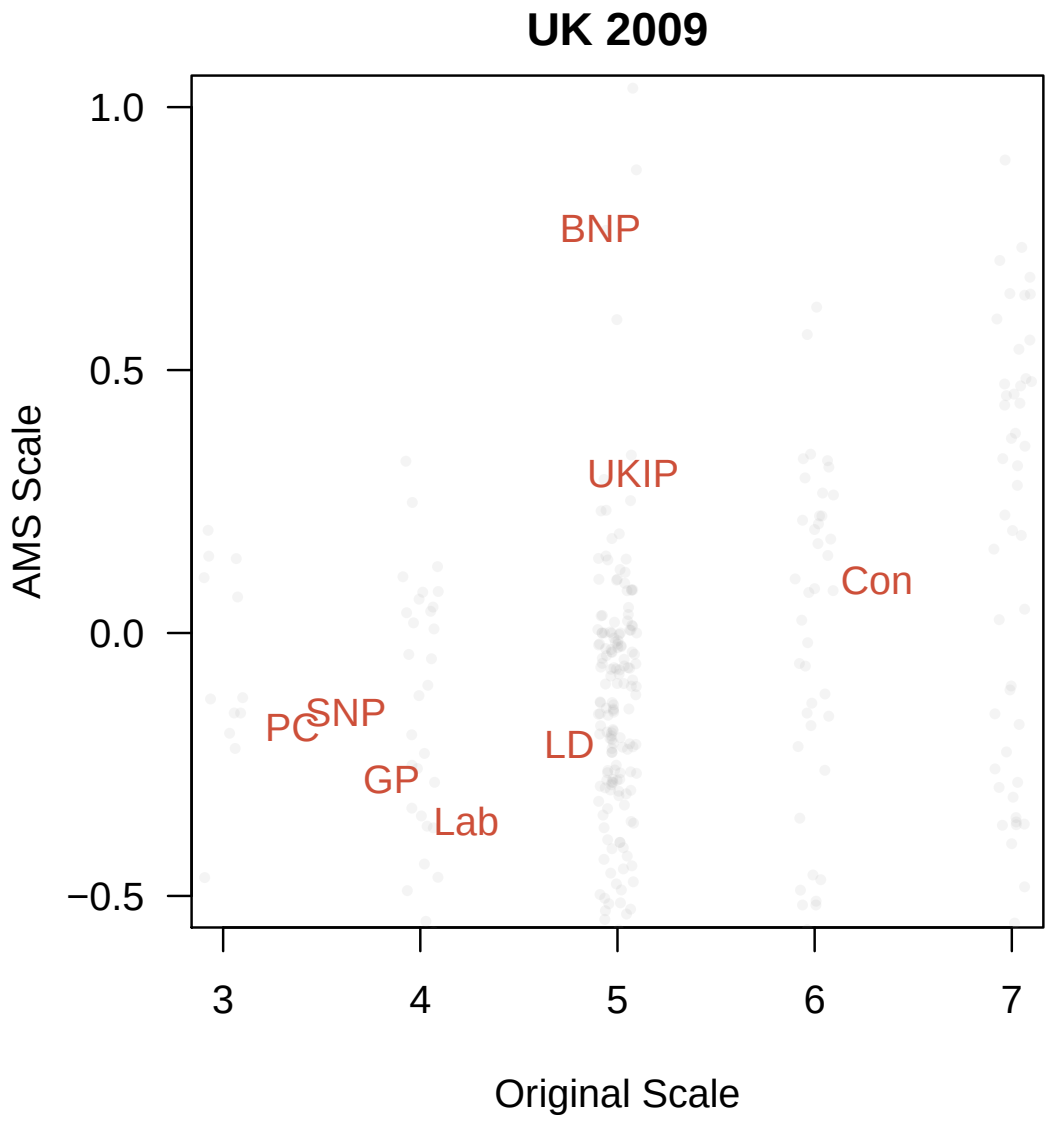
Anchoring items model

$$
\begin{aligned}
y_{ro} &\sim \text{Normal}(\hat{y}_{ro}, \sigma) \\
\hat{y}_{ro} &= \alpha_r + \beta_r \theta_o
\end{aligned}
$$

Scaling self-placements

$$
\zeta_r = \frac{s_r - \alpha_r}{\beta_r}
$$

CA: EES VS L/R Batteries

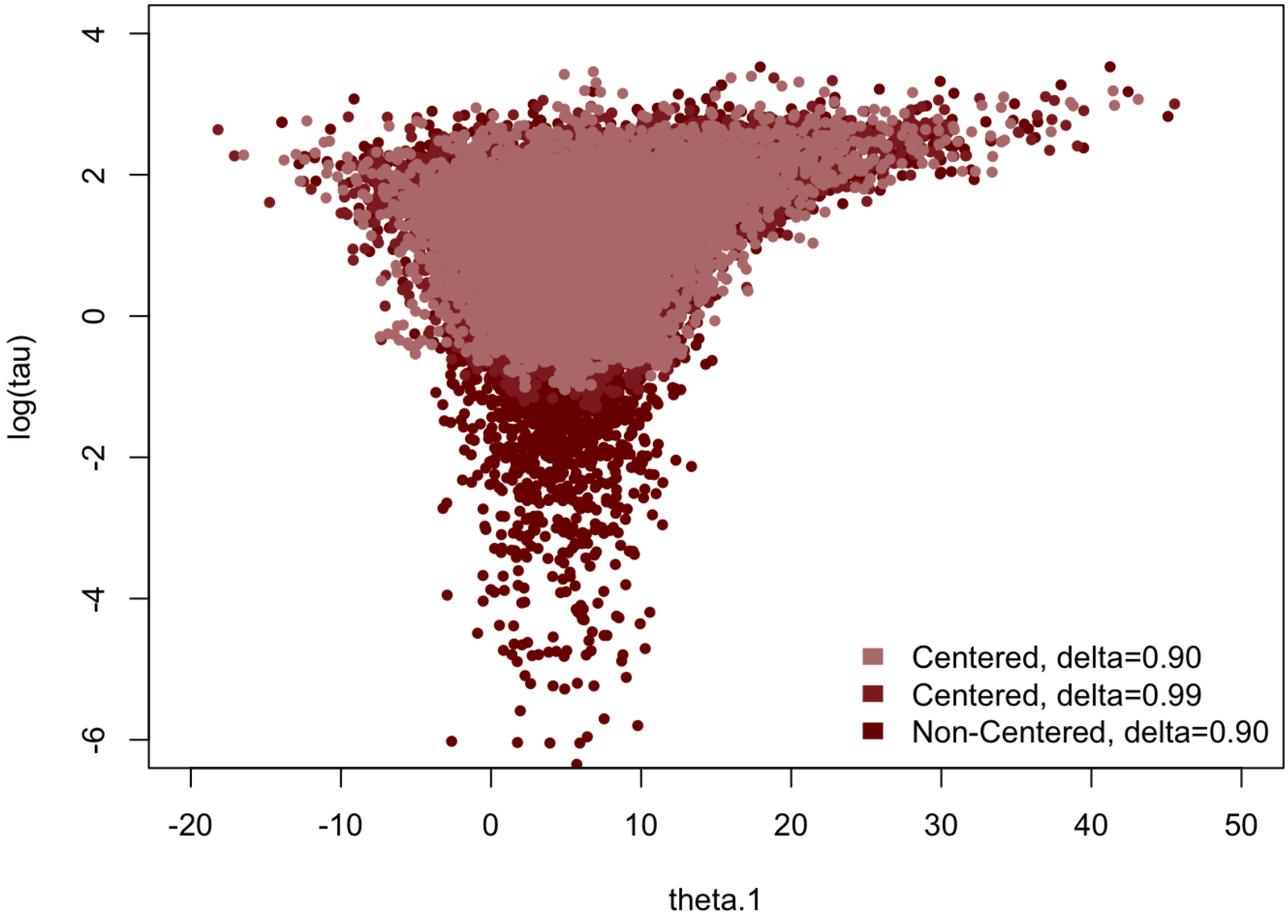


44% respondents have flipped scales

Bayesian AMS (Hare at al. 2015)

$$
\begin{aligned}
y_{ro} &\sim \text{Normal}(\alpha_r + \beta_r \theta_o,\ \sigma_r \sigma_o) \\
\sigma_o^{-1/2} &\sim \text{Gamma}(0.1,\ 0.1) \\
\sigma_r^{-1/2} &\sim \text{Gamma}(\gamma_1,\ \gamma_2) \\
\boldsymbol{\gamma} &\sim \text{Gamma}(0.1,\ 0.1) \\
\boldsymbol{\theta} &\sim \text{Normal}(0,\ 1) \\
\boldsymbol{\alpha} &\sim \text{Uniform}(-100,\ 100) \\
\boldsymbol{\beta} &\sim \text{Uniform}(-100,\ 100)
\end{aligned}
$$

HMC diagnostics: divergent transitions



Source: https://mc-stan.org/users/documentation/case-studies/divergences_and_bias.html

Fitting complex latent variable models

Peter Fischli and David Weiss's The First Blush of Morning, 1984.
Source: https://www.wmagazine.com/story/peter-fischli-david-weiss-merry-pranksters

A Mixture Framework

$$O = \pi M + (1 - \pi)C$$

- O observed responses
- M *informative* responses
- C *uninformative* responses
- π mixing weight, $\pi \in [0, 1]$

Measurement Model

Ordinal IRT

$$
\begin{aligned}
y_{ro} &\sim \text{Categorical}(\mathbf{p}_{ro}) \\
\mathbf{p}_{ro} &= f\left(\boldsymbol{\tau}_r, \boldsymbol{\theta}_o, \dots\right)
\end{aligned}
$$

Heterogeneity

M by respondent

- Scale flipping

C by respondent

- Straightlining
- Answering a different question

C by answer

- Midpoint as “Don’t know”/“Won’t say”
- Pseudoguessing

Measurement model 1.

$$
\begin{aligned}
y_{ro} &\sim \text{Categorical}\left(\boldsymbol{p}_{ro}\right) \\
p_{rok} &= \text{OrdLogit}\left(\tau_{rk} - \theta_o\right), \ \tau_{rk} < \tau_{r,k+1} \\
\boldsymbol{\tau_r} &\sim \text{Logistic}(0,\ 1) \\
\theta_o &\sim \text{Normal}(0,\ \sigma) \\
\zeta_r &\sim \text{Normal}(0,\ \sqrt{R}\sigma) \\
\sigma &\sim \text{HalfNormal}^{+}(0,\ 1)
\end{aligned}
$$

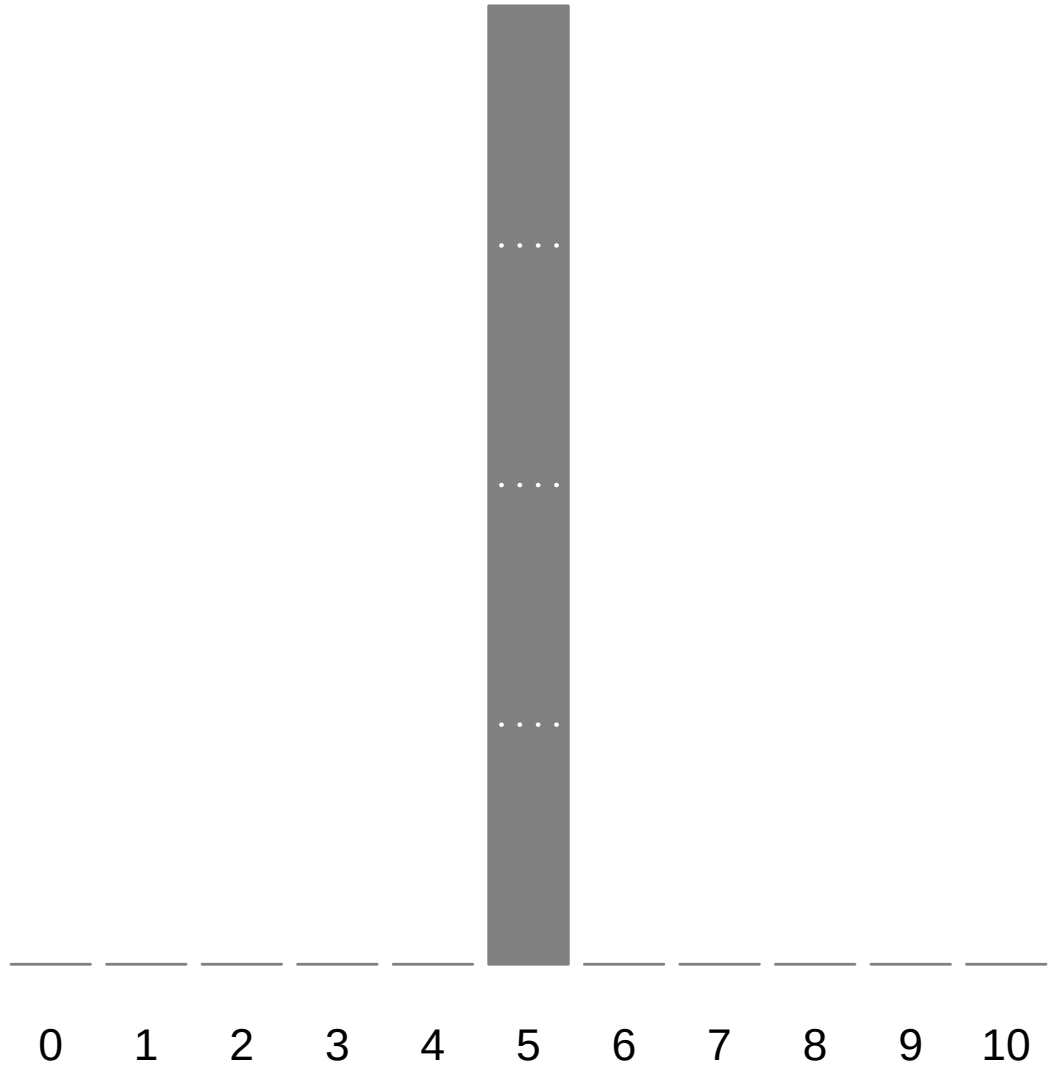
Measurement model 2.

$$
\begin{aligned}
p_{rok} &= \text{OrdLogit}\left(\beta_o \left(\tau_{rk} - \theta_o\right)\right) \\
\ln \beta_o &\sim \text{Normal}(0,\ 1)
\end{aligned}
$$

Respondent scale flipping

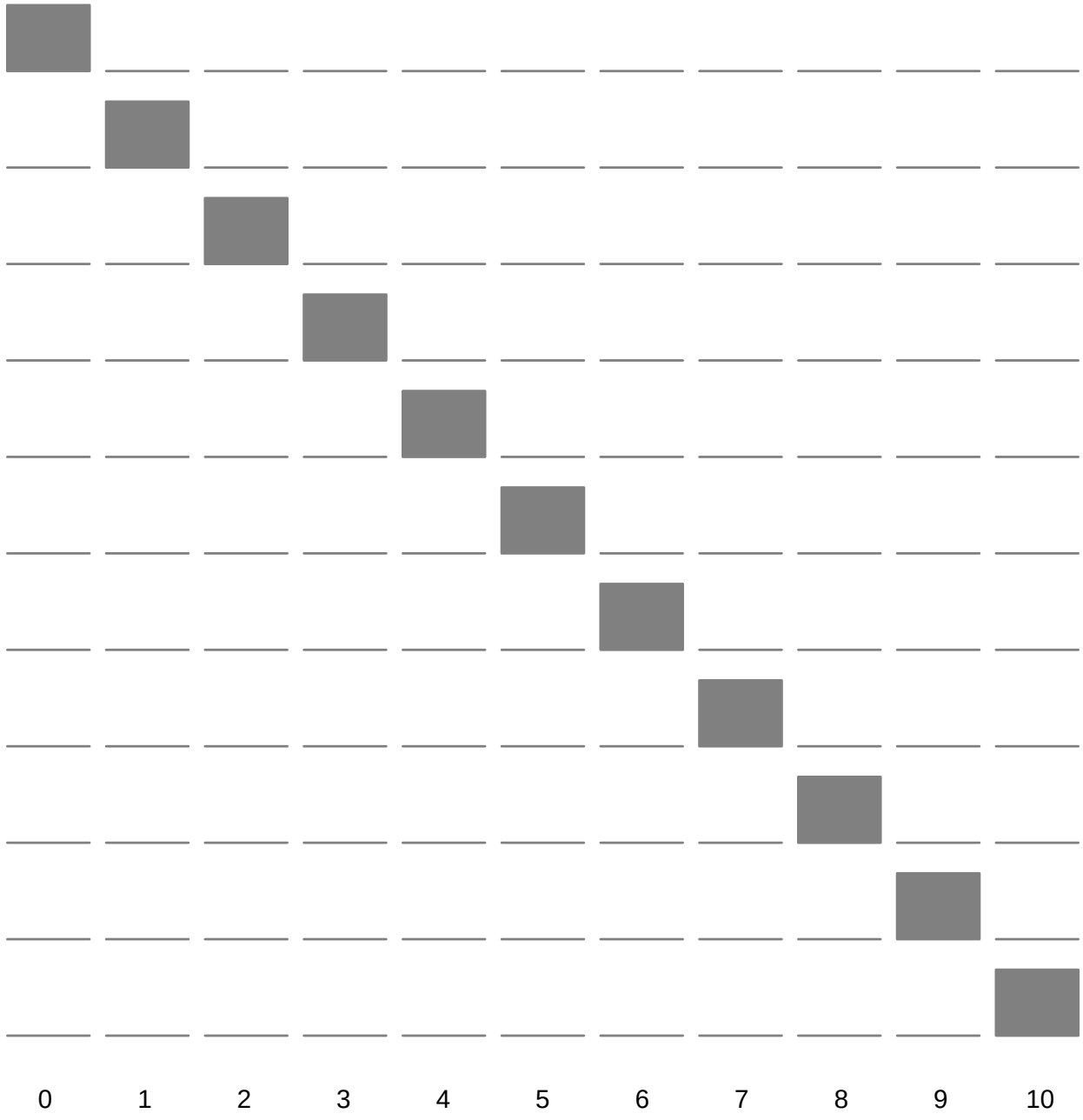
$$
\begin{aligned}
p_{rok} &= \text{OrdLogit}\left(\tau_{rk} - \gamma_r \theta_o\right) \\
\gamma_r &\in \{-1, +1\} \\
P(\gamma_r = -1) &= \pi^{\gamma} \\
\pi^{\gamma} &\sim \text{Beta}(0.5, 0.5)
\end{aligned}
$$

The Spike: Midpoint Inflation



The Slab: Pseudoguessing

The Spikes: Straightlining



Contamination

By respondent

$$g_r \sim \text{Dirichlet} - \text{Categorical}\,(0.5, \ldots, 0.5)$$

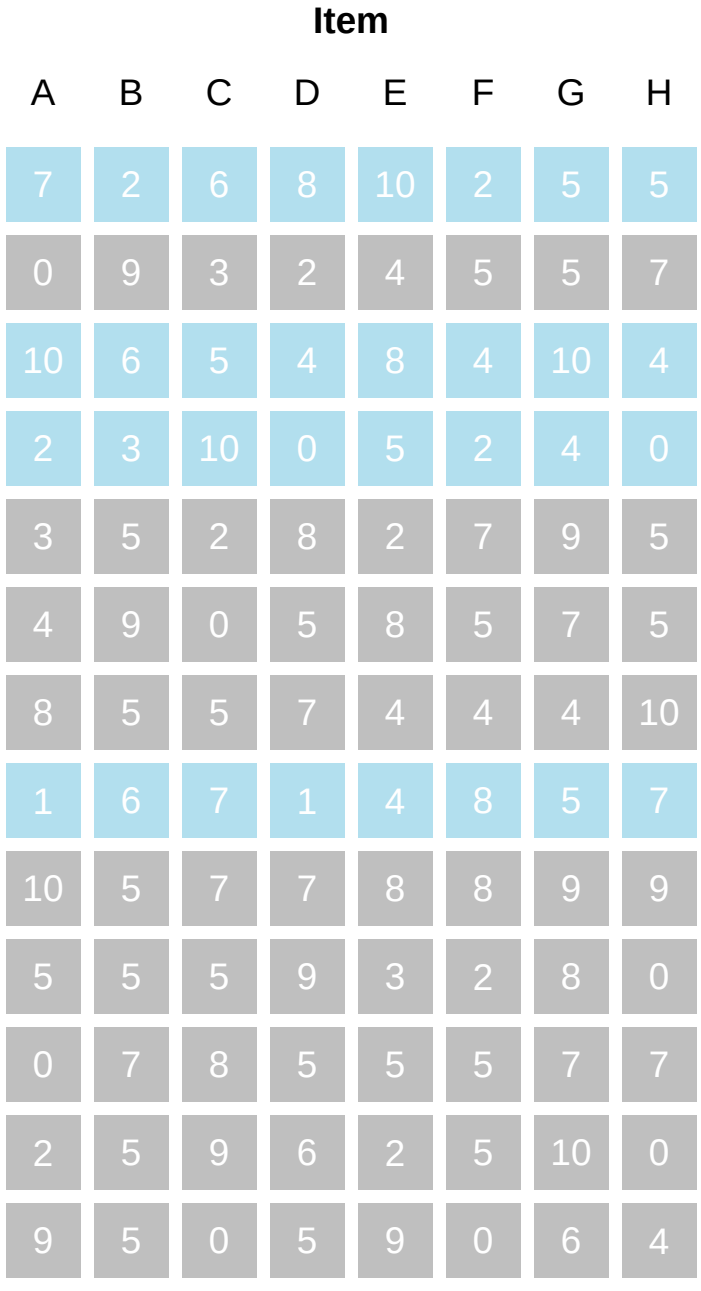
By answer

$$g_{ro} \sim \text{Dirichlet} - \text{Categorical}\,(0.5, \ldots, 0.5)$$

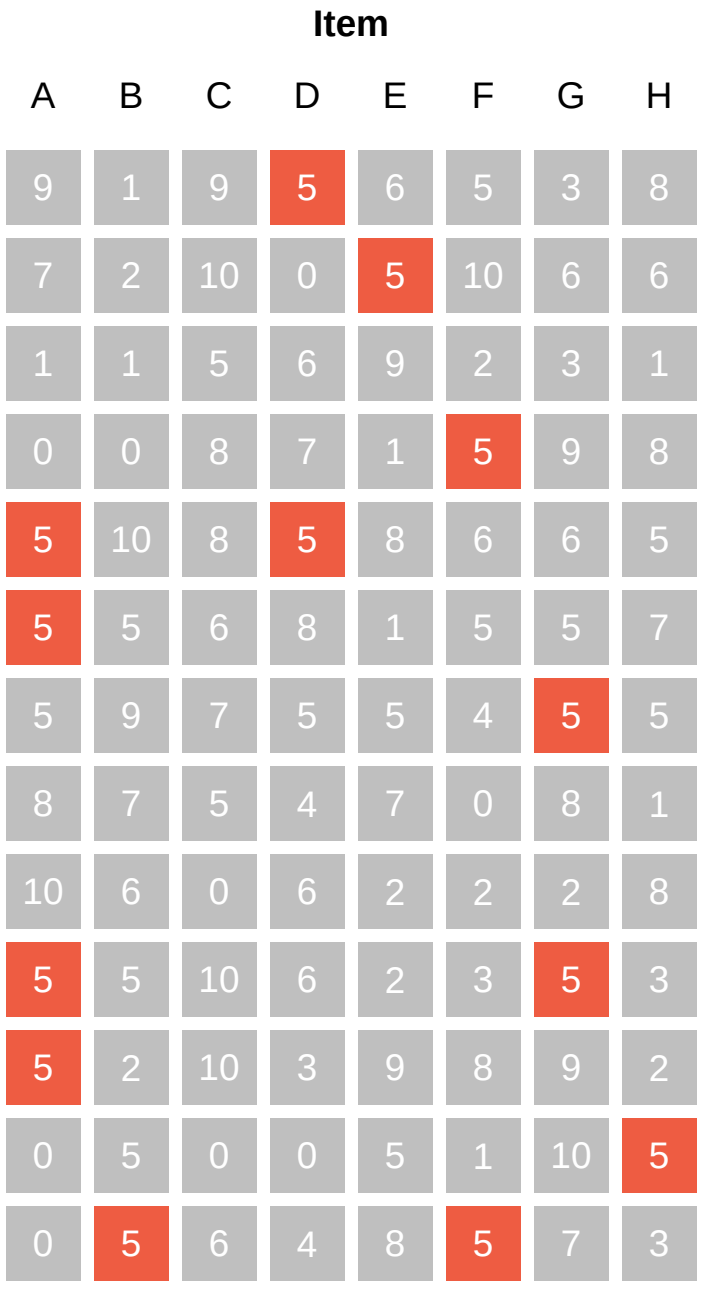
Baseline: No Flipping or Contamination

Item							
A	B	C	D	E	F	G	H
5	4	8	6	6	5	5	2
5	5	7	5	5	10	10	8
1	5	2	6	8	1	5	5
5	4	5	7	5	2	4	5
9	5	5	9	7	5	5	5
7	5	5	4	7	0	8	1
0	6	2	2	2	8	5	8
2	3	5	0	3	0	10	6
2	7	5	9	5	5	10	6
5	4	2	10	3	9	8	9
0	5	4	1	10	5	0	0
6	7	3	7	0	5	3	0
10	0	5	9	6	4	8	5

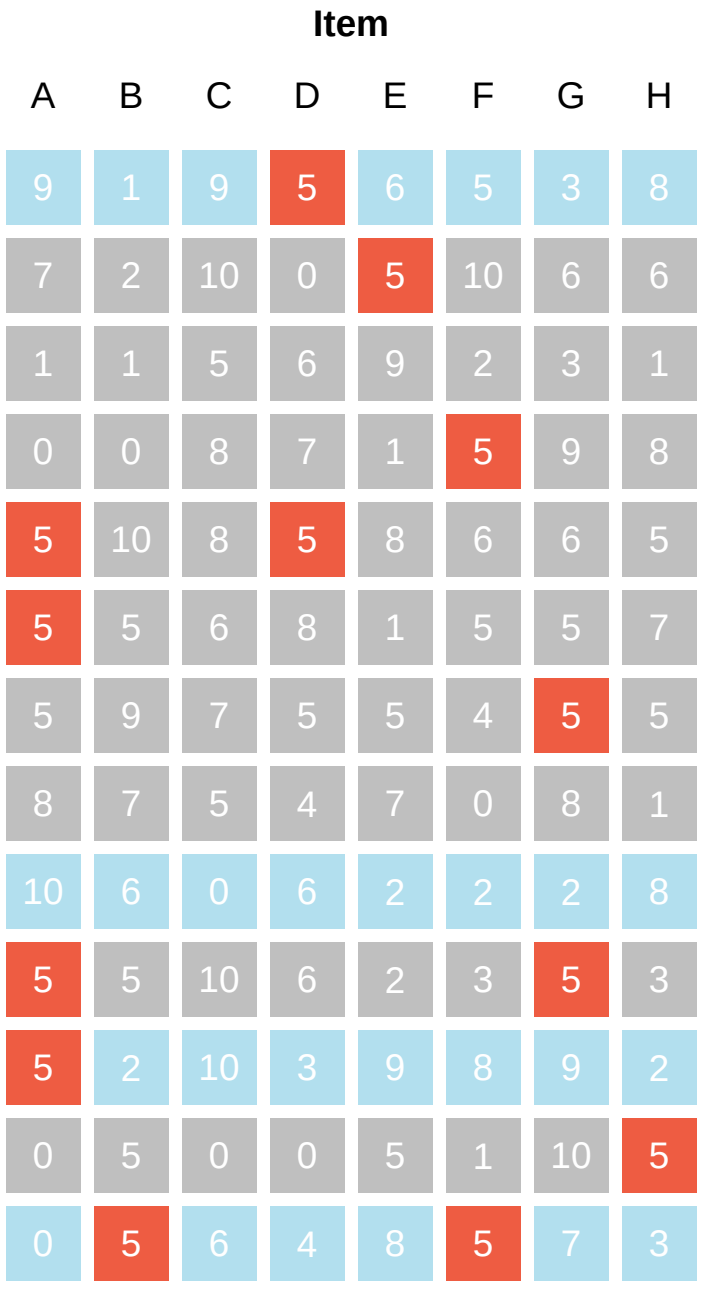
Flipping


Item
A B C D E F G H
7 2 6 8 10 2 5 5
0 9 3 2 4 5 5 7
10 6 5 4 8 4 10 4
2 3 10 0 5 2 4 0
3 5 2 8 2 7 9 5
4 9 0 5 8 5 7 5
8 5 5 7 4 4 4 10
1 6 7 1 4 8 5 7
10 5 7 7 8 8 9 9
5 5 5 9 3 2 8 0
0 7 8 5 5 5 7 7
2 5 9 6 2 5 10 0
9 5 0 5 9 0 6 4

Contamination: Midpoint Inflation



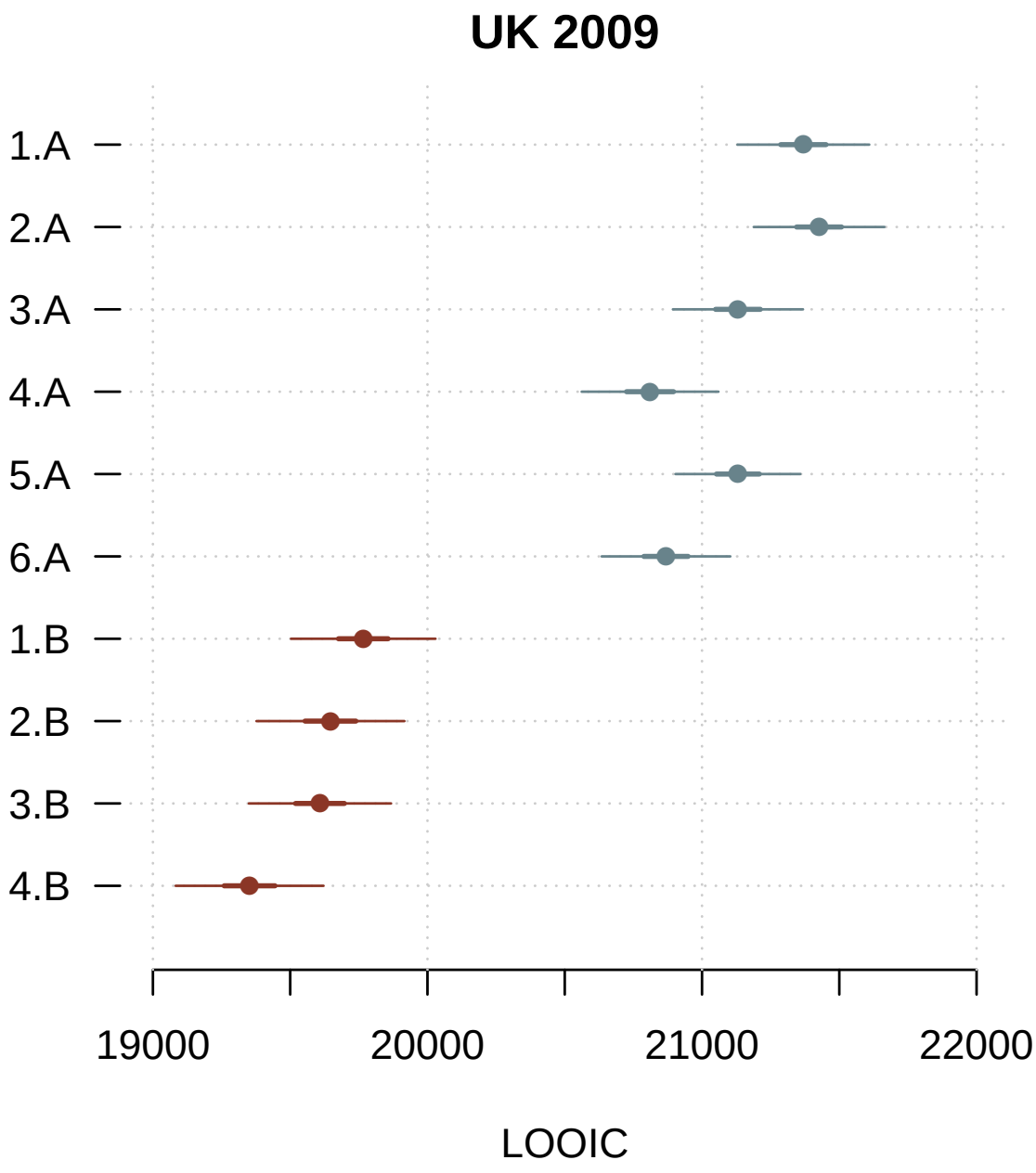
Flipping and Midpoint Inflation



Three EES VS Left-Right Batteries: Models

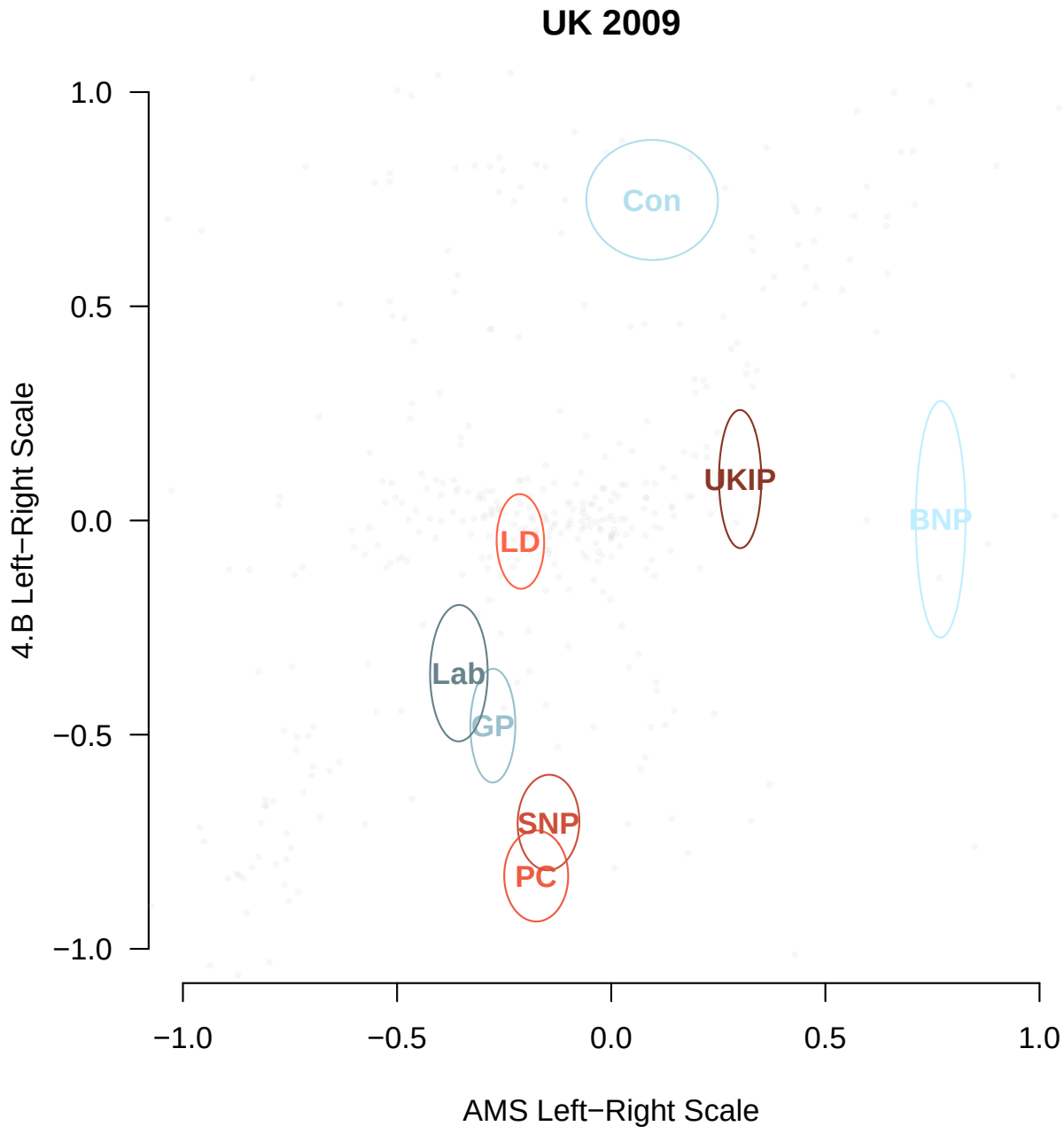
	A	B
1	respondent cut-points	+ item slopes
2	—"— + flipping	+ —"—
3	—"— + midpoint inflation	+ —"—
4	—"— rate by item	+ —"—

Three EES VS Left-Right Batteries: Model Fit

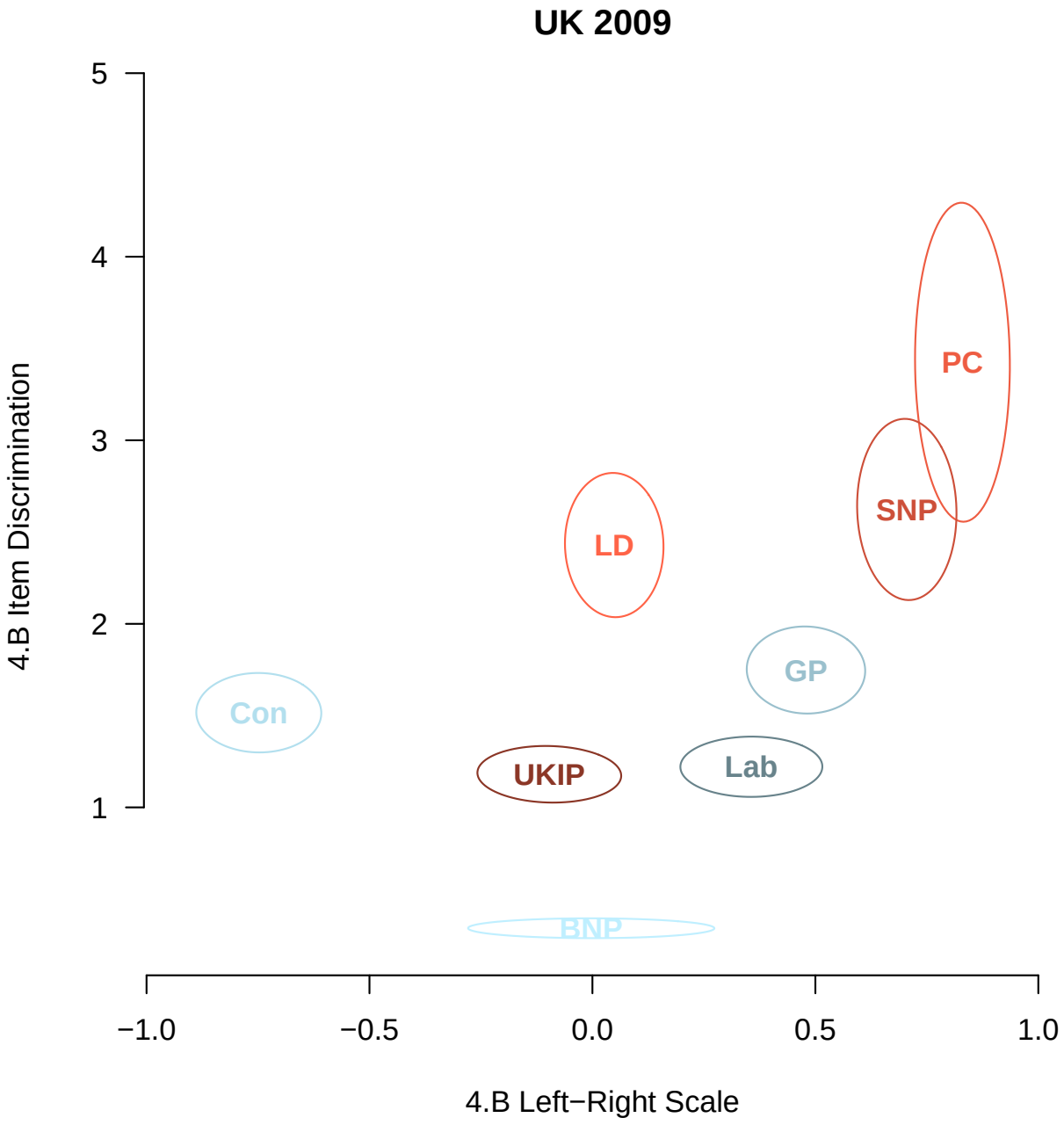


Three EES VS Left-Right Batteries: Mid-Point Inflation

AMS vs. 4.B


UK 2009
Con
UKIP
BNP
LD
Lab
GP
SNP
PC
1.0
0.5
0.0
−0.5
−1.0
4.B Left−Right Scale
−1.0
−0.5
0.0
0.5
1.0
AMS Left−Right Scale

4.B Item Slopes



Conclusion

No free lunch

- bias-variance trade-off
- estimating *well* costs

What didn't work so well deserves a paper

- GP cut-points
- respondent & item slopes
- Dirichlet-Categorical for ***y***

Next

- missing answers
- more structured memberships
- implications for party ambiguity

Thank you!

@medzihorsky