Better Anchoring and Ambiguity Measurement with Mixture Models

Juraj Medzihorsky

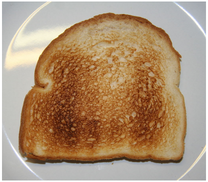
31 March 2020

The Problem

How crunchy do you like your toast?

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

Anchoring items



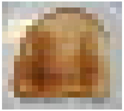
Source: Wikimedia Commons

How crunchy is this toast?

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

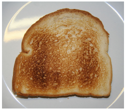
Anchoring batteries

Source: Wikimedia Commons



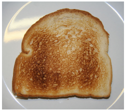
How crunchy is this toast?

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



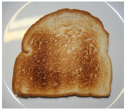
How [redacted] is this toast?

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



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**

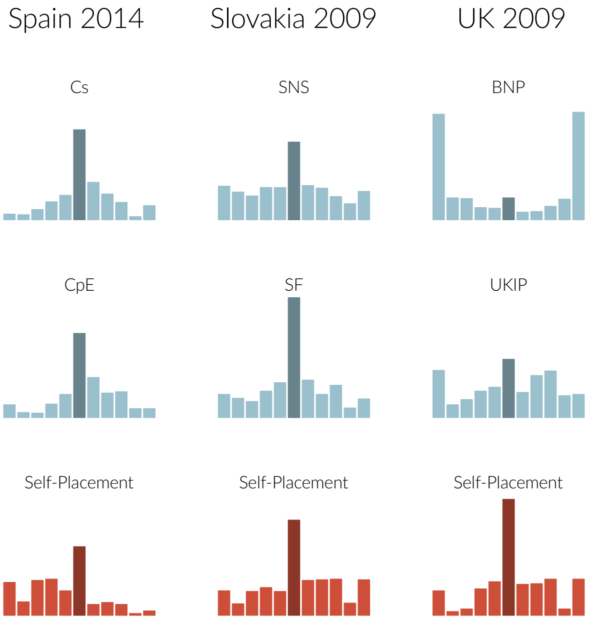


How crunchy is this toast?

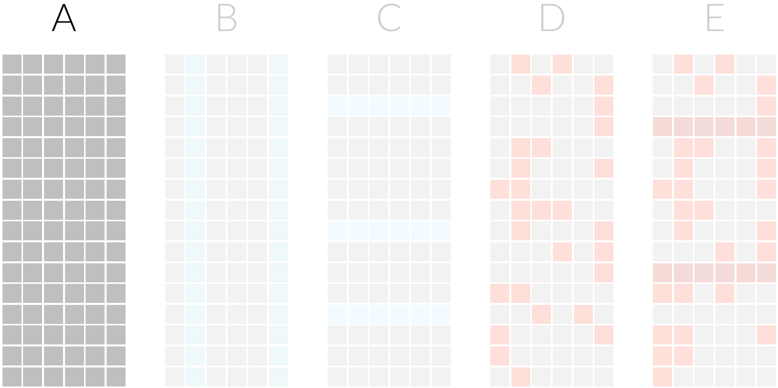
1. Very crunchy
2. Crunchy
3. Neither crunchy nor soft
4. Soft
5. Very soft

Source: telegraph.co.uk

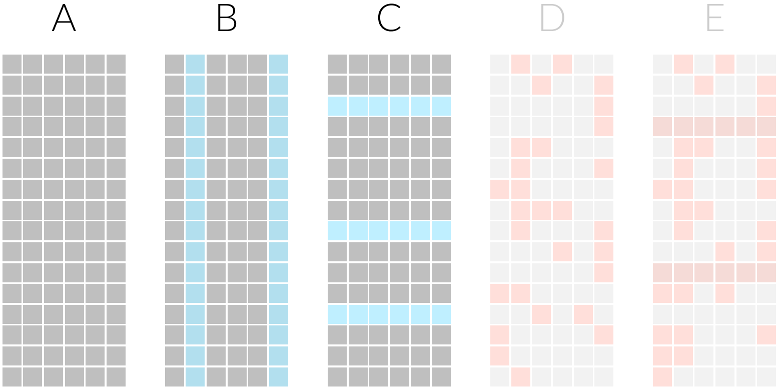
EES Left-Right batteries



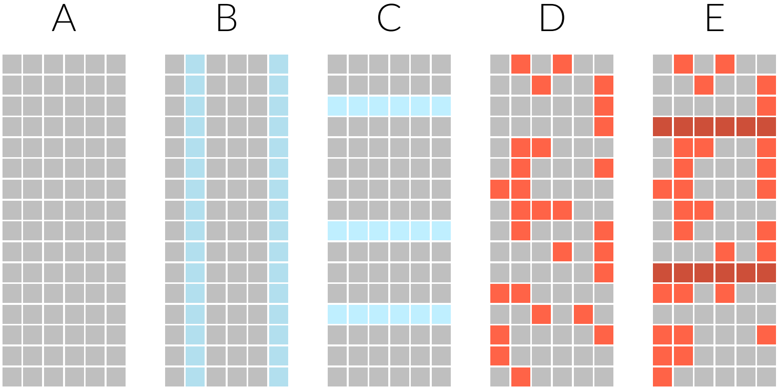
Low-quality responses



Low-quality responses



Low-quality responses



A Solution

A mixture framework

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

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

Modeling Informative Responses

Aldrich-McKelvey Scaling

Classical AMS

$$y_{ro} \sim \text{Normal}(\alpha_r + \beta_r \theta_o,\ \sigma)$$

Bayesian AMS (Hare at al. 2015)

$$y_{ro} \sim \text{Normal}(\alpha_r + \beta_r \theta_o,\ \sigma_r \sigma_o)$$

Scaling self-placements

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

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

Measurement model

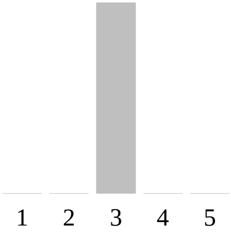
$$
\begin{aligned}
y_{ro} &\sim \text{Categorical}\left(\boldsymbol{p}_{ro}\right) \\
p_{rok} &= \text{OrdLogit}\left(\left(\tau_{rk} - \gamma_r \theta_o\right) \beta_o\right) \\
\boldsymbol{\tau_r} &\sim \text{Logistic}(0,\ 1),\ \tau_{rk} < \tau_{r,k+1} \\
P(\gamma_r = -1) &\sim \text{Beta}(0.5,\ 0.5),\ \gamma_r \in \{-1,\ +1\} \\
\ln \beta_o &\sim \text{Normal}(0,\ 1) \\
\theta_o &\sim \text{Normal}(0,\ \sigma) \\
\zeta_r &\sim \text{Normal}(0,\ \sqrt{R}\sigma) \\
\sigma &\sim \text{HalfNormal}^{+}(0,\ 1)
\end{aligned}
$$

Modeling Uninformative Responses

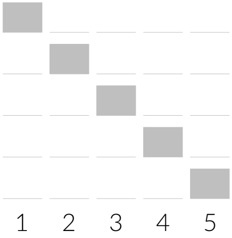
Contamination models

i. Pseudoguessing

ii. Midpoint



iii. Straightlining



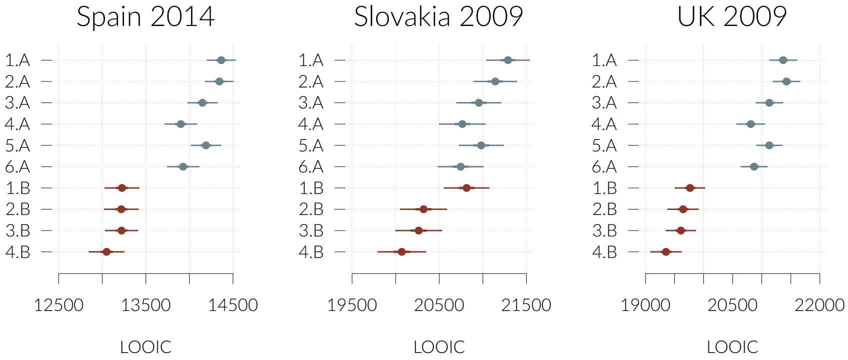
iv. Between-respondent multidimensionality

Building Mixture Models

10 mixture models

	Scale flipping	Midpoint inflation	Pseudo-guessing	Item slope
1.A				
2.A	Y			
3.A	Y	battery		
4.A	Y	item		
5.A	Y	battery	item	
6.A	Y	item	item	
1.B				Y
2.B	Y			Y
3.B	Y	battery		Y
4.B	Y	item		Y

EES: Model fit

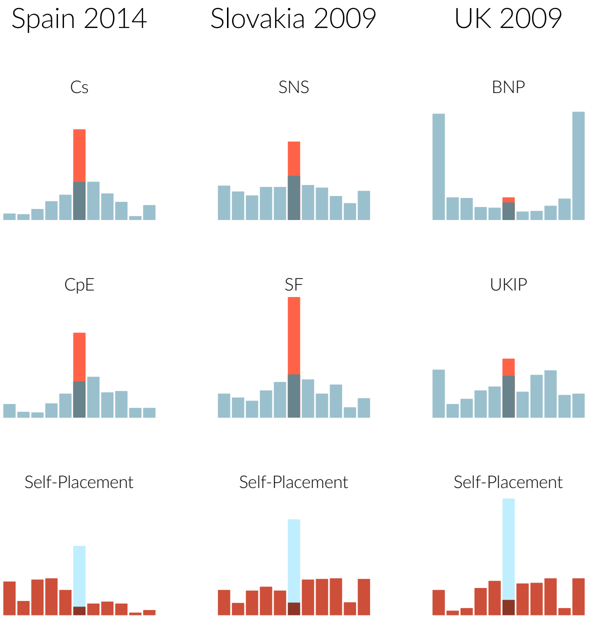


EES: Best fit

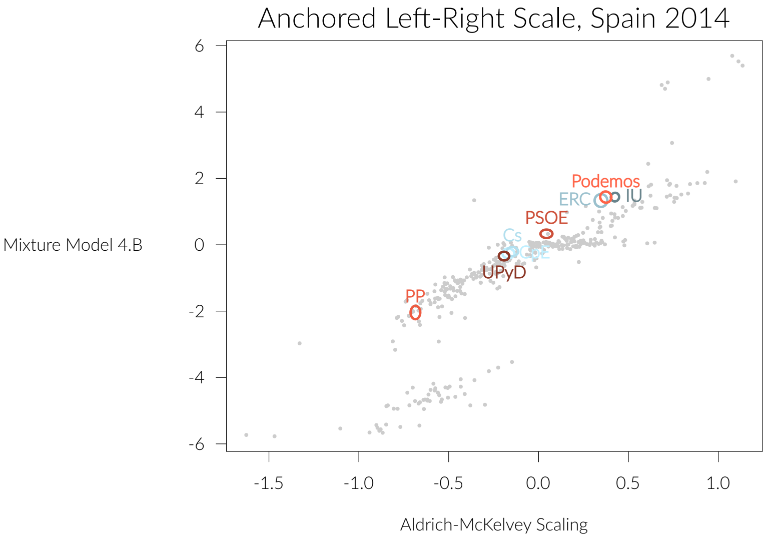
Model 4.B

- Respondent thresholds
- Scale flipping by respondent, population rate
- Midpoint inflation by response, item rate
- Item discrimination

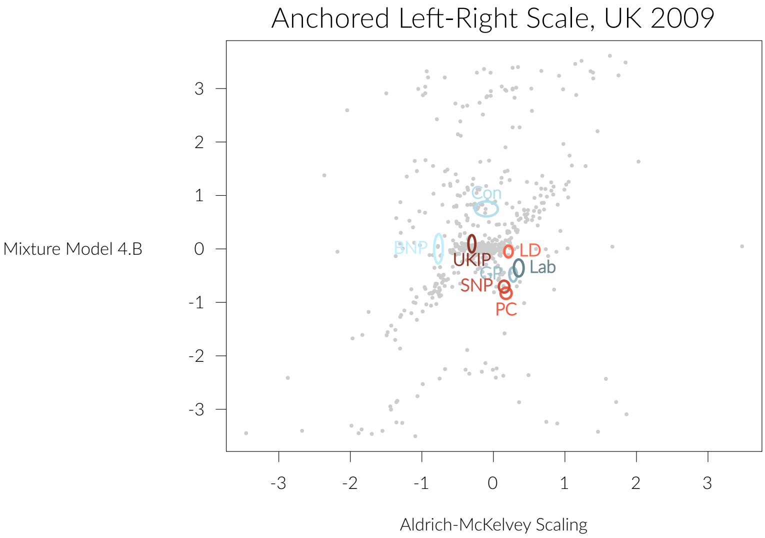
EES: Midpoint inflation estimated by model 4.B



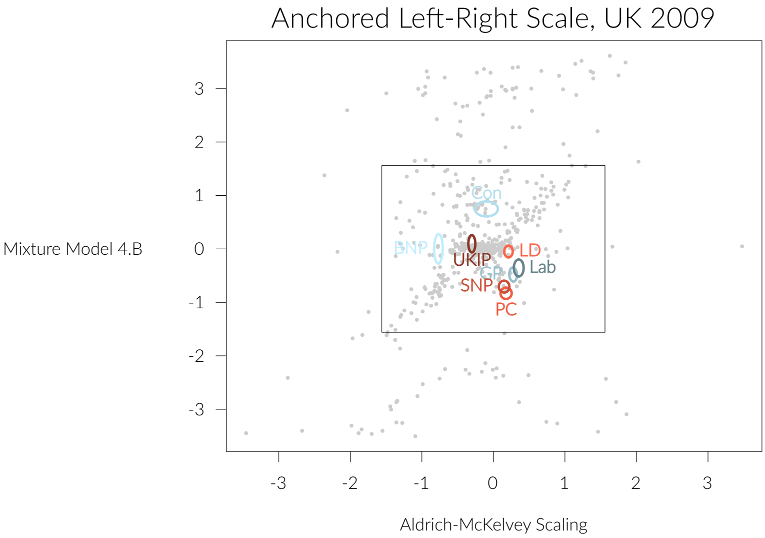
Model-based anchoring



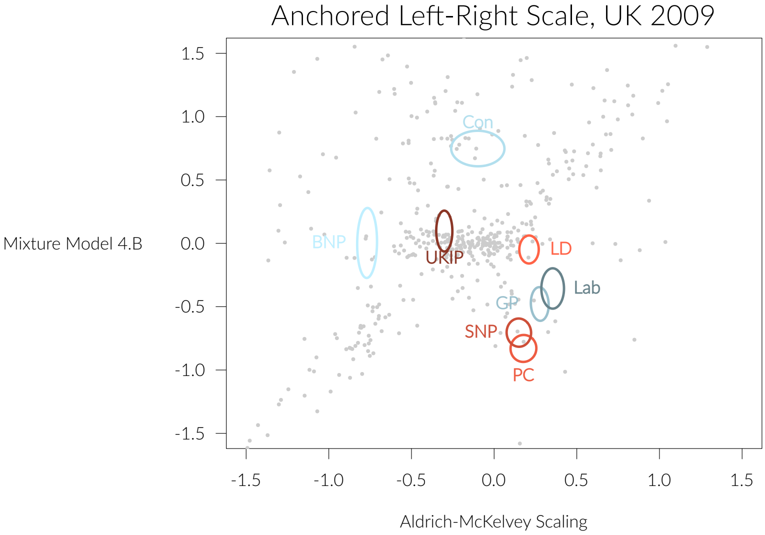
Model-based anchoring



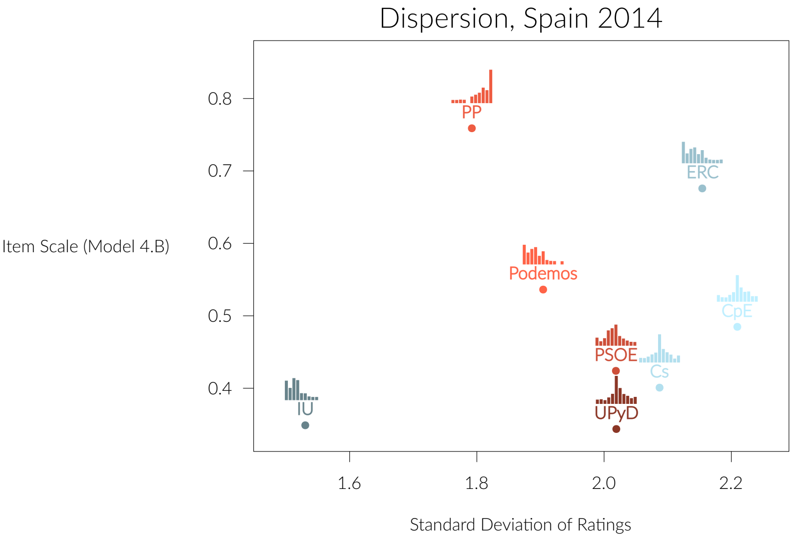
Model-based anchoring



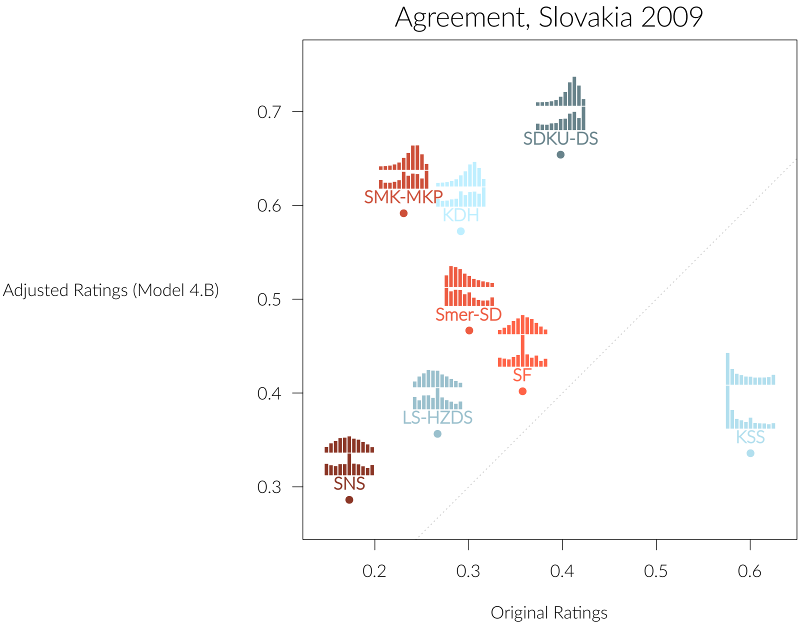
Model-based anchoring



Ambiguity and dispersion



Ambiguity and agreement



Conclusion

Conclusion

Developing the paper

- Two-step version
- Bayesian model validation
- Replicate causal analyses

Future research

- Modeling rationalization bias (Bølstad 2020)
- Nonresponse and ambiguity (Rozenas 2013)
- More structured component memberships

II. Research Agenda

Research agenda outline

0. Current ongoing work
1. Heterogeneous attitudinal structures
2. Minimum mixture estimation

0. Current ongoing work

- *R&R* with S. I. Lindberg & another invitation at *Cell: Patterns*
- Software implementing NOCNOC (DiDiD) with A. Glynn & N. Ichino, possibly *JSS*
- Grant application to RJ as the PI, on developing techniques for better anchoring, with D. Pemstein and K. L. Marquardt

1.1 Heterogeneous Attitudinal Structures

Theorizing heterogeneous shared mental maps

Spatial theories

- Social spaces & fields
- Ideological & policy spaces
- Mapping between them

Theorizing heterogeneous shared mental maps

- Belief systems and entropy (Martin 1999)
  - Constraint = Consensus + *Tightness*
  - Tightness = *Association*
- Association and *relational meaning* (Goldberg 2011, Boutyline & Vaisey 2017)
- Associations *diffuse* (Goldberg & Smith 2018)
- Diffusion by *imitation* (Henrich 2015)

Relational & Correlational Class Analysis:

- Pathbreaking

Relational & Correlational Class Analysis:

- Pathbreaking
- Simple mental maps
- Do not match the theorized DGP
- CCA's theorized model:

$$y_{ro} = \alpha_r + \beta_r \theta_{om[r]} + \epsilon_{ro}$$

- DIF-sensitive
- Uncertainty

Mixture IRT unfolding models:

$$
\begin{aligned}
y_{ro} &\sim f(\lambda_{ro}, \boldsymbol{\tau}_r, \ldots) \\
\lambda_{ro} &= \sum_{j=1}^{M} \mathrm{I}(j = m_r) d(\boldsymbol{\zeta}_r, \boldsymbol{\theta}_{oj}) \\
\boldsymbol{\tau}_r &\sim g(\mathbf{z}_r, \boldsymbol{\gamma}_1) \\
m_r &\sim h(\mathbf{x}_r, \boldsymbol{\gamma}_2) \\
\boldsymbol{\zeta}_r &\sim k(\mathbf{w}_r, \boldsymbol{\gamma}_3)
\end{aligned}
$$

Realignment and innovation

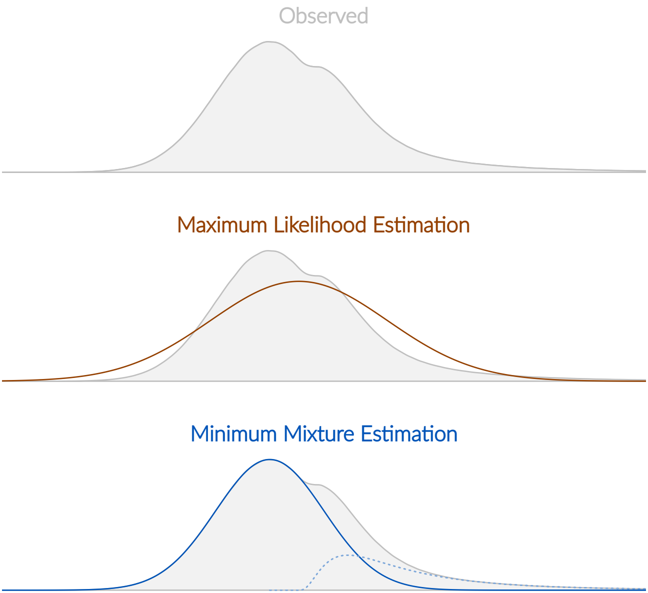
- Changing parties costs
- New entrants under polarization (Segatti)
- Ambiguity expansions & contractions
- Context and falling levels (Možný)
- *Dynamic* unfolding models
- *ItaNES* panel data (F. Vegetti)
- Supply-side

The international stage and attitudes

- With M. Popovic
- From *Culture in Power Transitions*
- Attitudes towards foreign countries
- Conventionally individual countries or pairs
- Scaling *GAP* and *TTS* thermometers
- National spaces and the UN
- Variation in time and over dimensions

1.2 Minimum Mixture Estimation

Minimum mixture estimation



Minimum mixture estimation

$$\pi^*(\mathrm{O}, \mathcal{M}) = \inf\{\pi : \quad \begin{aligned} &\mathrm{O} = (1-\pi)\mathrm{M} + \pi \mathrm{R}, \\ &\mathrm{M} \in \mathcal{M}, \\ &\mathrm{R} \text{ residual} \end{aligned} \quad \}$$

MME: Voter transitions and swing

Best

			Constant '66'	Constant '70	Swing '66'	Swing '70	Residual '66'	Residual '70
N	Kingston upon Hull, North	C	20	20			0	1
		L	29	29			1	0
P	Lancashire, Newton	C	18	18		3	0	0
		L	29	29	2		0	0
U	Hertfordshire, Hertford	C	25	25		3	0	1
		L	21	21	3		1	0

Worst

			Constant '66'	Constant '70	Swing '66'	Swing '70	Residual '66'	Residual '70
N	Merthyr Tydfil	C	8	8			3	0
		L	24	24			33	0
P	Merthyr Tydfil	C	7	7		1	4	0
		L	24	24	2		31	0
U	Merthyr Tydfil	C	6	6		2	5	0
		L	24	24	2		31	0

MME: Further applications

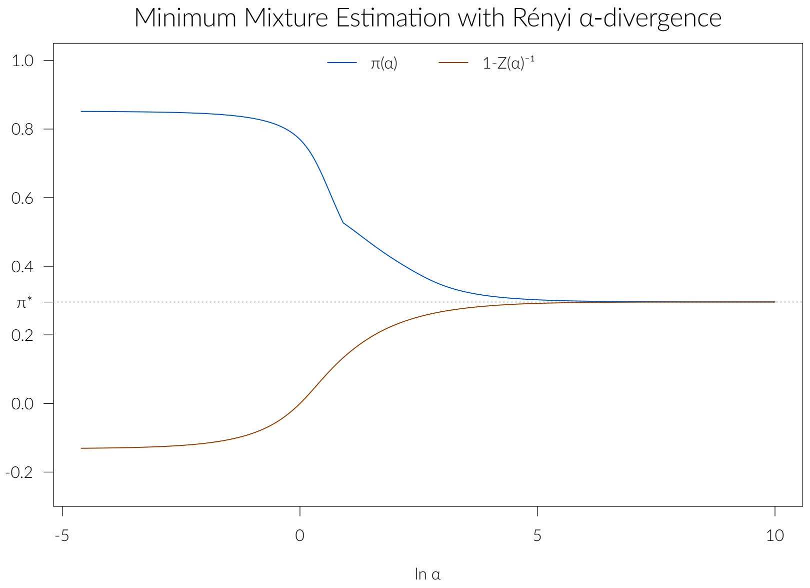
Roll call voting

- Cohesion under cross-cutting groups
- Interpreting latent spaces

Political text

- Identifying ideological or policy content
- Programmatic overlap and stability

New minimum mixture estimators



III. Teaching

Detectives, farmers, and causal questions

Activity:

- Write down a *causal* question that you find interesting. Try making it a social science question.

Detectives, farmers, and causal questions

Detective	Farmer
Who's done it?	*Will the fertilizer help?*
Responsibility	Intervention
What caused y_i?	Does X affect Y?
Backwards	Forwards
Causes of effects	Effects of causes
Most humans	R. A. Fisher

Detectives, farmers, and causal questions

Activity:

- Who had a detective question?

Detectives, farmers, and causal questions

Activity:

- Who had a detective question?
- Who had a farmer question?

Detectives, farmers, and causal questions

Activity:

- Who had a detective question?
- Who had a farmer question?
- Did we get more detective or farmer questions?

Detectives, farmers, and causal questions

Source: Wikimedia Commons

Detectives, farmers, and causal questions

Activity:

- Did anyone have a question that is neither a detective nor a farmer question?

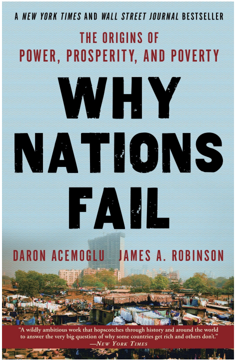
Detectives, farmers, and causal questions

Activity:

- Did anyone have a question that is neither a detective nor a farmer question?
- Is it really neither?
- What would you call it?
- What is it closer to?

Detectives, farmers, and causal questions

The Class Action Detective



"What causes all y_i?" are the most attractive but also the most difficult causal questions to answer.

Detectives, farmers, and causal questions

Activity:

- Let's take some of your detective questions and find some farmer questions that are related to them.

Homework 7:

- Prepare three causal questions related to your course paper.
- At least one of the questions *must* be an effects-of-causes question and at least one a causes-of-effects question.
- Briefly describe why each question is theoretically interesting in about 150 words per question, not counting the references.

Thank you!