

Aggregating Expert Ratings with IRT Models: The V-Dem Methodology

Juraj Medzihorsky

V-Dem Institute & Durham University

10 February 2022

Based on work by V-Dem's measurement and data teams

- Dan Pemstein,
- Kyle L. Marquardt
- Eitan Tzelgov
- Yi-ting Wang
- Joshua Krusell
- Farhad Miri
- Johannes von Röemer

The Data

Elections (Austria)

1. Read Question. 2. Click & drag to select years. 3. Apply or Edit specific dates, if desired. 4. Apply or type response. 5. Rate Confidence. 6. Submit. 7. Repeat for remaining years. 8. Click "Next".

Undo

(Regional government elected) :

At the regional level, are **government** offices elected in practice?

"Government offices" here refers to a regional executive and a regional assembly, not a judiciary and not minor bureaucrats. An executive is a single individual (or a very small group) (e.g., a governor). An assembly is a larger body of officials, who may be divided into two chambers. "Elected" refers to offices that are directly elected by citizens or indirectly elected by a regional elected assembly. All other methods of obtaining office – including appointment by higher or lower levels of government – are considered to be non-elected. In classifying a position as elected one is making no judgments about the freeness/fairness of the election or the relative extent of suffrage. One is simply indicating that there is an election and that the winner of that election (however conducted) generally takes office.

Min: 0 Max: 5

- ☐ (0) Generally, offices at the regional level are not elected.
- ☐ (1) Generally, the regional executive is elected but not the assembly.
- ☐ (2) Generally, the regional assembly is elected but not the executive.
- ☐ (3) Generally, the regional executive is elected and there is no assembly.
- ☐ (4) Generally, the regional assembly is elected and there is no executive.
- ☐ (5) Generally, the regional executive and assembly are elected.

Confidence: 0%

I have no idea at all. [Any scores accompanied by a confidence level of zero will be treated as missing data.]

Submit

Jump To Question :

Regional government elected

Previous

Next

Exit

	00	01	02	03	04	05	06	07	08	09
2005						2005	2006	2007	2008	2009
2010	2010	2011	2012	2013	2014	2015	2016	2017		

Click and drag to select range of years.

Specific Dates:

Add

Del

Edit

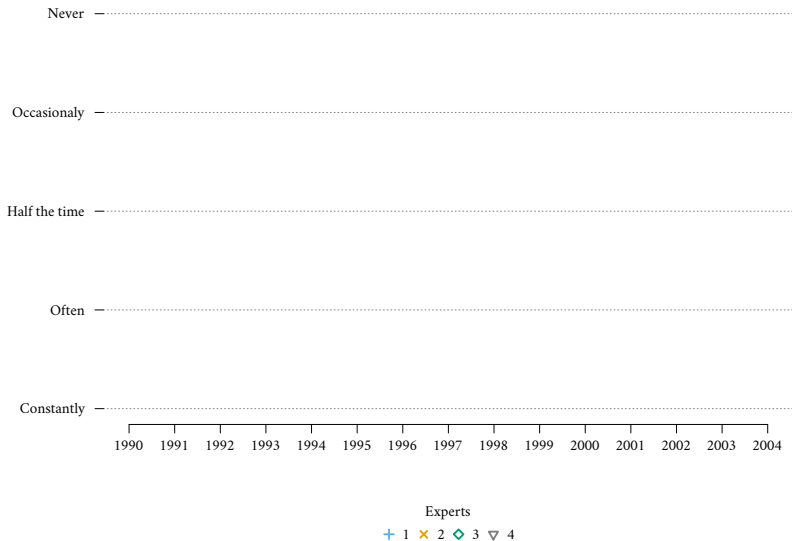
Legend

'▲': At least one date in this cell does not have an answer.

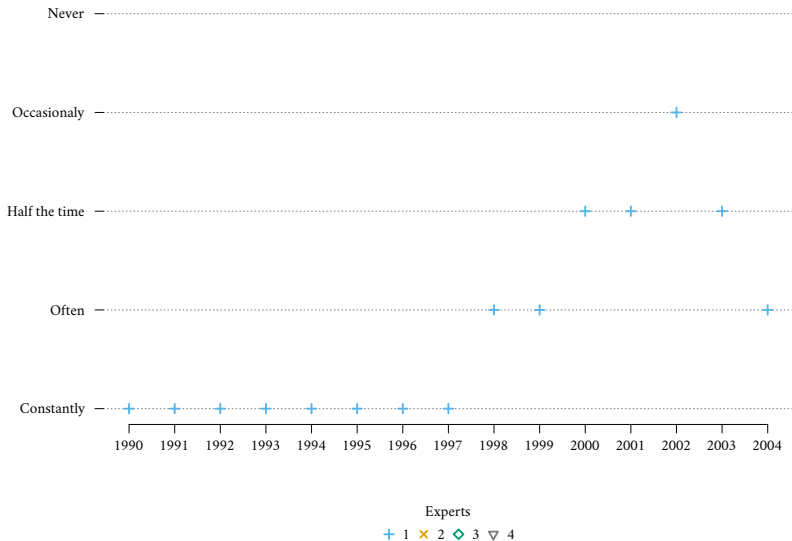
'⚠': All dates in this cell have an answer, but at least one does not have a confidence rating.

'✔': All dates in this cell have both an answer and a confidence rating.

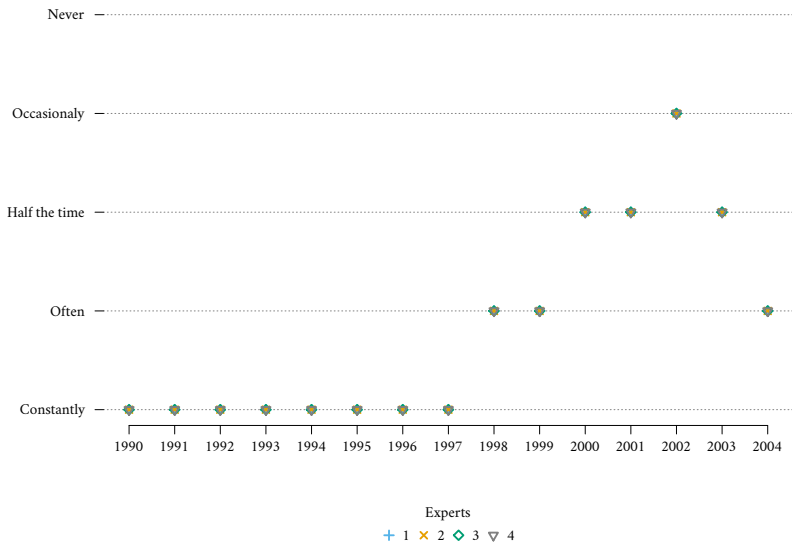
Executive Embezzlement in Džundža



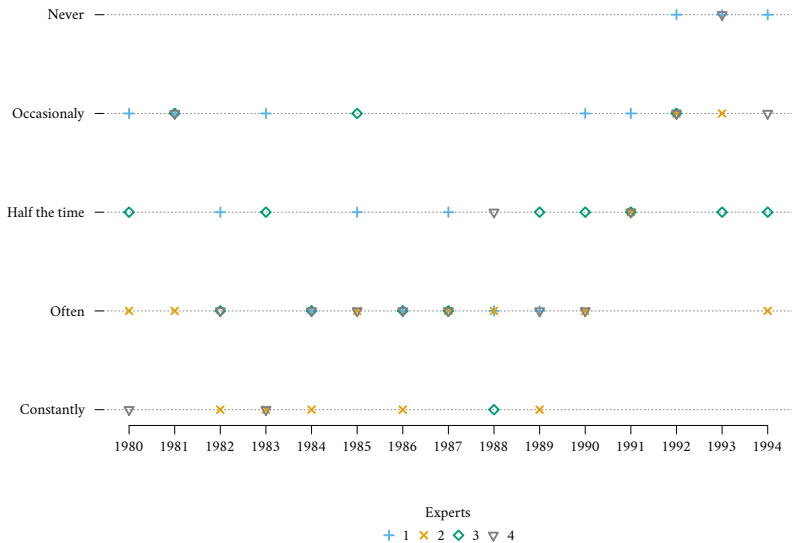
Executive Embezzlement in Džundža



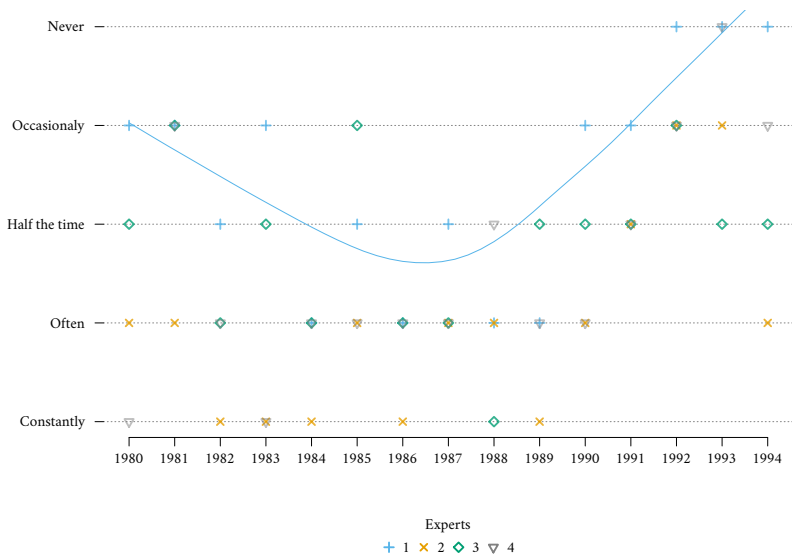
Executive Embezzlement in Džundža



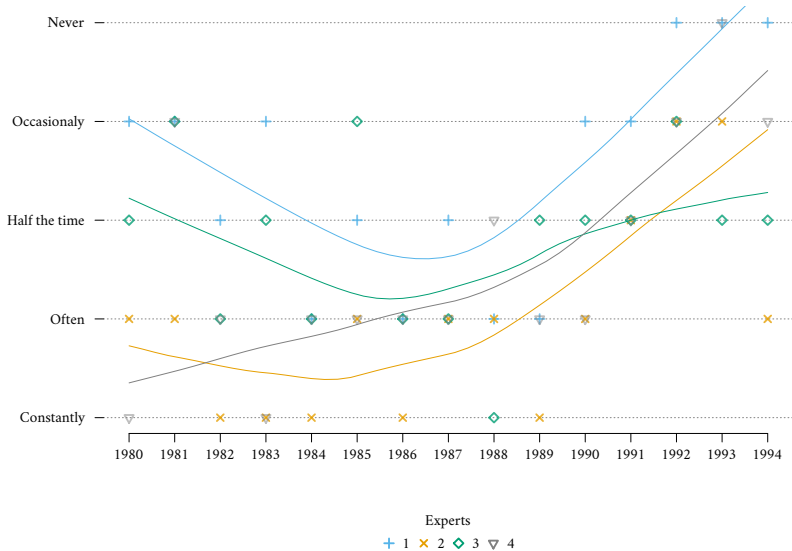
Executive Embezzlement in Molvemark



Executive Embezzlement in Molvemark



Executive Embezzlement in Molvamar



The Problem

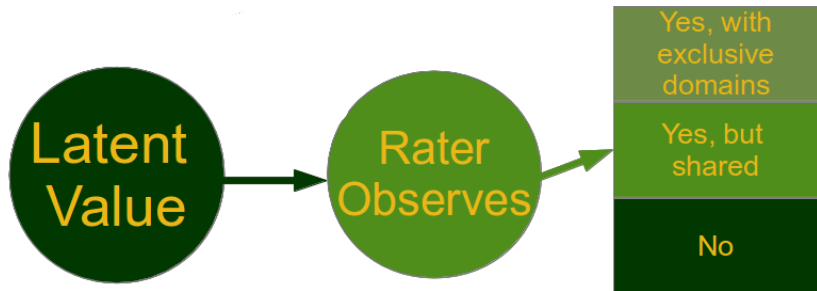
Rater
Observes

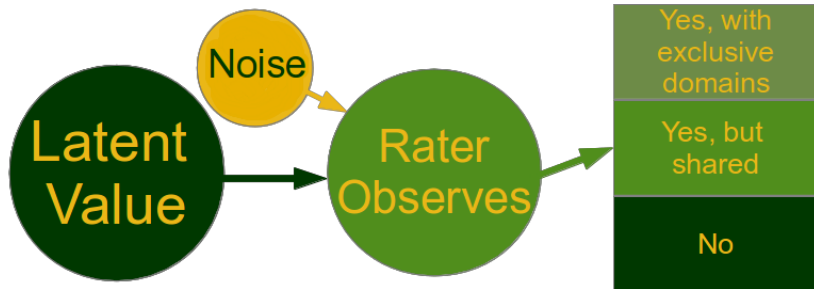


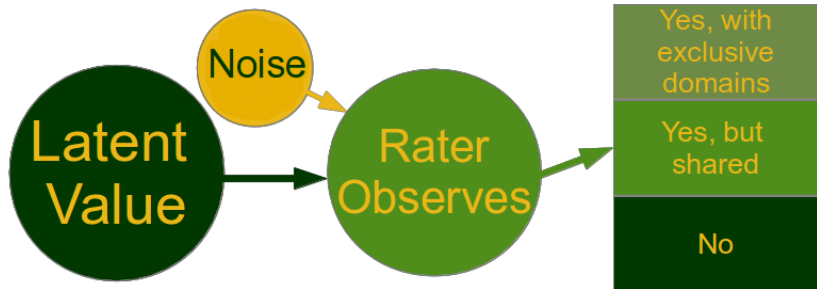
Yes, with
exclusive
domains

Yes, but
shared

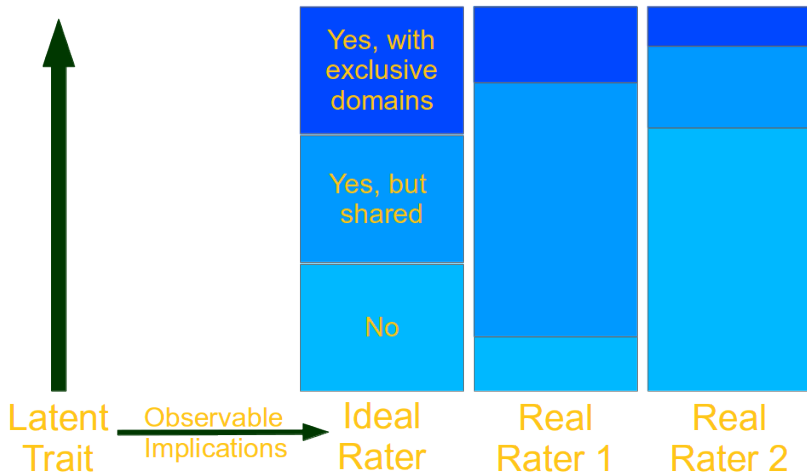
No







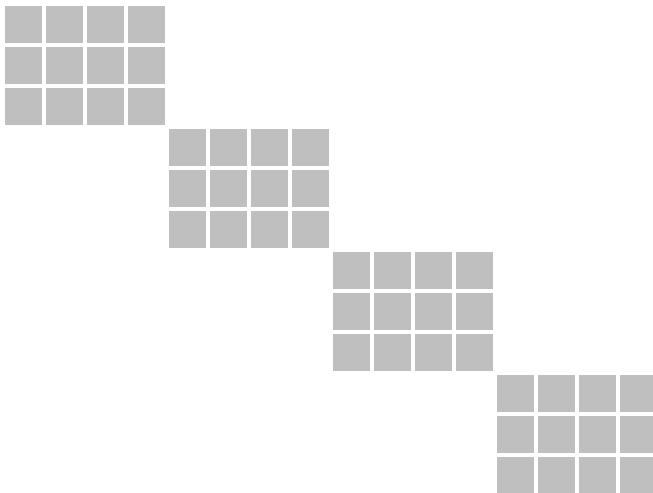
This is not all ...



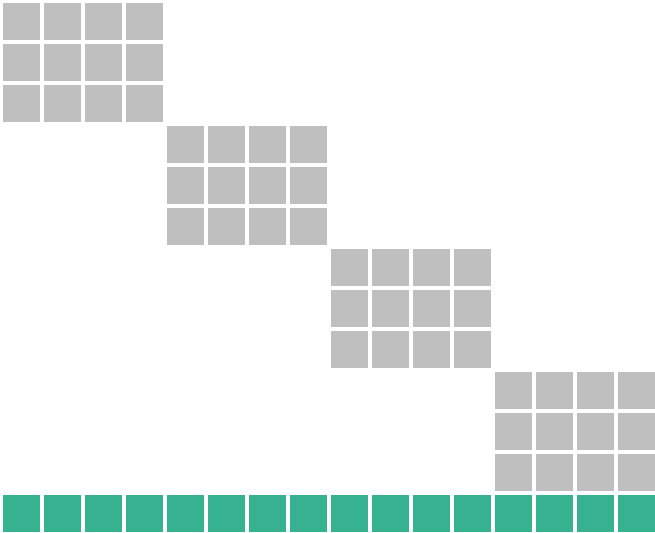
Survey Design

Expert

A B C D E F G H I J K L M N O P

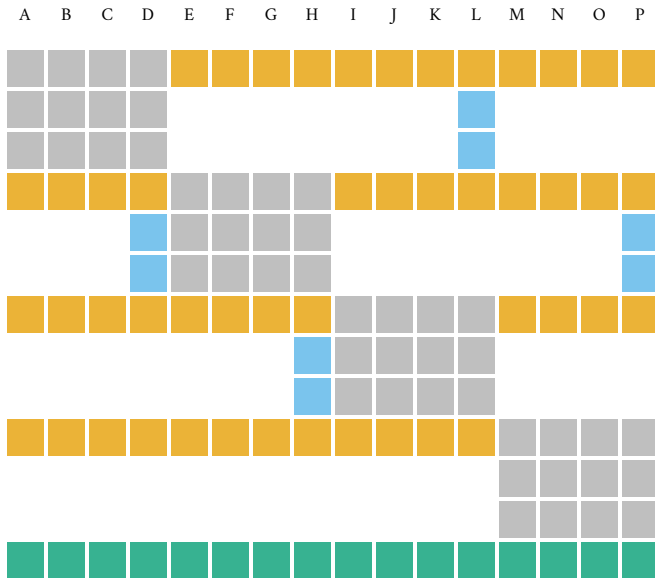


A B C D E F G H I J K L M N O P



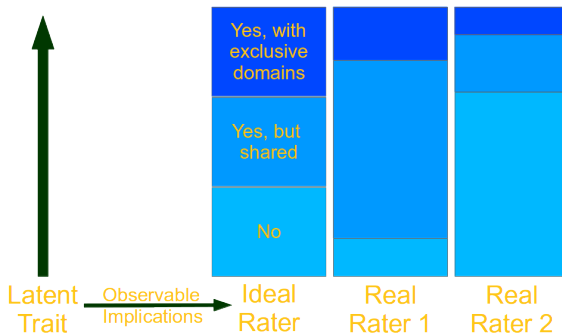
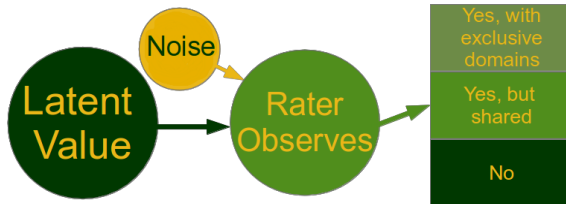
The diagram illustrates a 16-bit bus architecture with 16 columns labeled A through P. The grid shows various patterns of colored squares (gray, yellow, teal) representing different bus configurations. The bottom row is entirely teal. The top row is entirely yellow. The middle rows show complex patterns of gray and yellow squares.

Expert



The Model

Modeling The Data Generating Process



The V-Dem Measurement Model

$$\text{rating}[\text{object}, \text{expert}] = f(\text{truth}[\text{object}], \text{reliability}[\text{expert}], \text{thresholds}[\text{expert}])$$

The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \textit{Categorical}(\mathbf{p}_{ie}) \\ \mathbf{p}_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\ p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \boldsymbol{\tau}_e &= \{\tau_{e1} = -\infty, \dots, \tau_{eK-1}, \tau_{eK+1} = \infty\} \\ \tau_{ek} &\sim \textit{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\ \tau_{ck}^C &\sim \textit{Normal}(\tau_k^W, 0.25) \\ \tau_k^W &\sim \textit{Uniform}(-6, 6) \\ \beta_e &\sim \textit{Normal}^+(1, 1) \\ \zeta_i &\sim \textit{Normal}(z_i, 1)\end{aligned}$$



Articles

About 178,000 results (0.13 sec)

Any time

Since 2022

Since 2021

Since 2018


Custom range...

Sort by relevance

Sort by date

Any type

Review articles

☐ include patents☒ include citations Create alert**[BOOK] Item response theory**[SE Embretson, SP Reise - 2013 - taylorfrancis.com](#)

This book develops an intuitive understanding of IRT principles through the use of graphical displays and analogies to familiar psychological principles. It surveys contemporary IRT models, estimation methods, and computer programs. Polytomous IRT models are given central ...

[☆ Save](#) [🔖 Cite](#) [Cited by 6890](#) [Related articles](#) [All 8 versions](#) [🔗](#)**[BOOK] The basics of item response theory**[FB Baker - 2001 - ERIC](#)

... of **item response theory** would not be where it is today. ... on **item response theory**, which resulted in the first edition of this book in 1985. This suggestion allowed me to fulfill a long-standing desire to develop an instructional software package dealing with **item response theory** for ...

[☆ Save](#) [🔖 Cite](#) [Cited by 2490](#) [Related articles](#) [All 2 versions](#) [🔗](#)**Item response theory.**[F Drasgow, CL Hulin - 1990 - psycnet.apa.org](#)

review a variety of IRT [**item response theory**] models that can be used with items scored dichotomously or polychotomously/methods of estimation and appropriate computer programs are also described/emphasis... is placed on applications of IRT to important measurement ...

[☆ Save](#) [🔖 Cite](#) [Cited by 234](#) [Related articles](#)**[BOOK] Item response theory: Principles and applications**[RK Hambleton, H Swaminathan - 2013 - books.google.com](#)

... Despite these early research efforts, interest in **item response theory** lay dormant until the late 1960s and took a backseat to the emerging ... , " gave rise to a resurgence of interest in **item response theory**. Impetus for the development of **item response theory** as we now know it was ...

[☆ Save](#) [🔖 Cite](#) [Cited by 4131](#) [Related articles](#) [All 3 versions](#) [🔗](#)

The V-Dem Measurement Model

- Bayesian hierarchical IRT model
 - ▶ IRT: each expert gets their own thresholds τ and reliability β
 - ▶ Bayesian: full generative model with parameter distributions
 - ▶ Hierarchical: experts' thresholds are regularized towards their countries' thresholds which are in turn regularized towards the global thresholds
- Related to the Graded Response Model (Samejima)
- Alternative to Aldrich-McKelvey Scaling

The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \textit{Categorical}(\mathbf{p}_{ie}) \\ \mathbf{p}_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\ p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \boldsymbol{\tau}_e &= \{\tau_{e1} = -\infty, \dots, \tau_{eK-1}, \tau_{eK+1} = \infty\} \\ \tau_{ek} &\sim \textit{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\ \tau_{ck}^C &\sim \textit{Normal}(\tau_k^W, 0.25) \\ \tau_k^W &\sim \textit{Uniform}(-6, 6) \\ \beta_e &\sim \textit{Normal}^+(1, 1) \\ \zeta_i &\sim \textit{Normal}(z_i, 1)\end{aligned}$$

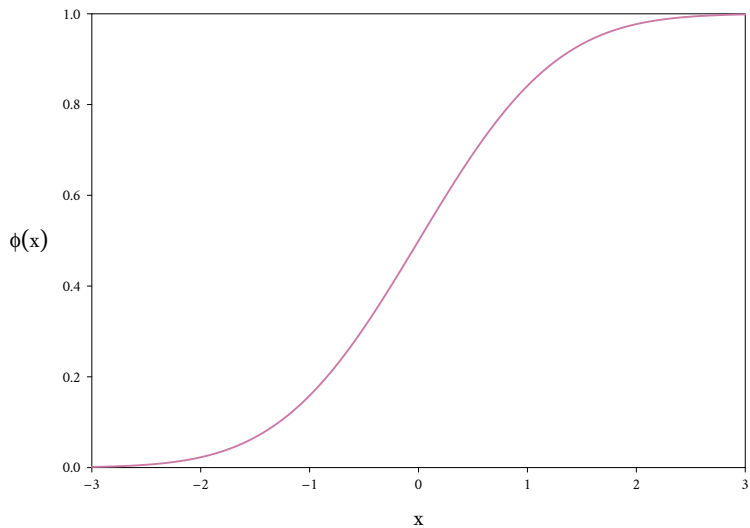
The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \textit{Categorical}(\mathbf{p}_{ie}) \\ \mathbf{p}_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\ p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \boldsymbol{\tau}_e &= \{\tau_{e1} = -\infty, \dots, \tau_{eK-1}, \tau_{eK+1} = \infty\} \\ \tau_{ek} &\sim \textit{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\ \tau_{ck}^C &\sim \textit{Normal}(\tau_k^W, 0.25) \\ \tau_k^W &\sim \textit{Uniform}(-6, 6) \\ \beta_e &\sim \textit{Normal}^+(1, 1) \\ \zeta_i &\sim \textit{Normal}(z_i, 1)\end{aligned}$$

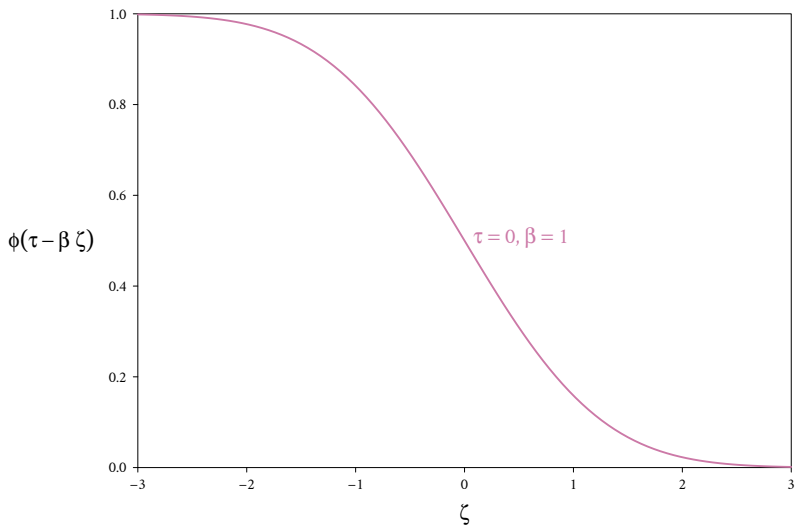
The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \text{Categorical}(p_{ie}) \\p_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\\tau_e &= \{\tau_{e1} = -\infty, \dots, \tau_{eK-1}, \tau_{eK+1} = \infty\} \\\tau_{ek} &\sim \text{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\\tau_{ck}^C &\sim \text{Normal}(\tau_k^W, 0.25) \\\tau_k^W &\sim \text{Uniform}(-6, 6) \\\beta_e &\sim \text{Normal}^+(1, 1) \\\zeta_i &\sim \text{Normal}(z_i, 1)\end{aligned}$$

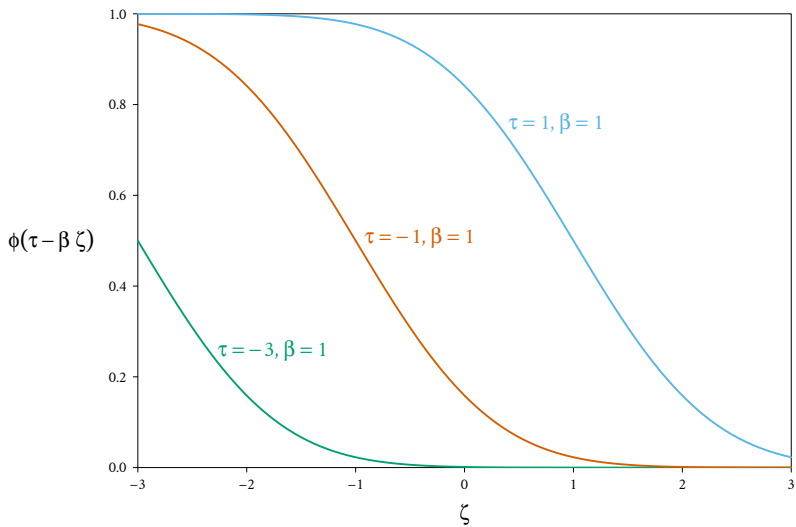
$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



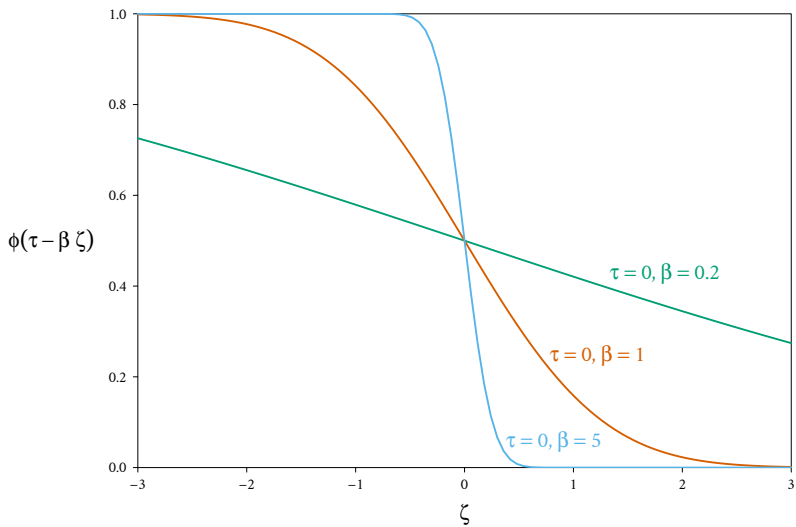
$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



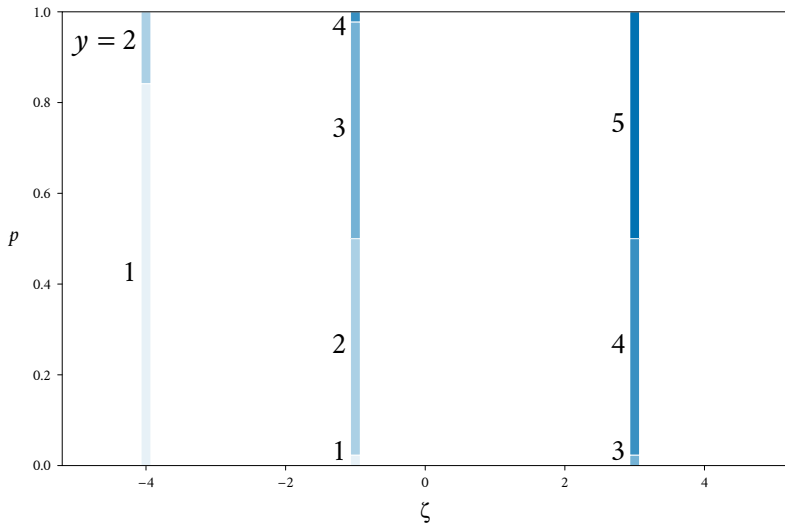
$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



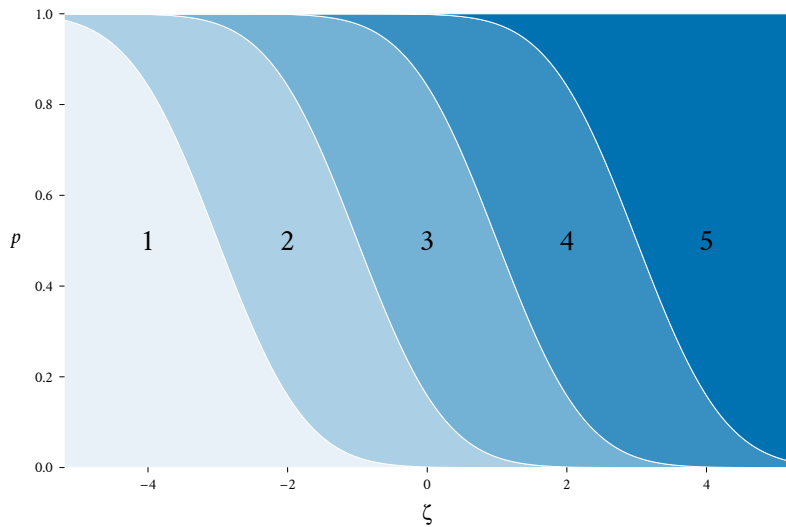
$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



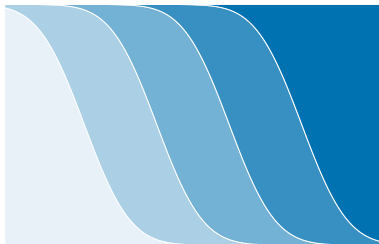
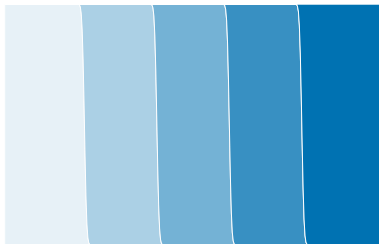
$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



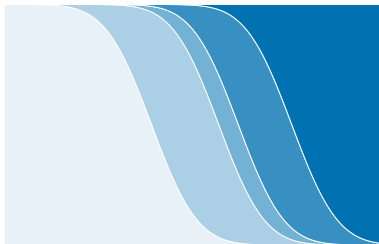
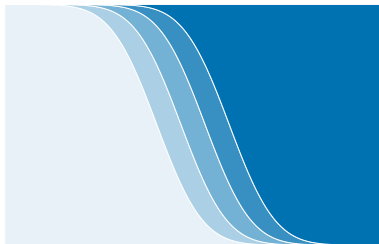
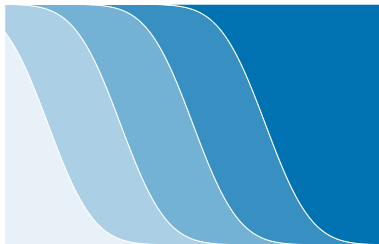
$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



$$p_{iek} = \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i)$$



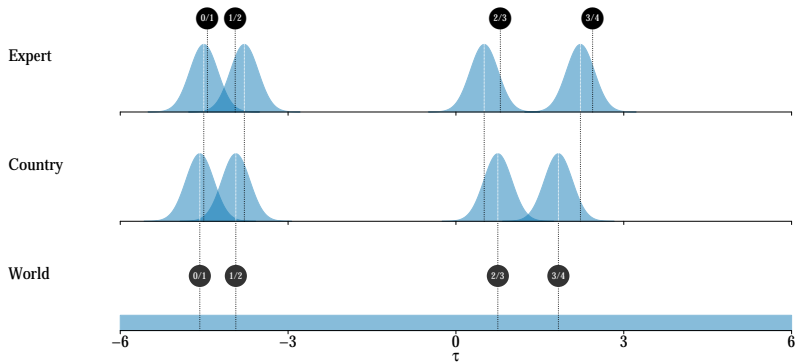
The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \textit{Categorical}(p_{ie}) \\p_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \boldsymbol{\tau}_e &= \{\tau_{e1} = -\infty, \dots, \tau_{eK-1}, \tau_{eK+1} = \infty\} \\\tau_{ek} &\sim \textit{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\\tau_{ck}^C &\sim \textit{Normal}(\tau_k^W, 0.25) \\\tau_k^W &\sim \textit{Uniform}(-6, 6) \\\beta_e &\sim \textit{Normal}^+(1, 1) \\\zeta_i &\sim \textit{Normal}(z_i, 1)\end{aligned}$$

$$\tau_{ek} \sim \text{Normal} \left(\tau_{c[e]k}^C, 0.25 \right), k \in \{2, \dots, K-1\}$$

$$\tau_{ck}^C \sim \text{Normal} \left(\tau_k^W, 0.25 \right)$$

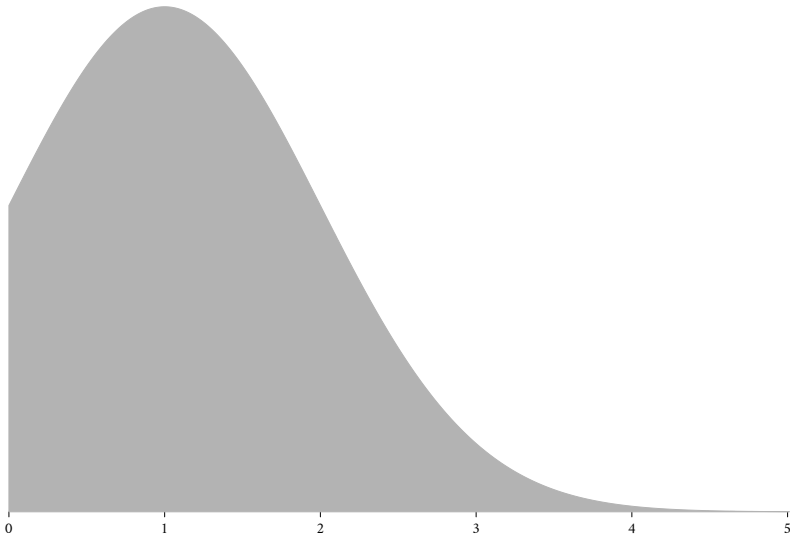
$$\tau_k^W \sim \text{Uniform}(-6, 6)$$



The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \textit{Categorical}(\mathbf{p}_{ie}) \\ \mathbf{p}_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\ p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \tau_{ek} &\sim \textit{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\ \tau_{ck}^C &\sim \textit{Normal}(\tau_k^W, 0.25) \\ \tau_k^W &\sim \textit{Uniform}(-6, 6) \\ \beta_e &\sim \textit{Normal}^+(1, 1) \\ \beta_e &\sim \textit{Normal}^+(1, 1) \\ \zeta_i &\sim \textit{Normal}(z_i, 1)\end{aligned}$$

$$\beta_e \sim \text{Normal}^+(1, 1)$$



The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \text{Categorical}(\mathbf{p}_{ie}) \\ \mathbf{p}_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\ p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \tau_{ek} &\sim \text{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\ \tau_{ck}^C &\sim \text{Normal}(\tau_k^W, 0.25) \\ \tau_k^W &\sim \text{Uniform}(-6, 6) \\ \beta_e &\sim \text{Normal}^+(1, 1) \\ \beta_e &\sim \text{Normal}^+(1, 1) \\ \zeta_i &\sim \text{Normal}(z_i, 1)\end{aligned}$$

The implementation



- Probabilistic programming language
- Bayesian and penalized likelihood estimation
- No-U-turns sampler (NUTS), Hamiltonian Monte Carlo
- Many interfaces: R, python, Stata etc.



The V-Dem Measurement Model

$$\begin{aligned}y_{ie} &\sim \textit{Categorical}(\mathbf{p}_{ie}) \\ \mathbf{p}_{ie} &= \{p_{ie1}, \dots, p_{ieK}\} \\ p_{iek} &= \phi(\tau_{e,k+1} - \beta_e \zeta_i) - \phi(\tau_{ek} - \beta_e \zeta_i) \\ \boldsymbol{\tau}_e &= \{\tau_{e1} = -\infty, \dots, \tau_{eK-1}, \tau_{eK+1} = \infty\} \\ \tau_{ek} &\sim \textit{Normal}(\tau_{c[e]k}^C, 0.25), \quad k \in \{2, \dots, K-1\} \\ \tau_{ck}^C &\sim \textit{Normal}(\tau_k^W, 0.25) \\ \tau_k^W &\sim \textit{Uniform}(-6, 6) \\ \beta_e &\sim \textit{Normal}^+(1, 1) \\ \zeta_i &\sim \textit{Normal}(z_i, 1)\end{aligned}$$


```
z_star ~ normal(0, 1);
for (j in 1:J)
  beta[j] ~ normal(1, 1)T[0,];
for (c in 1:C)
  tau_raw_country[c] ~ normal(tau_raw_world, 0.25);
for (j in 1:J)
  tau_raw_expert[j] ~ normal(tau_raw_country[expert_country[j]], 0.25);
for (r in 1:R) {
  p = Phi_approx(tau[j_id[r], y[r]+1] - z[cy_id[r]]*beta[j_id[r]])-
      Phi_approx(tau[j_id[r], y[r]] - z[cy_id[r]]*beta[j_id[r]]);
  target += log(p);
}
```

Demonstration

Demonstration

- https://github.com/jmedzihorsky/ELTE_29_01_2014/vdem_mm
- R and Stan code
- Data: a small slice of old V-Dem data

Thank you!