Agenda

1. **Presentations and proposals**
2. **Three-level models**
3. **Non-nested multilevel models**

Presentations

Format

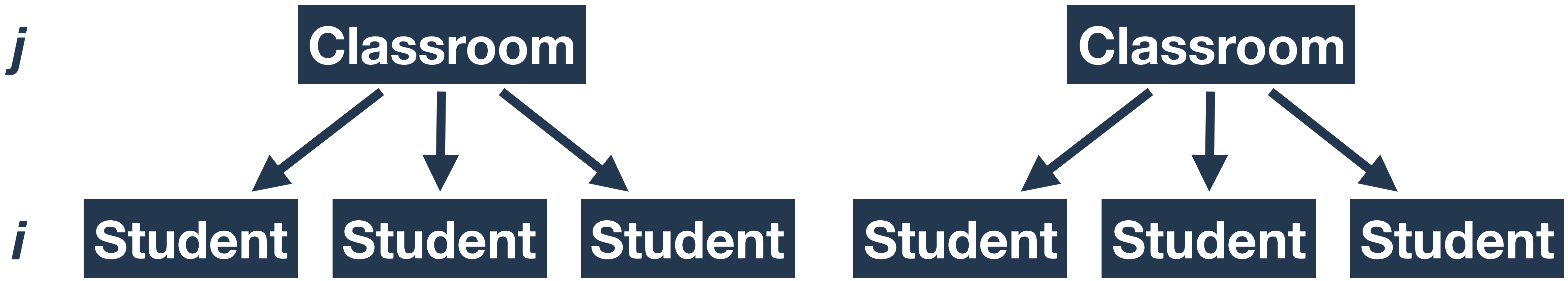
20 slides, automatically advancing every 20 seconds.
(Practice!)

Slot	Tue, April 9	Thu, April 11
1	Yildirim, Irem	Moloney, Kate
2	McCormack, Andrew	Hequet, Céline
3	Traves, Samantha	Nossek, Sean
4	Jutras, Kevin	Yang, Winnie
5	Carter-Rau, Rohan	Lee, Martha
6	Song, Sumin	Gounden Rock, Alyson
7	Amsden, Ryan	Zhao, Qiao
8	Jeong, Tay	Ng, Ka U
9	Isaac, Maike	Zhou, Lingyu
10	Moody, Alayne	

Three-level models

$$Math_{ij} \sim \text{MVNorm}(\mu_{ij}, \sigma)$$
$$\mu_{ij} = \beta_{0j} + \beta_{1j} Age_i$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Size_j + \eta_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11} Size_j + \eta_{1j}$$



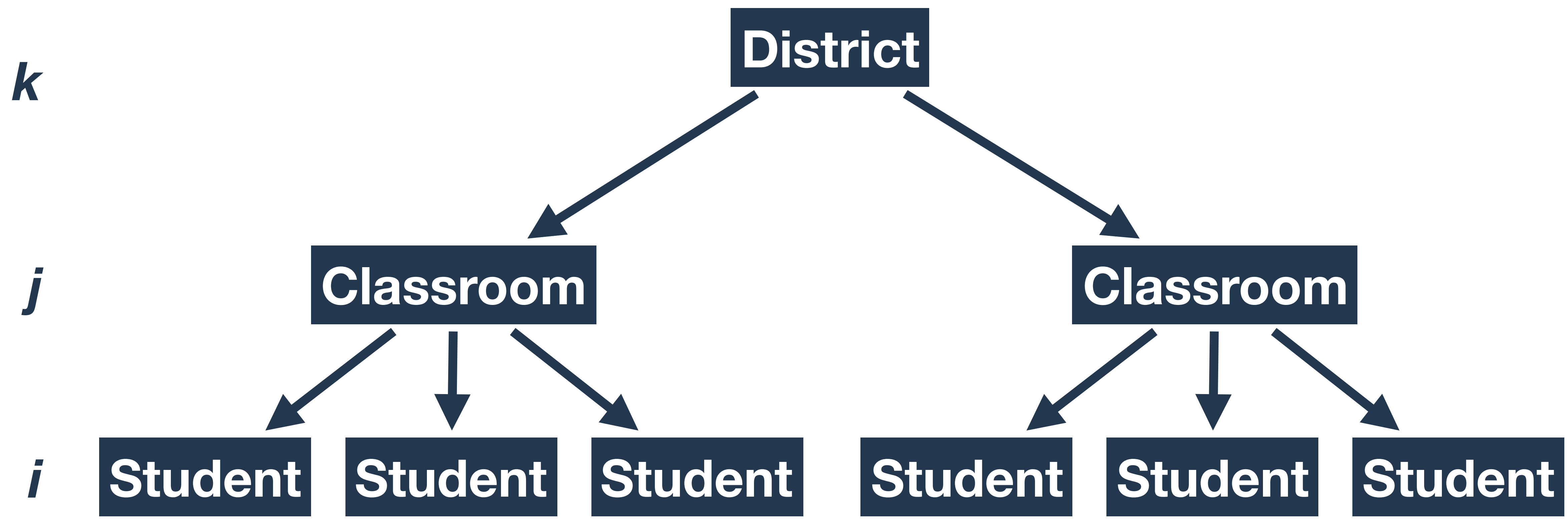
Three-level models

$$Math_{ij} \sim \text{MVNorm}(\mu_{ij}, \sigma)$$

$$\mu_{ij} = \beta_{0j} + \beta_{1j} Age_i$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Size_j + \eta_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Size_j + \eta_{1j}$$



Three-level models

$$
\begin{aligned}
Math_{ijk} &\sim \text{MVNorm}(\mu_{ijk}, \sigma) \\
\mu_{ijk} &= \beta_{0jk} + \beta_{1jk} Age_i \\
\\
\beta_{0jk} &= \gamma_{00k} + \gamma_{01k} Size_j + \eta_{0jk} \\
\beta_{1jk} &= \gamma_{10k} + \gamma_{11k} Size_j + \eta_{1jk} \\
\\
\gamma_{00k} &= \alpha_{000} + \nu_{00k} \\
\gamma_{01k} &= \alpha_{010} + \nu_{01k} \\
\gamma_{10k} &= \alpha_{100} + \nu_{10k} \\
\gamma_{11k} &= \alpha_{110} + \nu_{11k}
\end{aligned}
$$

Three-level models

Math score for student *i* in class *j* in district *k*.

$$\boxed{Math_{ijk}} \sim \text{MVNorm}(\mu_{ijk}, \sigma)$$

$$\mu_{ijk} = \beta_{0jk} + \beta_{1jk} Age_i$$

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} Size_j + \eta_{0jk}$$

$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k} Size_j + \eta_{1jk}$$

$$\gamma_{00k} = \alpha_{000} + \nu_{00k}$$

$$\gamma_{01k} = \alpha_{010} + \nu_{01k}$$

$$\gamma_{10k} = \alpha_{100} + \nu_{10k}$$

$$\gamma_{11k} = \alpha_{110} + \nu_{11k}$$

Three-level models

$$Math_{ijk} \sim \text{MVNorm}(\mu_{ijk}, \sigma)$$

$$\mu_{ijk} = \beta_{0jk} + \beta_{1jk} Age_i$$

Each district has its own average score.

$$\beta_{0jk} = \boxed{\gamma_{00k}} + \gamma_{01k} Size_j + \eta_{0jk}$$

$$\beta_{1jk} = \boxed{\gamma_{10k}} + \gamma_{11k} Size_j + \eta_{1jk}$$

The effect of age varies from district to district.

$$\gamma_{00k} = \alpha_{000} + v_{00k}$$

$$\gamma_{01k} = \alpha_{010} + v_{01k}$$

$$\gamma_{10k} = \alpha_{100} + v_{10k}$$

$$\gamma_{11k} = \alpha_{110} + v_{11k}$$

Three-level models

$$Math_{ijk} \sim \text{MVNorm}(\mu_{ijk}, \sigma)$$

$$\mu_{ijk} = \beta_{0jk} + \beta_{1jk} Age_i$$

The effect of class size varies from district to district.

$$\beta_{0jk} = \gamma_{00k} + \boxed{\gamma_{01k}} Size_j + \eta_{0jk}$$

$$\beta_{1jk} = \gamma_{10k} + \boxed{\gamma_{11k}} Size_j + \eta_{1jk}$$

The interaction between age and class size *also* varies by district.

$$\gamma_{00k} = \alpha_{000} + \nu_{00k}$$

$$\gamma_{01k} = \alpha_{010} + \nu_{01k}$$

$$\gamma_{10k} = \alpha_{100} + \nu_{10k}$$

$$\gamma_{11k} = \alpha_{110} + \nu_{11k}$$

Three-level models

$$Math_{ijk} \sim \text{MVNorm}(\mu_{ijk}, \sigma)$$

$$\mu_{ijk} = \beta_{0jk} + \beta_{1jk} Age_i$$

$$\beta_{0jk} = \gamma_{00k} + \gamma_{01k} Size_j + \boxed{\eta_{0jk}}$$

$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k} Size_j + \boxed{\eta_{1jk}}$$

Teacher-level random effects.

$$\gamma_{00k} = \alpha_{000} + \boxed{v_{00k}}$$

$$\gamma_{01k} = \alpha_{010} + \boxed{v_{01k}}$$

$$\gamma_{10k} = \alpha_{100} + \boxed{v_{10k}}$$

$$\gamma_{11k} = \alpha_{110} + \boxed{v_{11k}}$$

District-level random effects.

Three-level models

$$
\begin{aligned}
Math_{ijk} &\sim \text{MVNorm}(\mu_{ijk}, \sigma) \\
\mu_{ijk} &= \beta_{0jk} + \beta_{1jk} Age_i
\end{aligned}
$$

$$
\begin{aligned}
\beta_{0jk} &= \gamma_{00k} + \gamma_{01k} Size_j + \eta_{0jk} \\
\beta_{1jk} &= \gamma_{10k} + \gamma_{11k} Size_j + \eta_{1jk}
\end{aligned}
$$

$$
\begin{aligned}
\gamma_{00k} &= \alpha_{000} + v_{00k} \\
\gamma_{01k} &= \alpha_{010} + v_{01k} \\
\gamma_{10k} &= \alpha_{100} + v_{10k} \\
\gamma_{11k} &= \alpha_{110} + v_{11k}
\end{aligned}
$$

$$
\begin{aligned}
Math_{ijk} = \; & \alpha_{000} + \alpha_{010} Size_j + \alpha_{100} Age_i + \alpha_{110} Size_j Age_i \\
& v_{00k} + v_{01k} Size_j + v_{10k} Age_i + v_{11k} Size_j Age_i + \\
& \eta_{0jk} + \eta_{1jk} Age_i + \varepsilon_{ijk}
\end{aligned}
$$

Three-level models in R

$$
\begin{aligned}
Math_{ijk} = {} & \alpha_{000} + \alpha_{010} Size_j + \alpha_{100} Age_i + \alpha_{110} Size_j Age_i \\
& v_{00k} + v_{01k} Size_j + v_{10k} Age_i + v_{11k} Size_j Age_i + \\
& \eta_{0jk} + \eta_{1jk} Age_i + \varepsilon_{ijk}
\end{aligned}
$$

R formula

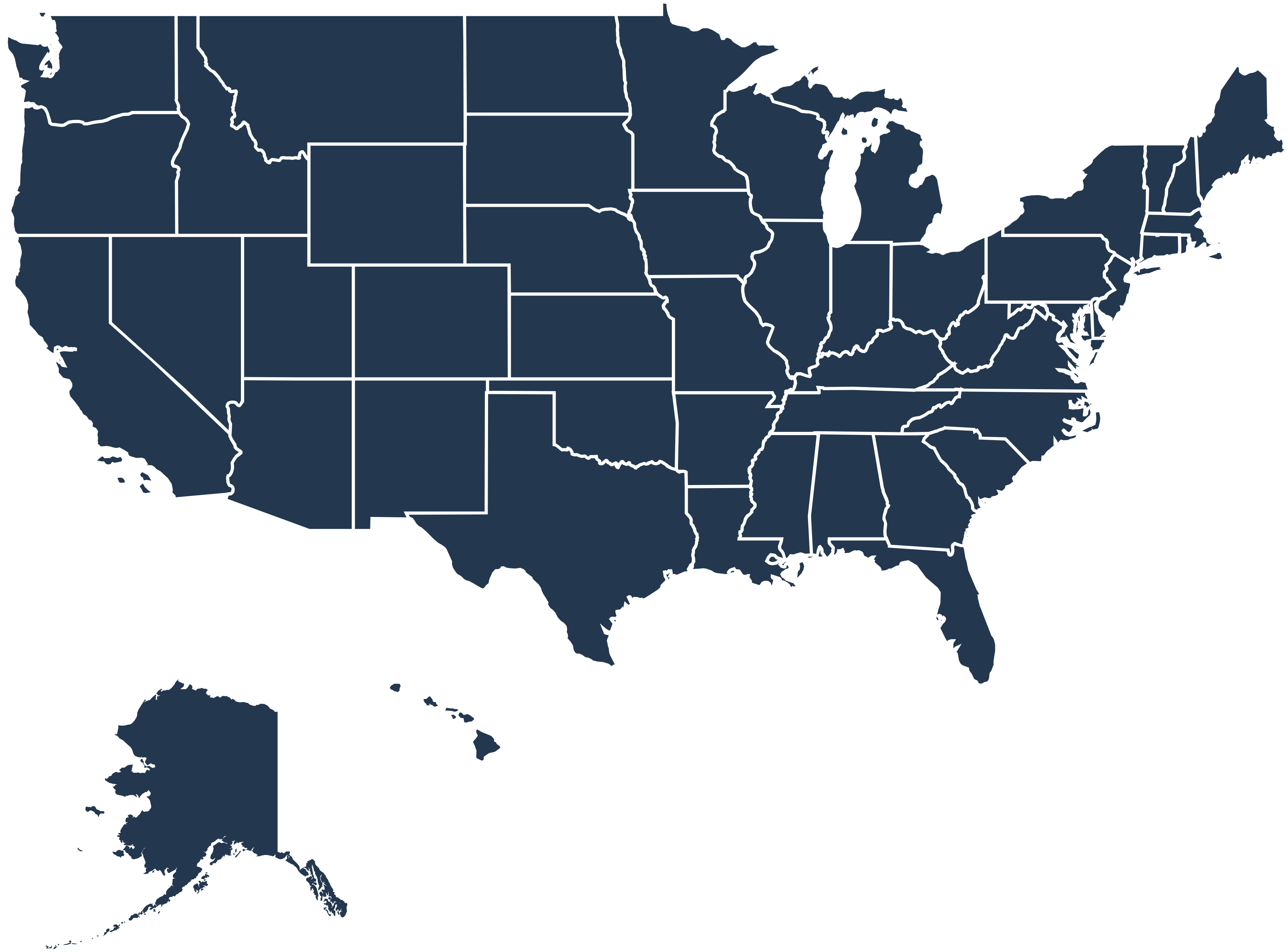
```
student_math_score ~
    student_age_s*class_size_c +
    (1 + student_age_s | teacher_id:district_id) +
    (1 + class_size*student_age_s | district_id)
```

Three-level models in R

$$
\begin{aligned}
Math_{ijk} &\sim \text{MVNorm}(\mu_{ijk}, \sigma) \\
\mu_{ijk} &= \beta_{0jk} + \beta_{1jk} Age_i \\
\\
\beta_{0jk} &= \gamma_{00k} + \gamma_{01k} Size_j + \eta_{0jk} \\
\beta_{1jk} &= \gamma_{10k} + \gamma_{11k} Size_j + \eta_{1jk} \\
\\
\gamma_{00k} &= \alpha_{000} + v_{00k} \\
\gamma_{01k} &= \alpha_{010} + v_{01k} \\
\gamma_{10k} &= \alpha_{100} + v_{10k} \\
\gamma_{11k} &= \alpha_{110} + v_{11k}
\end{aligned}
$$

	Mean	***90% credible interval***	
$\boldsymbol{\alpha_{000}}$	538.9	533.6	544.3
$\boldsymbol{\alpha_{010}}$	–1.38	–1.95	–0.79
$\boldsymbol{\alpha_{100}}$	–2.52	–4.05	–1.02
$\boldsymbol{\alpha_{110}}$	0.05	–0.21	0.32
$\boldsymbol{\phi_{\eta 0}}$	17.01	15.36	18.88
$\boldsymbol{\phi_{\eta 1}}$	1.40	0.06	3.23
$\boldsymbol{\phi_{v00}}$	13.62	6.18	19.46
$\boldsymbol{\phi_{v01}}$	0.28	0.01	0.68
$\boldsymbol{\phi_{v10}}$	1.86	0.08	4.27
$\boldsymbol{\phi_{v11}}$	0.09	0.01	0.19

Predicting inter-state migration



Predicting inter-state migration

Standard linear regression

$$\log(Flow_{ij}) \sim \text{Norm}(\mu_{ij}, \sigma)$$
$$\mu_{ij} = \beta_0 + \beta_1 Adj_{ij} + \beta_2 \log(SPop_i) + \beta_3 \log(SPop_j)$$

Flow$_{ij}$	Number of people that moved from state i to state j, 2015–16
Adj$_{ij}$	Indicator: state *i* shares a border with state *j*
SPop$_i$	Number of people that remained in state i, 2015–16

Attractive states

Two-level model can identify popular states to move to.

$$\log(Flow_{ij}) \sim \text{Norm}(\mu_{ij}, \sigma)$$

$$\mu_{ij} = \beta_{0j} + \beta_1 Adj_{ij} + \beta_2 \log(SPop_i)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \log(SPop_j) + \eta_{0j}$$

$\boldsymbol{\eta_{0j}}$ Unexplained attractiveness of state *j* as a destination

Non-nested model

Non-nested model identifies popular states to move into and to move out of.

$$\log\left(Flow_{ij}\right) \sim \text{Norm}\left(\mu_{ij}, \sigma\right)$$

$$\mu_{ij} = \beta_0 + \alpha_i + \omega_j + \beta_2 Adj_{ij}$$

$$\alpha_i = \gamma_{\alpha 1} \log\left(SPop_i\right) + \eta_{\alpha i}$$

$$\omega_j = \gamma_{\omega 1} \log\left(SPop_j\right) + \eta_{\omega j}$$

$\boldsymbol{\beta_0}$ Overall intercept (average log migration)

$\boldsymbol{\alpha_i}$ Effects specific to source state

$\boldsymbol{\omega_j}$ Effects specific to destination state

$\boldsymbol{\eta_{\alpha i}}$ Unexplained attractiveness of state *i* as a place to leave

$\boldsymbol{\eta_{\omega j}}$ Unexplained attractiveness of state *j* as a destination

Non-nested model

$$\log(Flow_{ij}) \sim \text{Norm}(\mu_{ij}, \sigma)$$

$$\mu_{ij} = \beta_0 + \alpha_i + \omega_j + \beta_2 Adj_{ij}$$

Second-level equations for α_i and ω_j have no intercept.

$$\alpha_i = \gamma_{\alpha 1} \log(SPop_i) + \eta_{\alpha i}$$

$$\omega_j = \gamma_{\omega 1} \log(SPop_j) + \eta_{\omega j}$$

$\boldsymbol{\beta_0}$ Overall intercept (average log migration)

$\boldsymbol{\alpha_i}$ Effects specific to source state

$\boldsymbol{\omega_j}$ Effects specific to destination state

$\boldsymbol{\eta_{\alpha i}}$ Unexplained attractiveness of state *i* as a place to leave

$\boldsymbol{\eta_{\omega j}}$ Unexplained attractiveness of state *j* as a destination

Non-nested model

$$\log(Flow_{ij}) \sim \mathrm{Norm}(\mu_{ij}, \sigma)$$

$$\mu_{ij} = \beta_0 + a_i + \omega_j + \beta_2 Adj_{ij}$$

$$a_i = \gamma_{a1} \log(SPop_i) + \eta_{ai}$$

$$\omega_j = \gamma_{\omega 1} \log(SPop_j) + \eta_{\omega j}$$

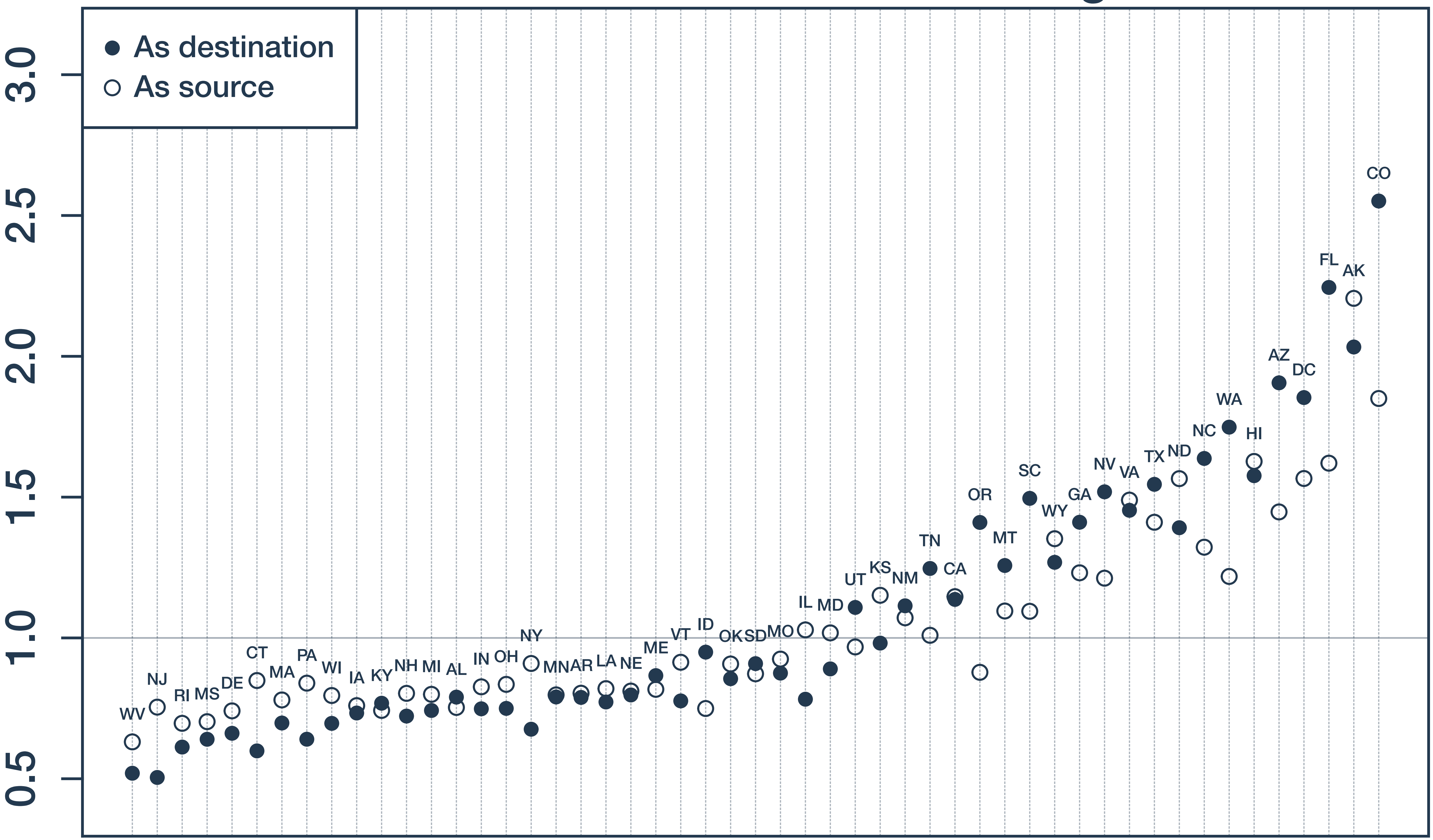
source state	dest. state	Adj	Log flow	Source log pop	Dest. log pop
AL	AK	0	5.3	14.3	12.5
AL	CA	0	7.3	14.3	16.4
AL	FL	1	8.8	14.3	15.8
AK	AL	0	5.4	12.5	14.3
AK	CA	0	7.3	12.5	16.4
AK	FL	0	6.7	12.5	15.8
CA	AL	0	7.4	16.4	12.5
…	…	…	…	…	…

R formula

```
log_flow ~ adjacent +
    (1 | source_state) + (1 | destination_state)
```

Non-nested model

State migration factors



Non-nested models

Multi-cohort panels of students	Each outcome (test score, e.g.) is associated with one student and one teacher. Students have multiple teachers and teachers have multiple classes.
Journal publications	Authors can contribute to multiple articles and multiple journals.
Multi-factor experiments	Research subjects exposed to multiple stimuli in multiple contexts.
Simple networked data	International trade, friendship nominations, Twitter mentions, bullying, …