Agenda

1. Models of time
2. Nesting time within students
3. Three-level models
4. Non-nested models

Presentations

Format

- 20 slides, auto-advancing every 20 seconds
- Slides can be duplicated for longer explanations
- You only have 6 minutes and 40 seconds, so be concise!
- Focus on motivation, (brief) data description, and results — don't get bogged down in model details
- Submit to me as PDF with 20 pages
- More details at https://soci620.netlify.app/pages/presentations.html

Models of time

Models of time

Common models of time

Autoregression models (and Gaussian processes)

Model outcome at time *t* as a function of covariates and outcome at time *t*–1.
$y_t = y_{t-1} + \beta X_{t-1} + \varepsilon_t$
$y_{\Delta t} = \beta X_{t-1} + \varepsilon_t$

Survival / event-history models

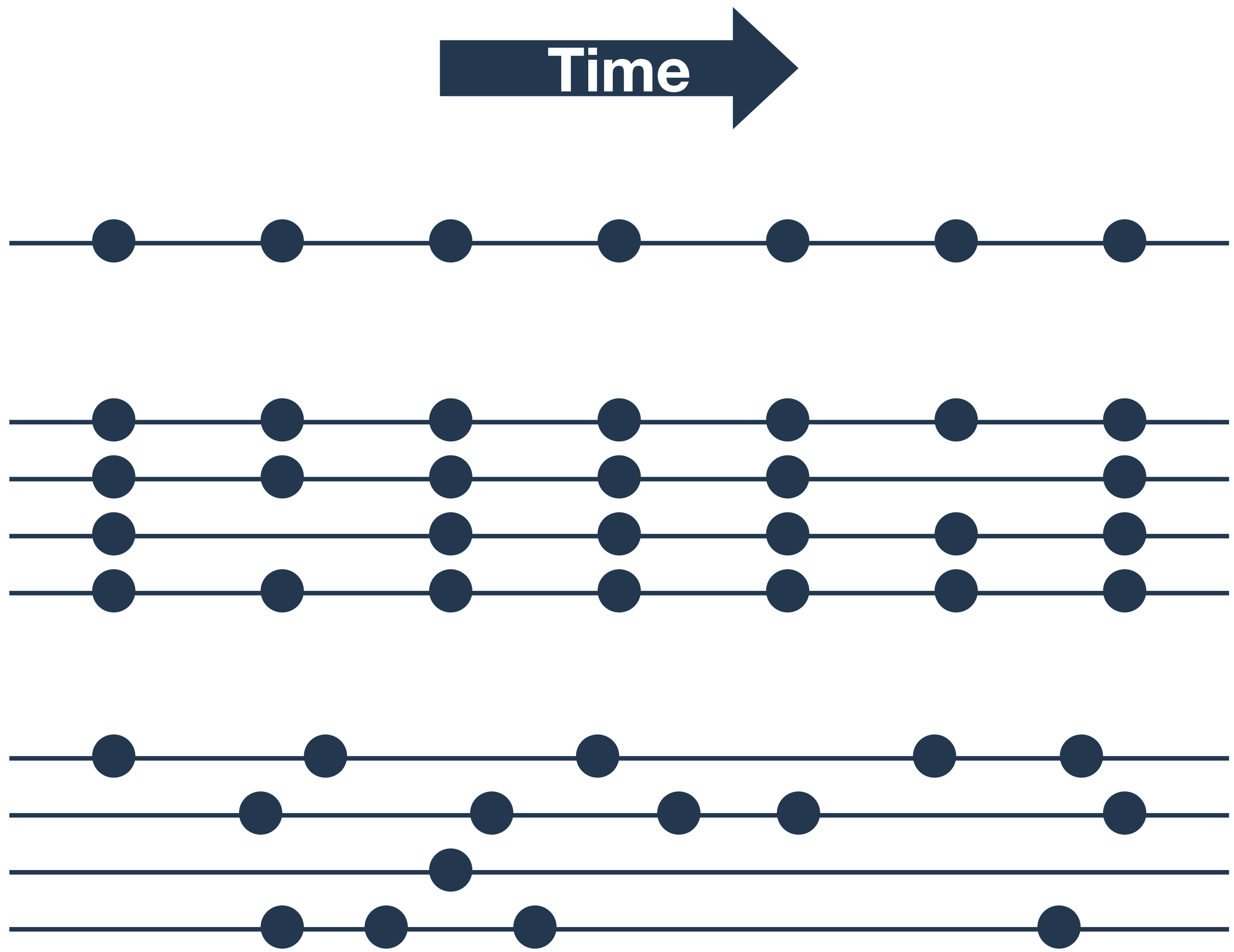
Model the timing of a one-time event (graduation, job acquisition, death).
$\lambda(t \mid X) = \lambda_0 \exp(\beta X)$

***Ad hoc* models**

Countless context-specific ways to model a randomly varying or functionally defined effect of time on outcomes.

Models of time

Common temporal data structures



Time series
Attitude surveys

Panel data
Tennessee STAR

Unstructured longitudinal data
Convenience samples; observational data

Student scores over time

Correlation between scores from the same student.

Student 1	540	531	564	563
Student 2	479	487	505	510
Student 3	503	505	501	503
Student 4	461	471	—	—
Student 5	525	—	506	541
	…	…	…	…
	1985	**1986**	**1987**	**1988**

Student scores over time

	1985	1986	1987	1988
Student 1	540	531	564	563
Student 2	479	487	505	510
Student 3	503	505	501	503
Student 4	461	471	—	—
Student 5	525	—	506	541
	…	…	…	…

Scores tend to increase over time.

Student scores over time

"Wide" vs "tall" tabular data:

	1985	1986	1987	1988
Student 1	540	531	564	563
Student 2	479	487	505	510
Student 3	503	505	501	503
Student 4	461	471	—	—
Student 5	525	—	506	541
	…	…	…	…

Wide

→

Student	Year	Score
1	1985	540
1	1986	531
1	1987	564
1	1988	563
2	1985	479
2	1986	487
2	1987	505
2	1988	510
3	1985	503
…	…	…

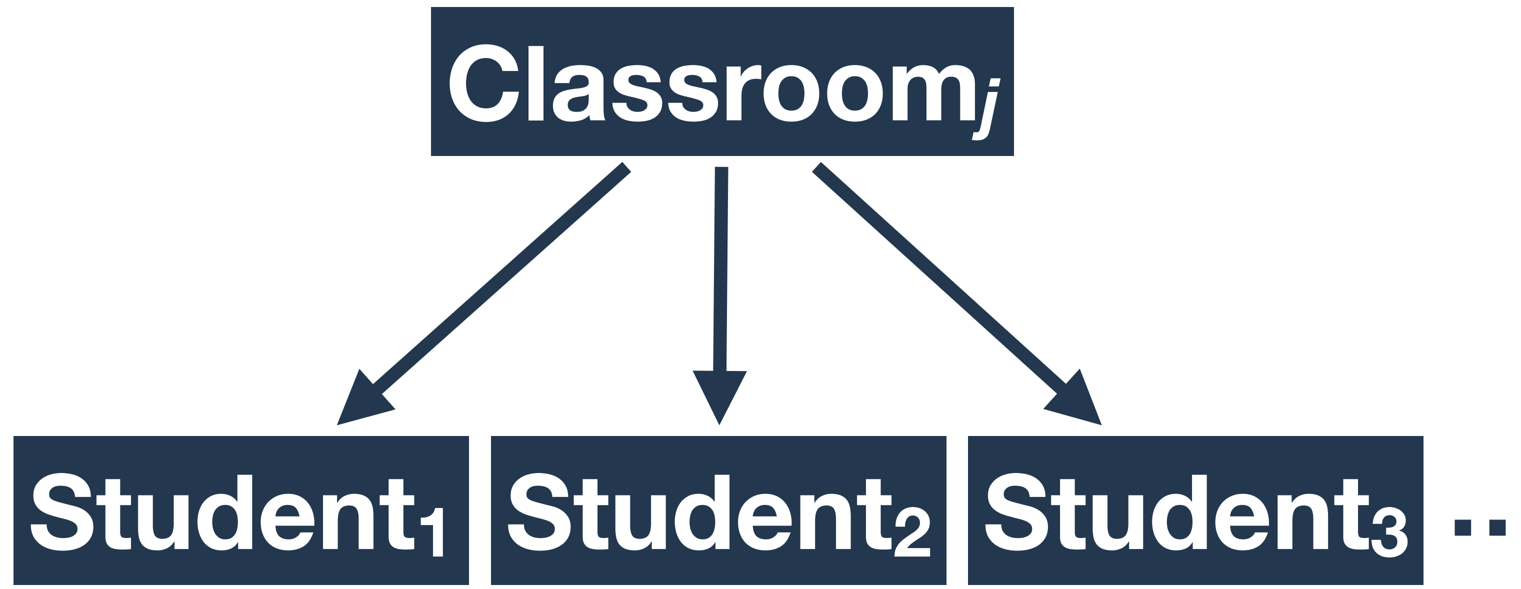
Tall

Nesting time within students

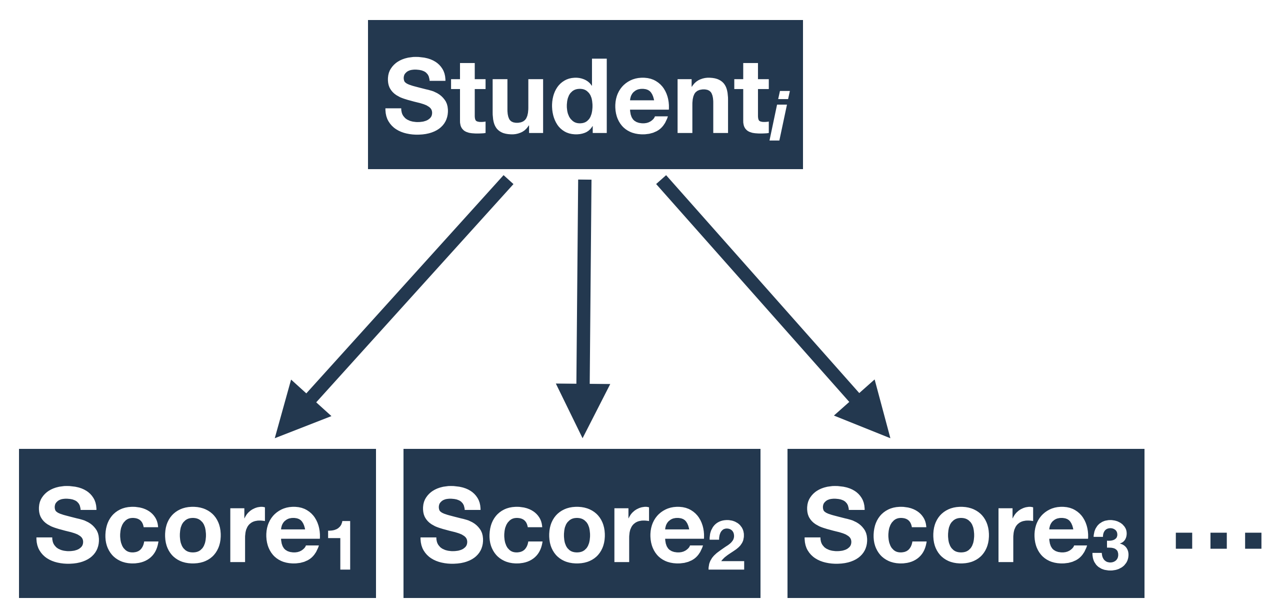
Student random effects

Students groups into classrooms

Single year



Test scores grouped by student



Student random effects

Score for student i at time t.

$$S_{ti} \sim \text{Norm}(\mu_{ti}, \sigma)$$

Average score for student i.

$$\mu_{ti} = \beta_{0i} + \beta_1 CSize_{ti}$$

$$\beta_{0i} = \gamma_{00} + \eta_{0i}$$

Average score across all students.

$$\eta_{0i} \sim \text{Norm}(0, \phi_0)$$

Student random effects

$$S_{ti} \sim \text{Norm}(\mu_{ti}, \sigma)$$
$$\mu_{ti} = \beta_{0i} + \beta_1 CSize_{ti}$$
$$\beta_{0i} = \gamma_{00} + \eta_{0i}$$

$$\eta_{0i} \sim \text{Norm}(0, \phi_0)$$

$$\gamma_{00} \sim \text{Norm}(500, 100)$$
$$\beta_1 \sim \text{Norm}(0, 50)$$
$$\sigma \sim \text{HalfCauchy}(0, 50)$$
$$\phi_0 \sim \text{HalfCauchy}(0, 50)$$

	Mean	*90% credible interval*	
$\boldsymbol{\gamma_{00}}$	547.06	544.05	549.94
$\boldsymbol{\beta_1}$	1.23	0.58	1.87
$\boldsymbol{\sigma}$	56.94	54.97	58.93
$\boldsymbol{\phi_0}$	33.91	30.16	37.60

Linear time trend

$$
\begin{aligned}
S_{ti} &\sim \text{Norm}(\mu_{ti}, \sigma) \\
\mu_{ti} &= \beta_{0i} + \beta_{1i} Year_{ti} + \beta_2 CSize_{ti} \\
\beta_{0i} &= \gamma_{00} + \eta_{0i} \\
\beta_{1i} &= \gamma_{10} + \eta_{1i}
\end{aligned}
$$

$$
[\eta_{0i}, \eta_{1i}] \sim \text{MVNorm}([0, 0], \Phi, R)
$$

Linear time trend

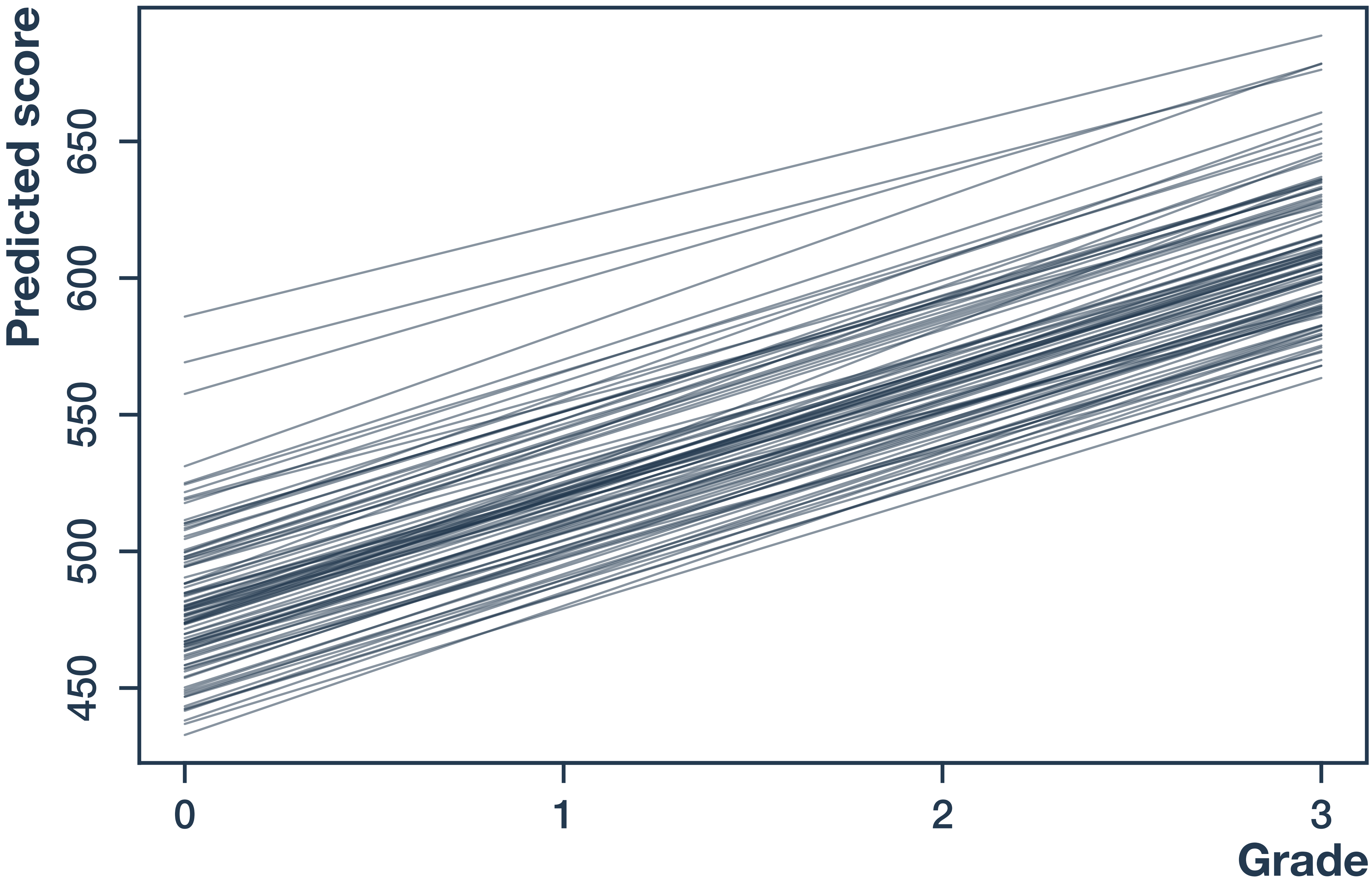
$$S_{ti} \sim \text{Norm}(\mu_{ti}, \sigma)$$
$$\mu_{ti} = \beta_{0i} + \beta_{1i} Year_{ti} + \beta_2 CSize_{ti}$$
$$\beta_{0i} = \gamma_{00} + \eta_{0i}$$
$$\beta_{1i} = \gamma_{10} + \eta_{1i}$$

$$[\eta_{0i}, \eta_{1i}] \sim \text{MVNorm}([0, 0], \Phi, R)$$

$$\gamma_{00} \sim \text{Norm}(500, 100)$$
$$\gamma_{10} \sim \text{Norm}(0, 50)$$
$$\beta_2 \sim \text{Norm}(0, 50)$$
$$\sigma \sim \text{HalfCauchy}(0, 50)$$
$$\phi_0, \phi_1 \sim \text{HalfCauchy}(0, 50)$$
$$R \sim LKJ(2, 2)$$

	Mean	***90% credible interval***	
$\boldsymbol{\gamma_{00}}$	484.64	481.84	487.39
$\boldsymbol{\gamma_{10}}$	42.88	41.75	44.02
$\boldsymbol{\beta_2}$	–0.33	–0.73	0.070
$\boldsymbol{\sigma}$	24.27	23.27	25.32
$\boldsymbol{\phi_0}$	38.25	35.84	40.65
$\boldsymbol{\phi_1}$	8.89	7.17	10.56
$\boldsymbol{\rho_{01}}$	–0.37	–0.48	–0.23

Linear time trend



Quadratic time trend

$$
\begin{aligned}
S_{ti} &\sim \text{Norm}(\mu_{ti}, \sigma) \\
\mu_{ti} &= \beta_{0i} + \beta_{1i} Year_{ti} + \beta_{2i} Year_{ti}^2 + \beta_3 CSize_{ti} \\
\beta_{0i} &= \gamma_{00} + \eta_{0i} \\
\beta_{1i} &= \gamma_{10} + \eta_{1i} \\
\beta_{2i} &= \gamma_{20} + \eta_{2i}
\end{aligned}
$$

$$[\eta_{0i}, \eta_{1i}, \eta_{2i}] \sim \text{MVNorm}([0, 0, 0], \Phi, R)$$

Quadratic time trend

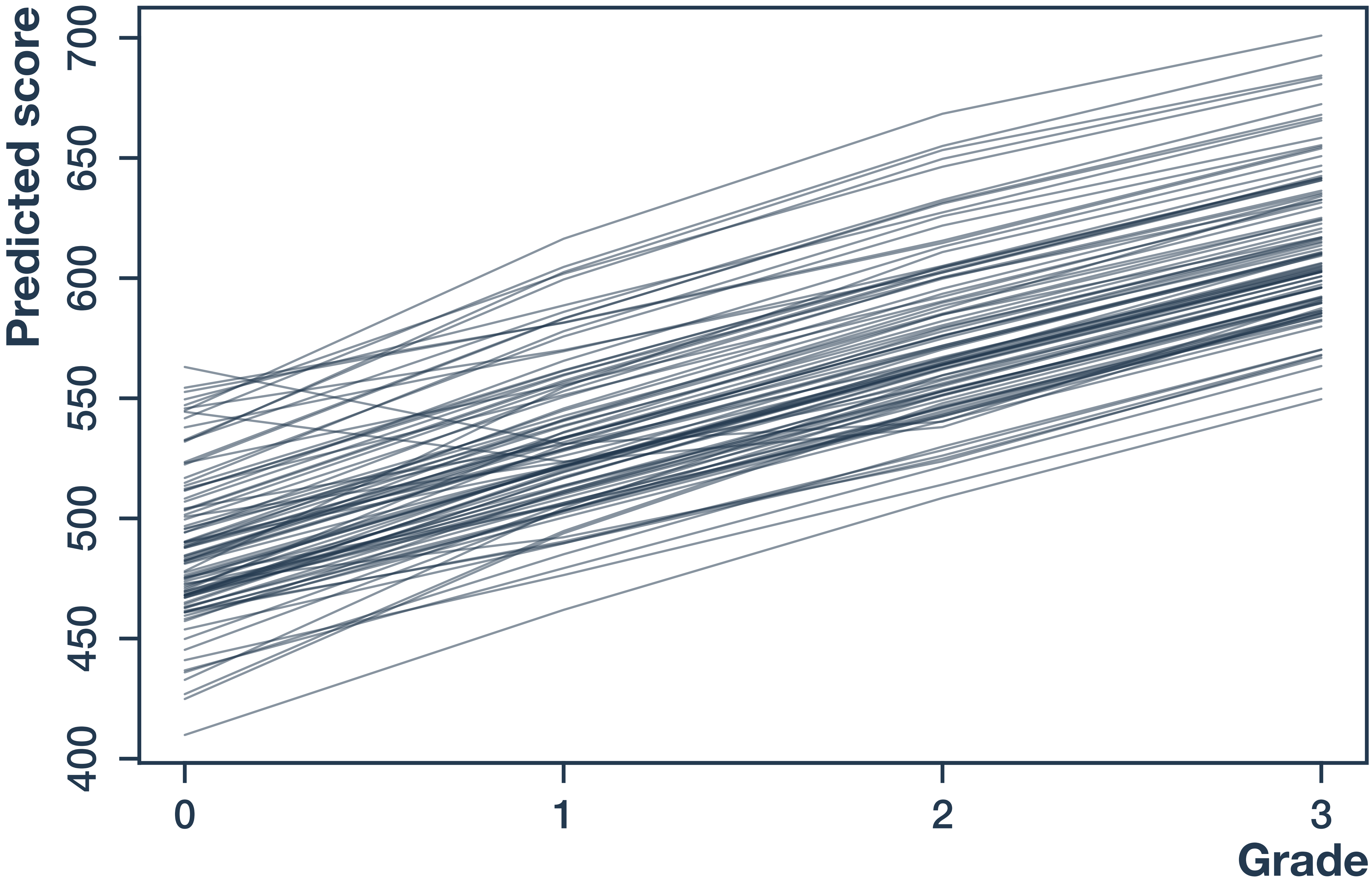
$$S_{ti} \sim \text{Norm}(\mu_{ti}, \sigma)$$
$$\mu_{ti} = \beta_{0i} + \beta_{1i} Year_{ti} + \beta_{2i} Year_{ti}^2 + \beta_3 CSize_{ti}$$
$$\beta_{0i} = \gamma_{00} + \eta_{0i}$$
$$\beta_{1i} = \gamma_{10} + \eta_{1i}$$
$$\beta_{2i} = \gamma_{20} + \eta_{2i}$$

$$[\eta_{0i}, \eta_{1i}, \eta_{2i}] \sim \text{MVNorm}([0,0,0], \Phi, R)$$

$$\gamma_{00} \sim \text{Norm}(500, 100)$$
$$\gamma_{10} \sim \text{Norm}(0, 50)$$
$$\gamma_{20} \sim \text{Norm}(0, 50)$$
$$\beta_3 \sim \text{Norm}(0, 50)$$
$$\sigma \sim \text{HalfCauchy}(0, 50)$$
$$\phi_0, \phi_1, \phi_2 \sim \text{HalfCauchy}(0, 50)$$
$$R \sim LKJ(2, 3)$$

	Mean	*90% credible interval*	
γ_{00}	482.32	479.37	485.30
γ_{10}	49.28	45.77	52.88
γ_{20}	–2.10	–3.14	–1.09
β_3	–0.36	–0.74	0.02
σ	21.91	20.76	23.04
ϕ_0	40.70	38.16	43.43
ϕ_1	31.48	26.35	36.38
ϕ_2	7.29	5.62	8.96
ρ_{01}	–0.45	–0.55	–0.35
ρ_{02}	0.40	0.25	0.53
ρ_{12}	–0.98	–1.00	–0.96

Quadratic time trend

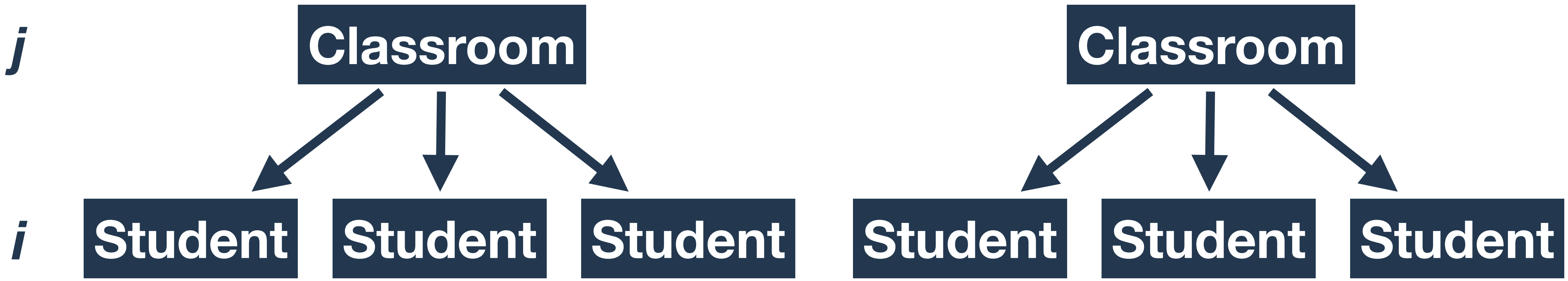


Three-level models

Three-level models

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

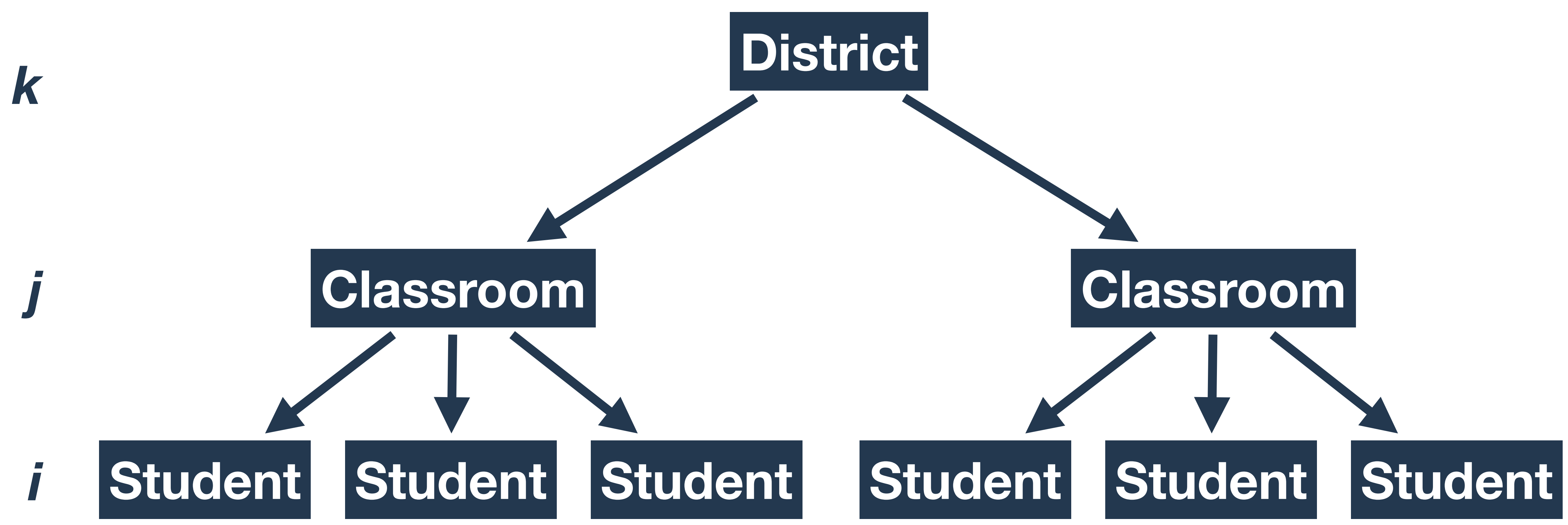
$$\beta_{0j} = \gamma_{00} + \gamma_{01} Size_j + \eta_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11} Size_j + \eta_{1j}$$



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

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

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

$$\gamma_{00k} = \alpha_{000} + \nu_{00k}$$
$$\gamma_{01k} = \alpha_{010} + \nu_{01k}$$
$$\gamma_{10k} = \alpha_{100} + \nu_{10k}$$
$$\gamma_{11k} = \alpha_{110} + \nu_{11k}$$

Three-level models

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

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

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

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

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

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

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

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

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

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.

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

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

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

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

Three-level models

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

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

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)
```

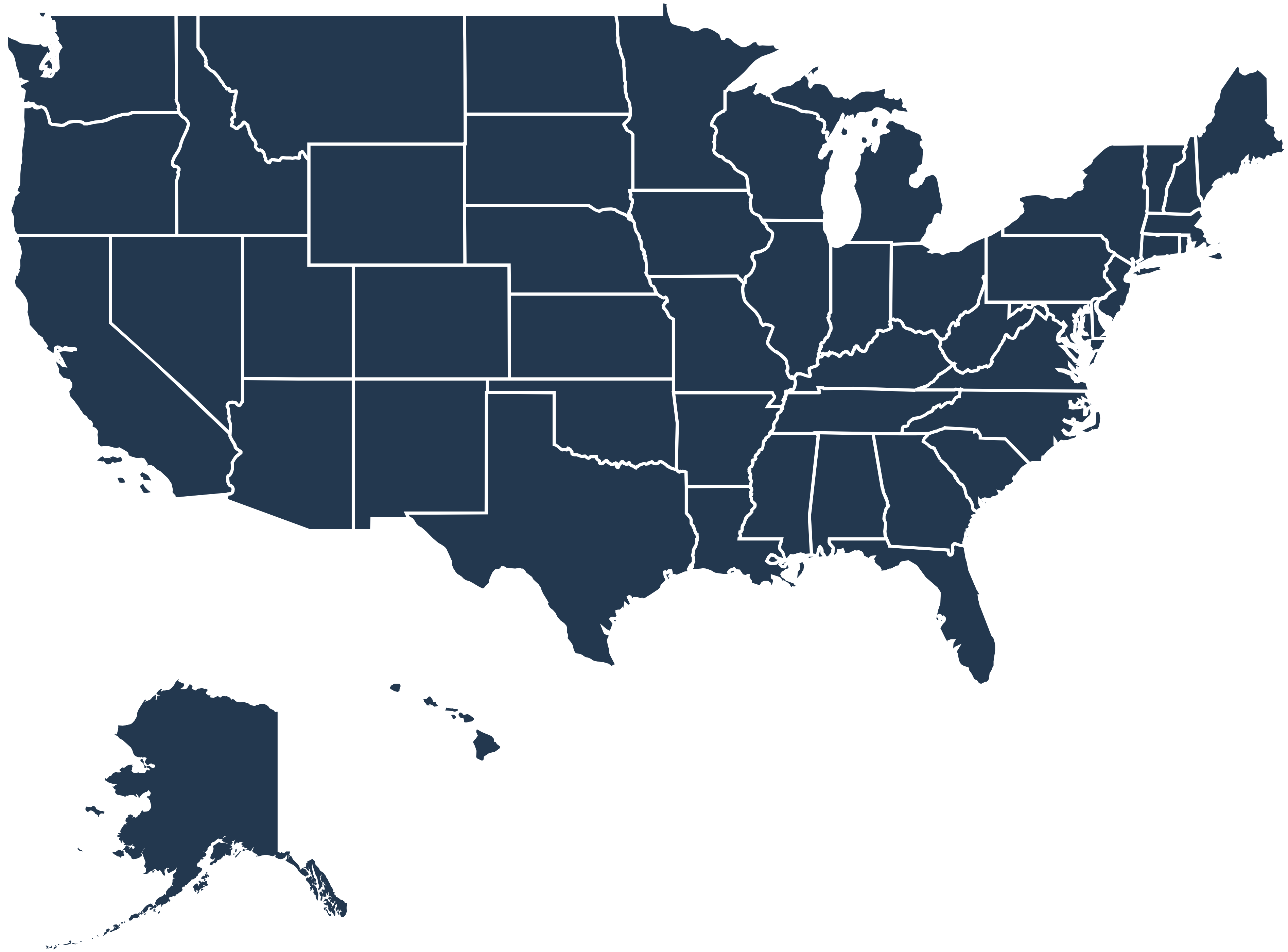
Three-level models in R

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

	Mean	*90% credible interval*	
α_{000}	538.9	533.6	544.3
α_{010}	–1.38	–1.95	–0.79
α_{100}	–2.52	–4.05	–1.02
α_{110}	0.05	–0.21	0.32
$\phi_{\eta 0}$	17.01	15.36	18.88
$\phi_{\eta 1}$	1.40	0.06	3.23
$\phi_{\nu 00}$	13.62	6.18	19.46
$\phi_{\nu 01}$	0.28	0.01	0.68
$\phi_{\nu 10}$	1.86	0.08	4.27
$\phi_{\nu 11}$	0.09	0.01	0.19

Non-nested models

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.

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

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

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

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

$\boldsymbol{\beta_0}$ 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

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

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

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

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

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

β_0 Overall intercept (average log migration)

α_i Effects specific to source state

ω_j Effects specific to destination state

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

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

Non-nested model

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

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

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

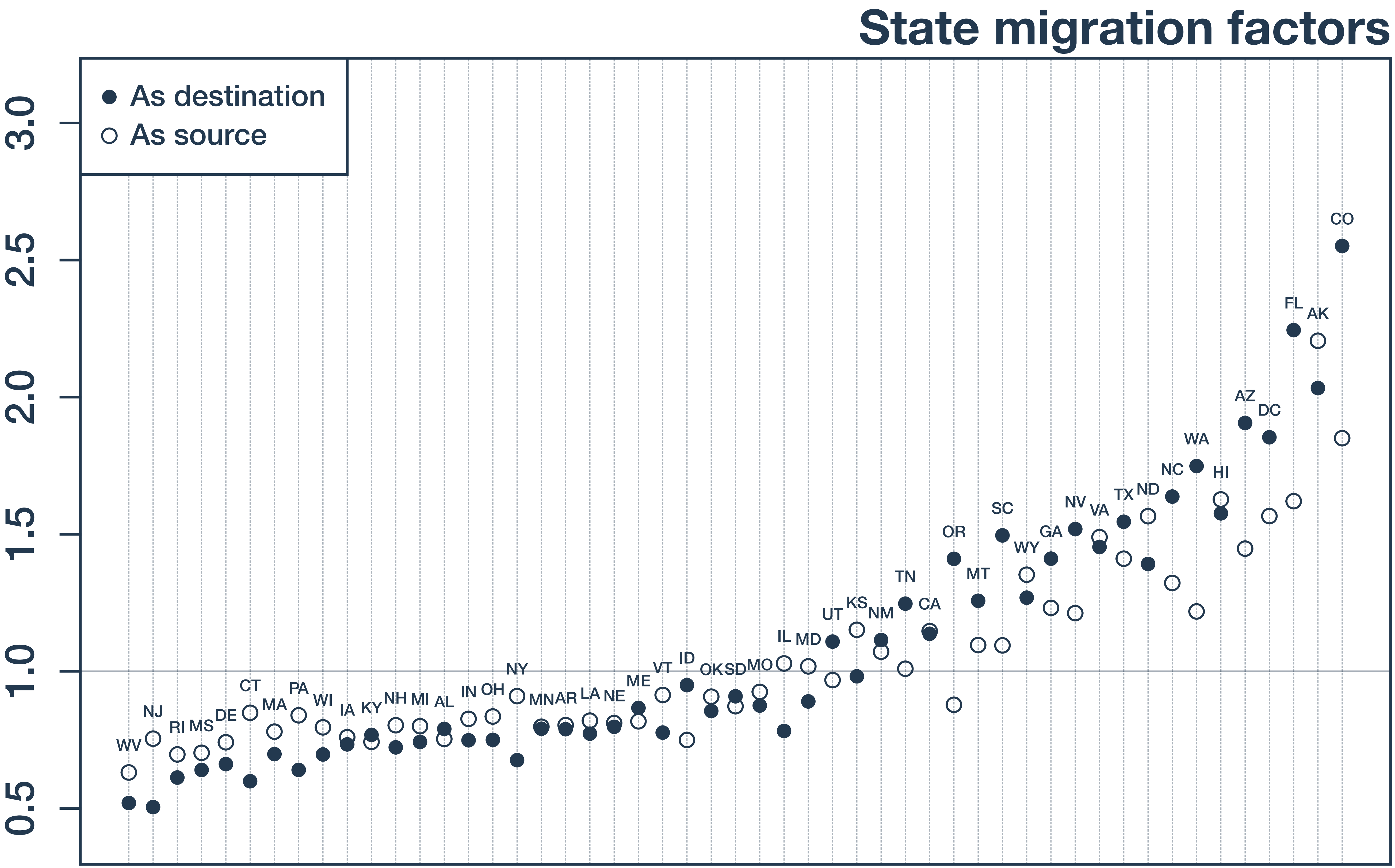
$$\omega_j = \gamma_{\omega 1} \log(SPop_j) + \eta_{\omega j}$$

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 + log_pop_src + log_pop_dest +
    (1 | source_state) + (1 | destination_state)
```

Non-nested model



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, …