HSS 8005

Quantitative analysis +

2024

Dr Chris Moreh

Week 7

Temporalities

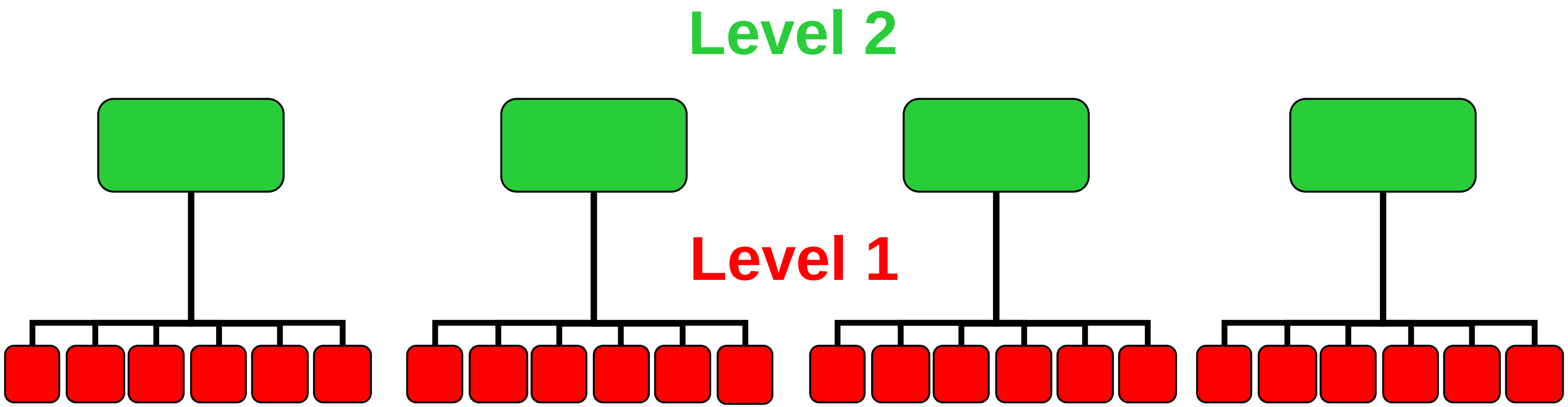
Time series, panel and longitudinal data analysis

Click and press F for full screen

View on ReCap

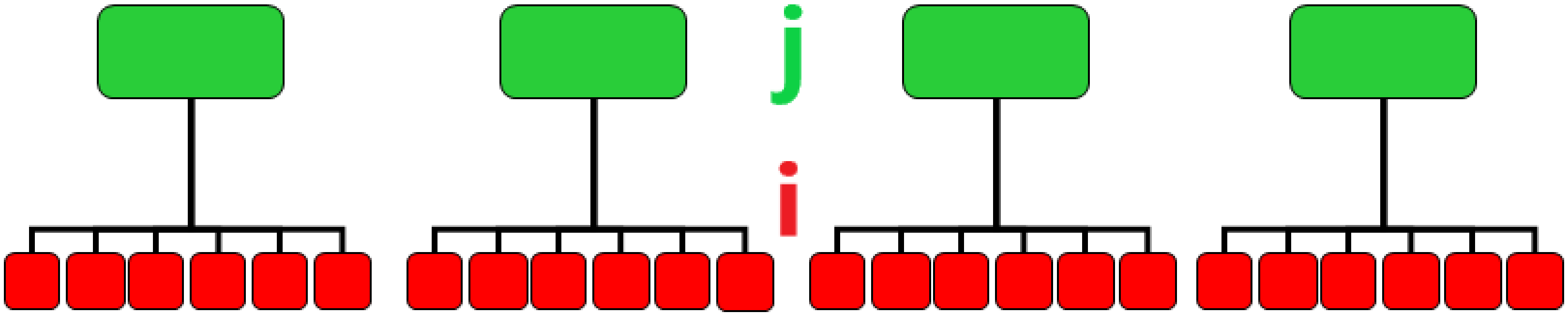
Hierarchical data: review

- We can refer to the nesting structure in terms of “Levels”…



Hierarchical data: review

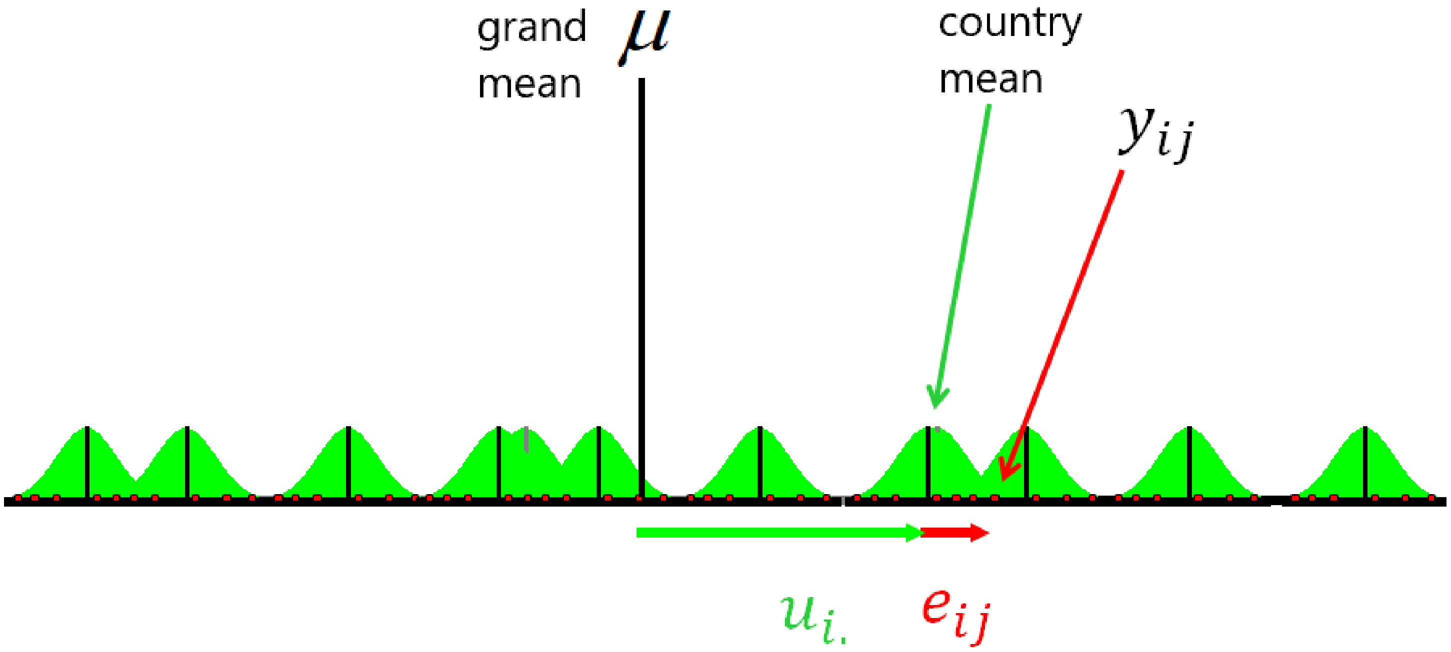
- For example: **individual respondents** nested (grouped) within **countries**



$$y_{ij} = y_{\text{respondent, country}}$$

Hierarchical models: review

- For example: **individual respondents** nested (grouped) within **countries**



$$y_{ij} = \mu + u_{i.} + e_{ij}$$

Hierarchical models: review

- For example: **individual respondents** nested (grouped) within **countries**

$$\underbrace{y_{ij}}_{\textbf{Observed}} = \underbrace{\mu}_{\textbf{Fixed}} + \underbrace{u_{i.} + e_{ij}}_{\textbf{Random}}$$

$$u_{i.} \sim N(0, \tau^2)$$

$$e_{ij} \sim N(0, \sigma^2)$$

Hierarchical data: temporality

- A minor change in perspective:
  - The Level 1 “observations” are measurements at different time-points
  - The Level 2 “groups” are the units (e.g. people, countries, schools, etc.) measured over time
- Example toy dataset: a simulated dataset contains repeated measurements of test scores for adolescents from ages 14 to 18.
- The repeated observations are nested within student and the students are nested within teachers’ classrooms.
- There are 75 students in total

Hierarchical data: temporality

```r
datawizard::data_peek(test)
```

```
Data frame with 375 rows and 9 variables

Variable   | Type    | Values
------------------------------------------------------------------------
teacher    | numeric | 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
student    | numeric | 1, 1, 1, 1, 1, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, ...
age        | numeric | 14, 15, 16, 17, 18, 14, 15, 16, 17, 18, 14, ...
cage       | numeric | -2, -1, 0, 1, 2, -2, -1, 0, 1, 2, -2, -1, 0, ...
sex        | factor  | Male, Male, Male, Male, Male, Male, ...
score      | numeric | 64, 68, 69, 71, 75, 58, 60, 71, 69, 72, 54, ...
pass       | numeric | 0, 0, 0, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, ...
experience | numeric | 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, ...
zero       | numeric | 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...
```

Hierarchical data: temporality

```
gt::gt_preview(test, top_n = 3, bottom_n = 3, incl_rownums = TRUE)
```

	teacher	student	age	cage	sex	score	pass	experience	zero
1	1	1	14	-2	Male	64	0	5	0
2	1	1	15	-1	Male	68	0	5	0
3	1	1	16	0	Male	69	0	5	0
4..372									
373	3	75	16	0	Female	81	1	8	0
374	3	75	17	1	Female	85	1	8	0
375	3	75	18	2	Female	86	1	8	0

Hierarchical data: temporality

```
test |> datawizard::to_numeric() |>
  datawizard::describe_distribution()
```

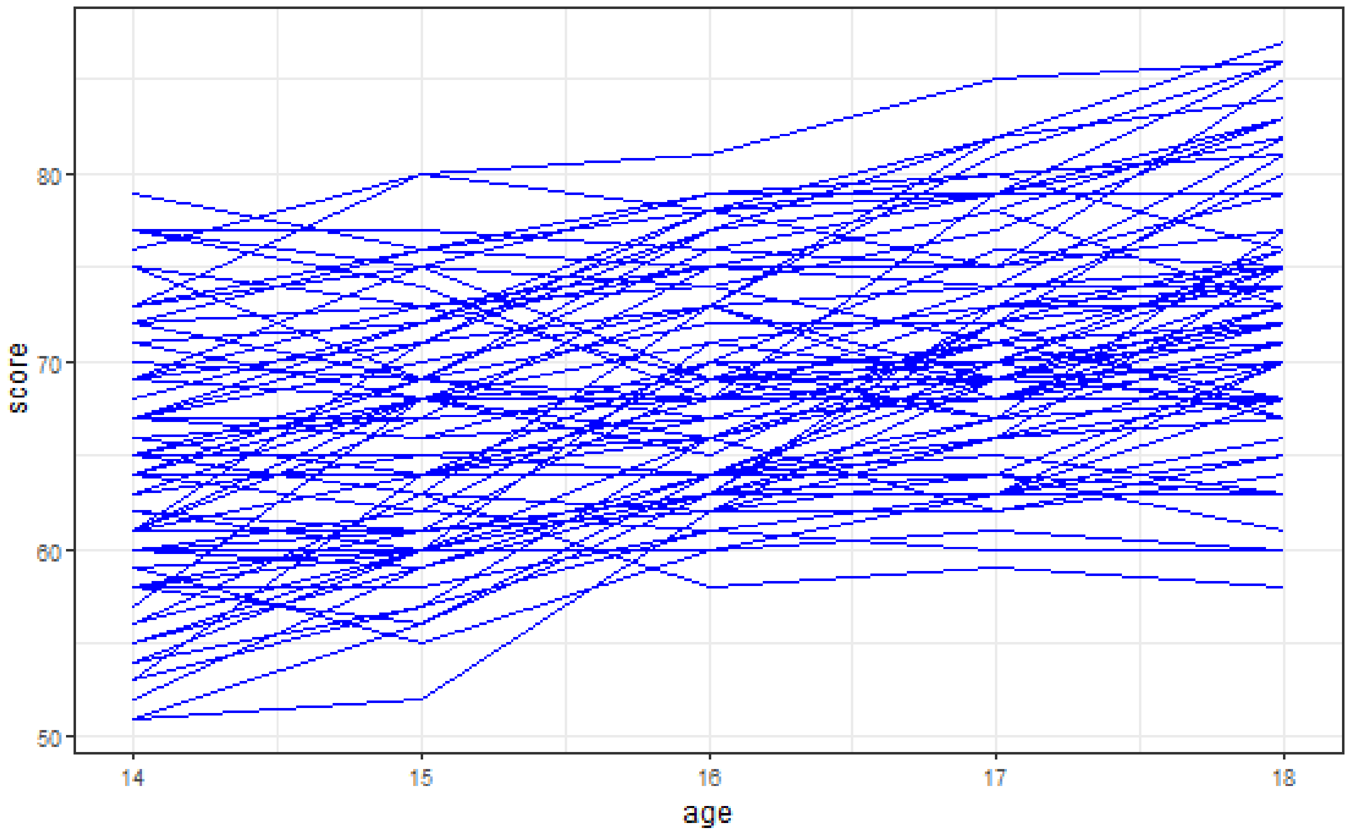
```
Variable    | Mean |    SD | IQR |          Range | Skewness | Kurtosis |   n | n_Missing
-----------------------------------------------------------------------------------------
teacher     |  2.00 |  0.82 |   2 |   [1.00, 3.00] |     0.00 |    -1.50 | 375 |         0
student     | 38.00 | 21.68 |  38 |  [1.00, 75.00] |     0.00 |    -1.20 | 375 |         0
age         | 16.00 |  1.42 |   2 | [14.00, 18.00] |     0.00 |    -1.30 | 375 |         0
cage        |  0.00 |  1.42 |   2 |  [-2.00, 2.00] |     0.00 |    -1.30 | 375 |         0
sex.Female  |  0.51 |  0.50 |   1 |   [0.00, 1.00] |    -0.03 |    -2.01 | 375 |         0
sex. Male   |  0.49 |  0.50 |   1 |   [0.00, 1.00] |     0.03 |    -2.01 | 375 |         0
score       | 68.71 |  7.20 |  10 | [51.00, 87.00] |     0.14 |    -0.28 | 375 |         0
pass        |  0.39 |  0.49 |   1 |   [0.00, 1.00] |     0.46 |    -1.80 | 375 |         0
experience  |  6.33 |  1.25 |   3 |   [5.00, 8.00] |     0.38 |    -1.50 | 375 |         0
zero        |  0.00 |  0.00 |   0 |   [0.00, 0.00] |          |          | 375 |         0
```

Hierarchical data: temporality

```
test |> datawizard::to_numeric() |>
  datawizard::describe_distribution()
```

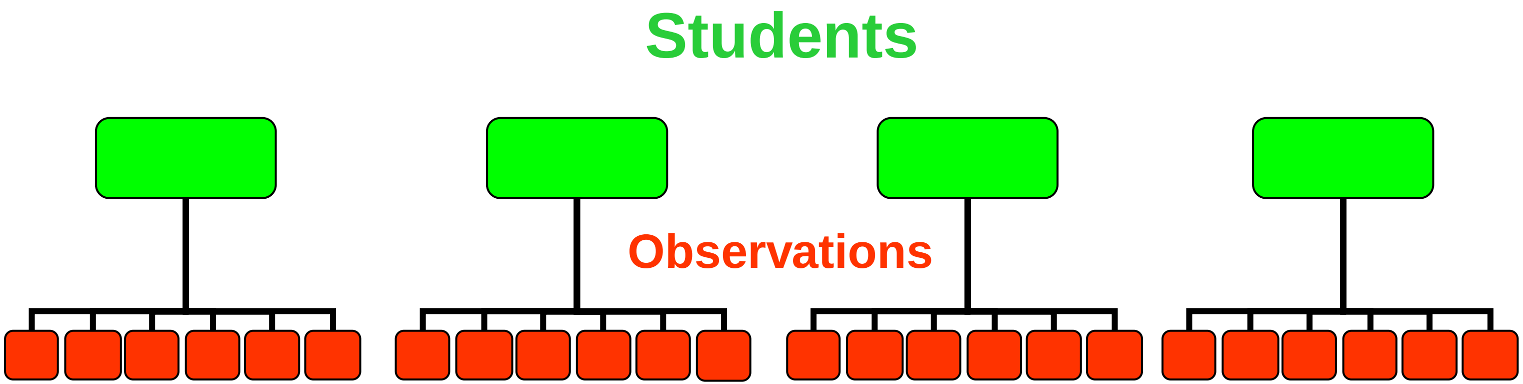
Variable	Mean	SD	IQR	Range	Skewness	Kurtosis	n	n_Missing
teacher	2.00	0.82	2	[1.00, 3.00]	0.00	-1.50	375	0
student	38.00	21.68	38	[1.00, 75.00]	0.00	-1.20	375	0
age	16.00	1.42	2	[14.00, 18.00]	0.00	-1.30	375	0
cage	0.00	1.42	2	[-2.00, 2.00]	0.00	-1.30	375	0
sex.Female	0.51	0.50	1	[0.00, 1.00]	-0.03	-2.01	375	0
sex. Male	0.49	0.50	1	[0.00, 1.00]	0.03	-2.01	375	0
score	68.71	7.20	10	[51.00, 87.00]	0.14	-0.28	375	0
pass	0.39	0.49	1	[0.00, 1.00]	0.46	-1.80	375	0
experience	6.33	1.25	3	[5.00, 8.00]	0.38	-1.50	375	0
zero	0.00	0.00	0	[0.00, 0.00]			375	0

Hierarchical data: temporality



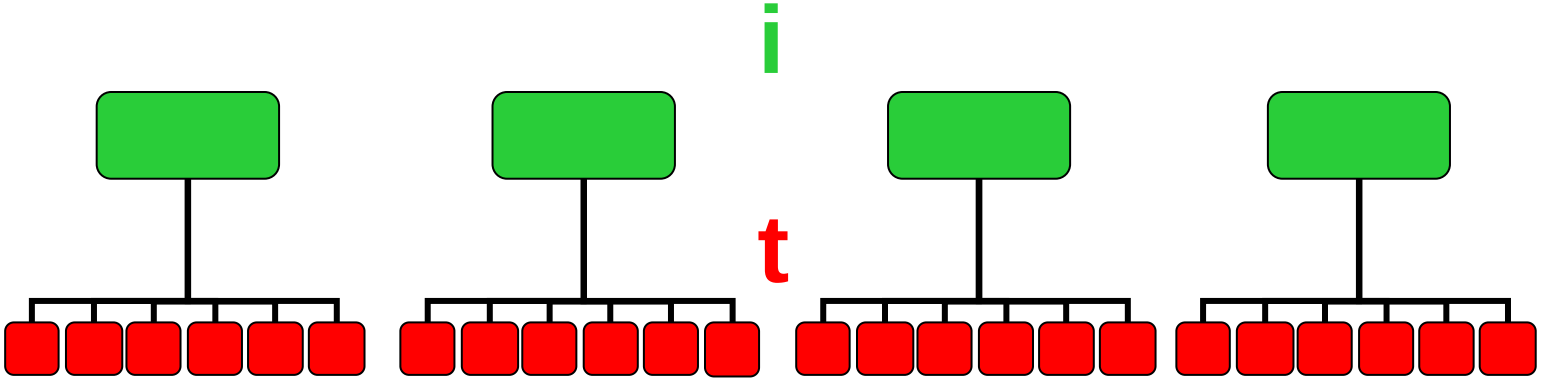
Hierarchical data: temporality

- Repeated observations within student are almost certainly correlated
- A two-level model:



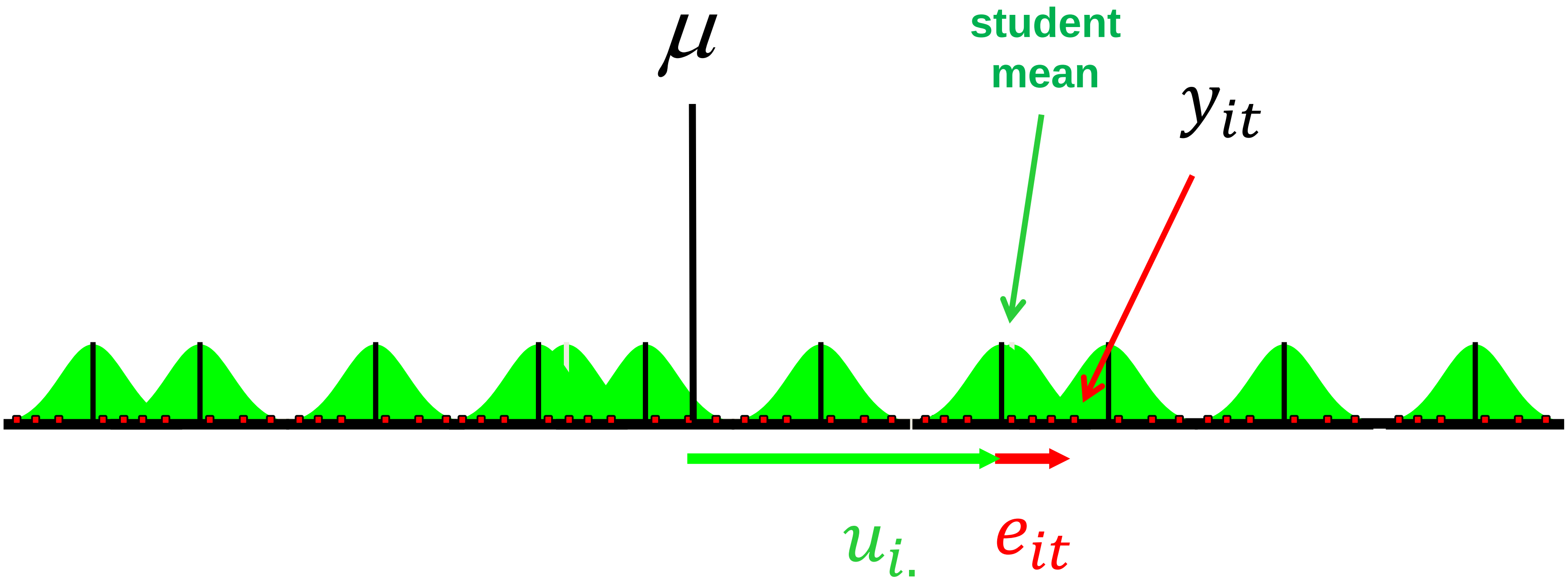
Hierarchical data: temporality

- Repeated observations within student are almost certainly correlated
- A two-level model:



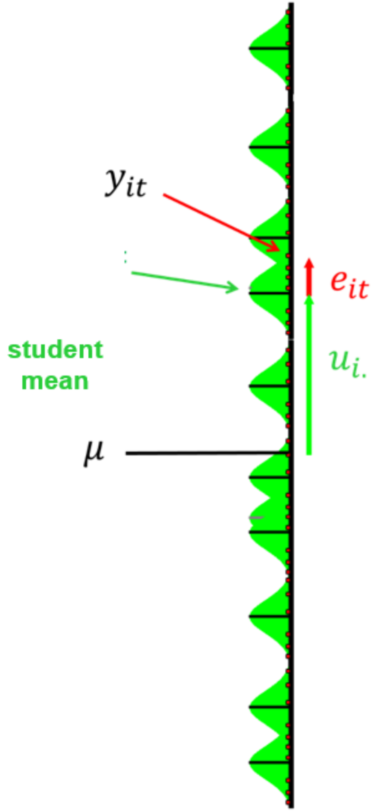
$$y_{it} = y_{\text{student},\text{observation}}$$

Hierarchical data: temporality



$$y_{it} = \mu + u_{i.} + e_{it}$$

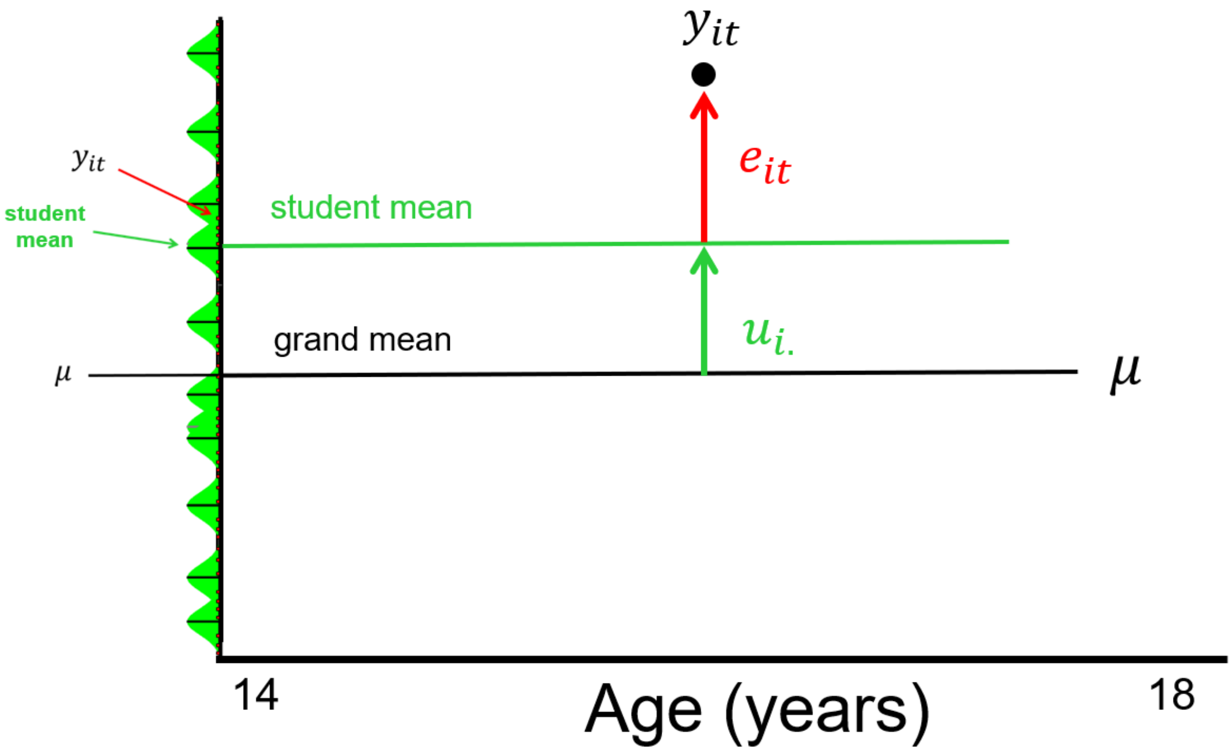
Hierarchical data: temporality



$$y_{it} = \mu + u_{i.} + e_{it}$$

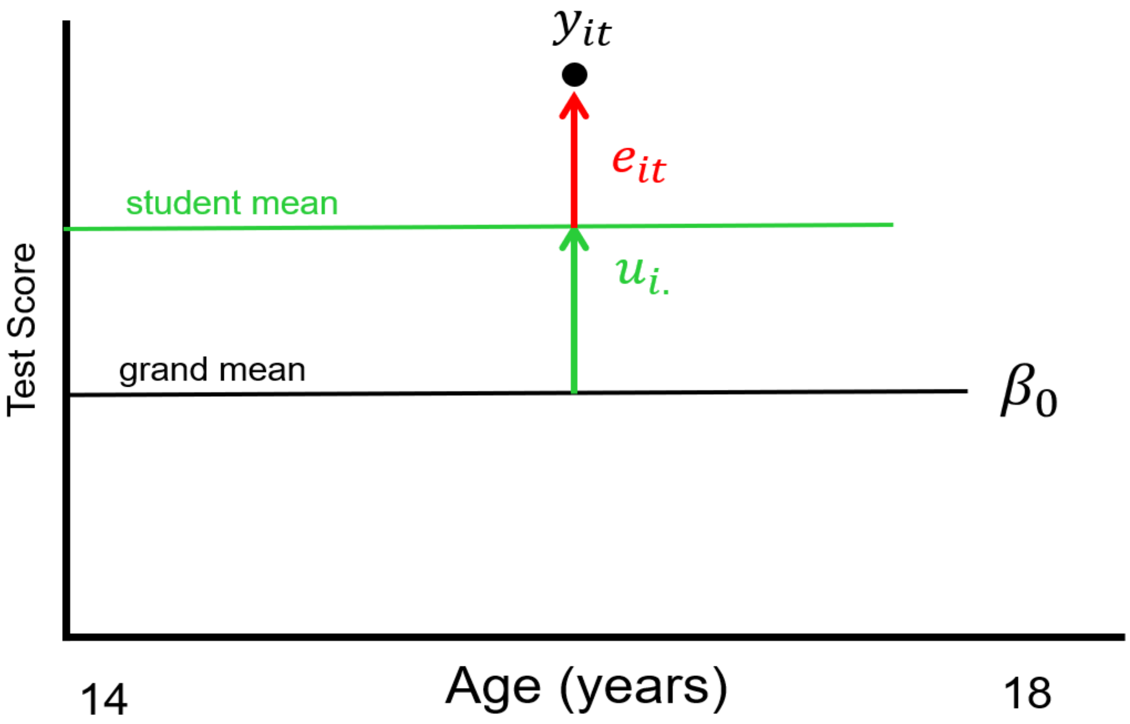
Hierarchical data: temporality

$$y_{it} = \mu + u_{i.} + e_{it}$$



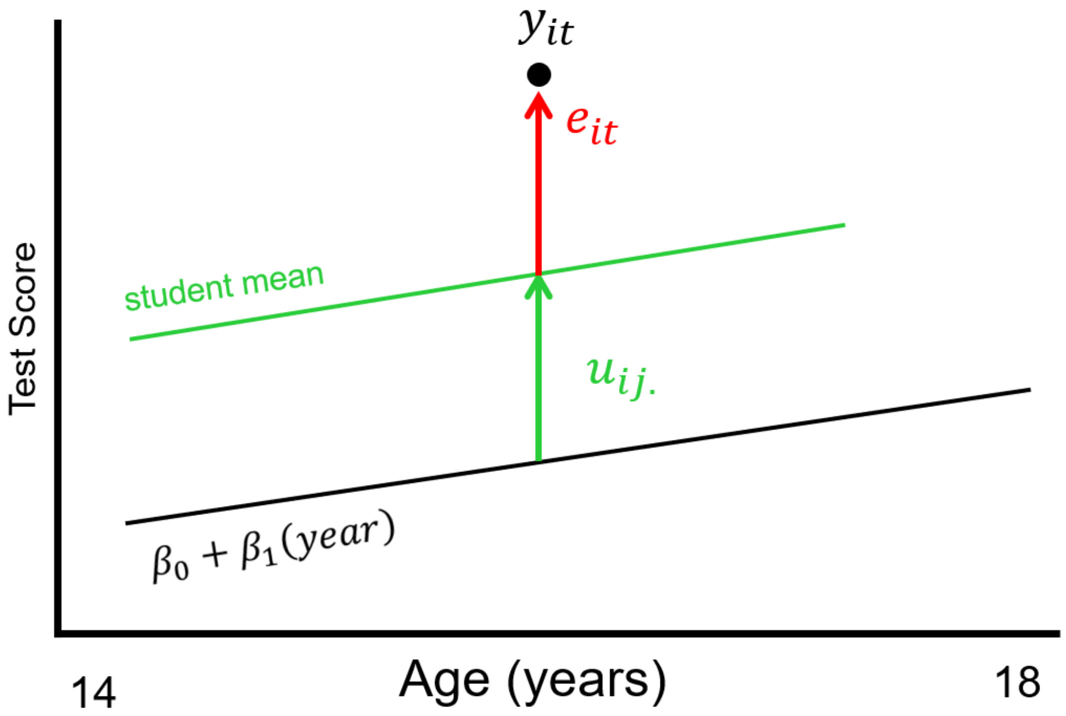
Hierarchical data: temporality

$$y_{it} = \beta_0 + u_{i.} + e_{it}$$



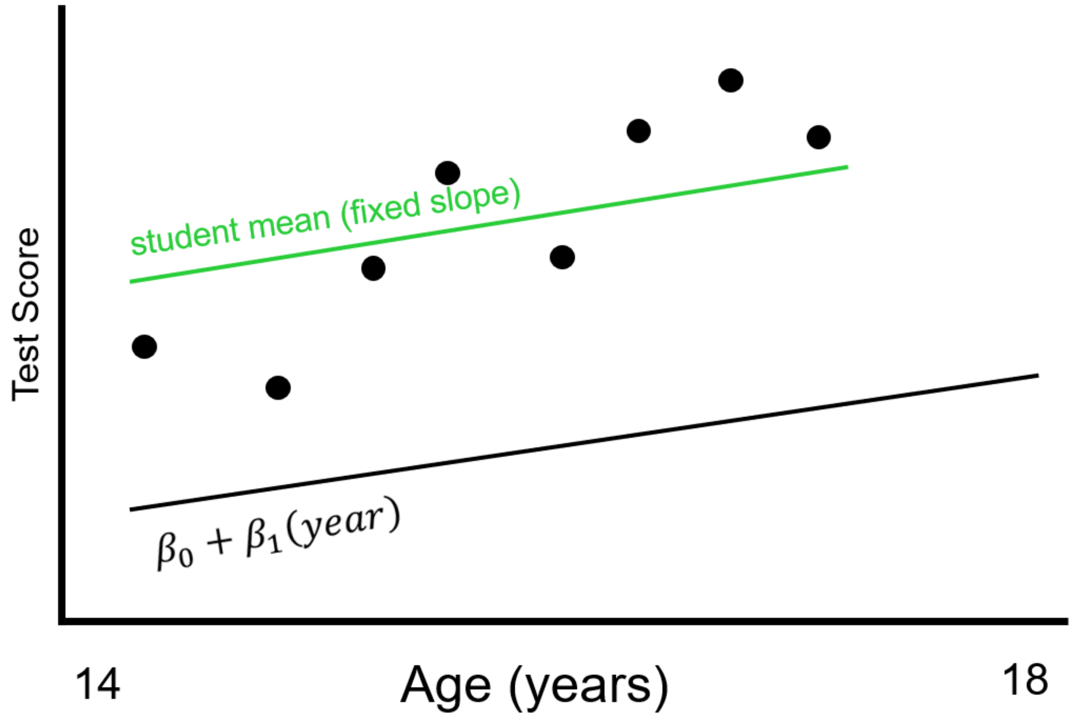
Hierarchical data: temporality

$$y_{it} = \beta_0 + \beta_1(age) + u_{i.} + e_{it}$$



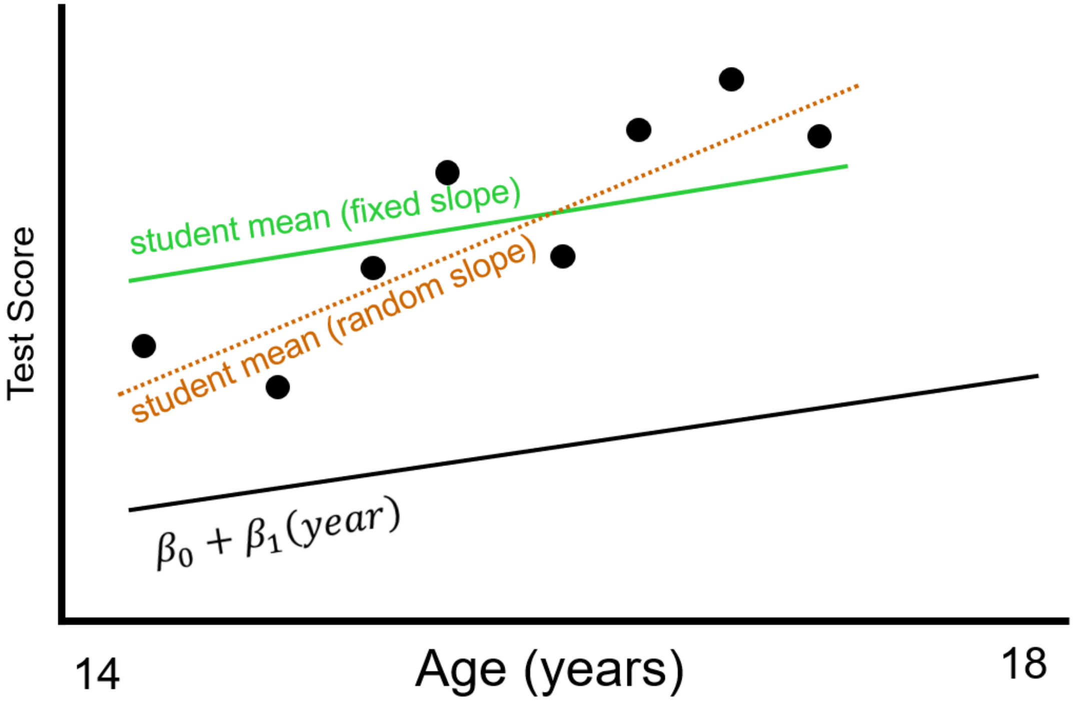
Hierarchical data: temporality

$$y_{it} = \beta_0 + \beta_1(age) + u_{i.} + e_{it}$$



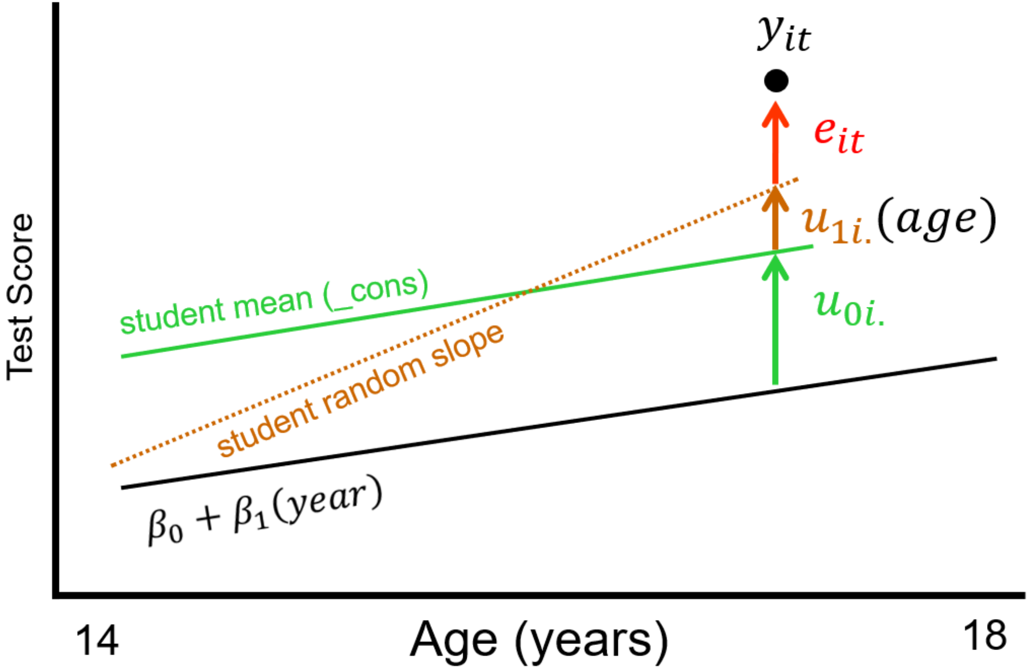
Hierarchical data: temporality

$$y_{it} = \beta_0 + \beta_1(age) + u_{0i.} + u_{1i.}(age) + e_{it}$$



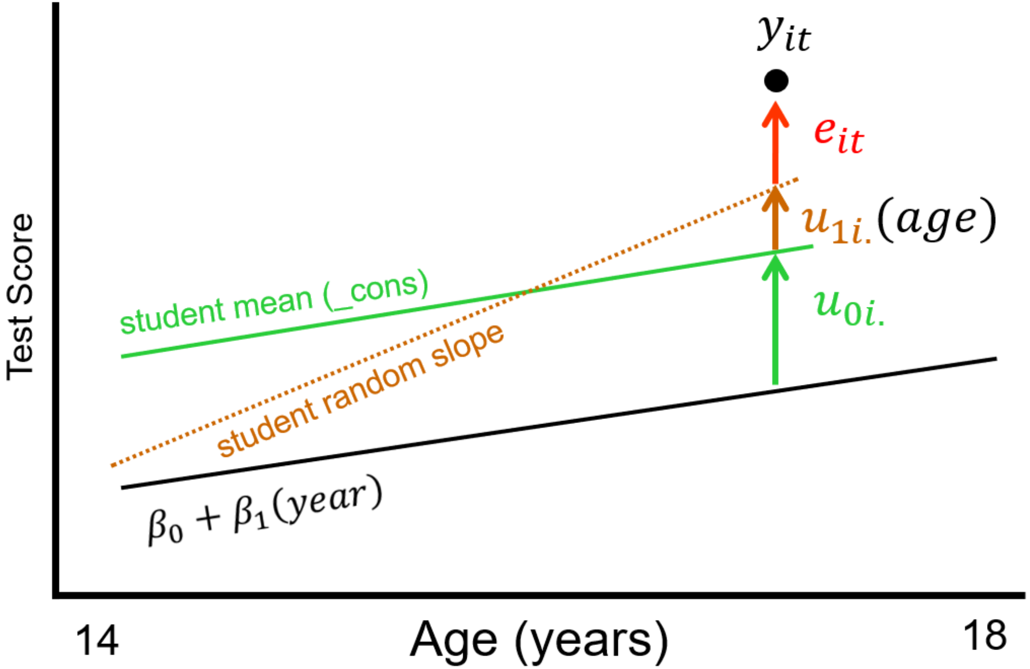
Hierarchical data: temporality

$$y_{ijk} = \beta_0 + \beta_1(age) + u_{0i.} + u_{1i.}(age) + e_{it}$$

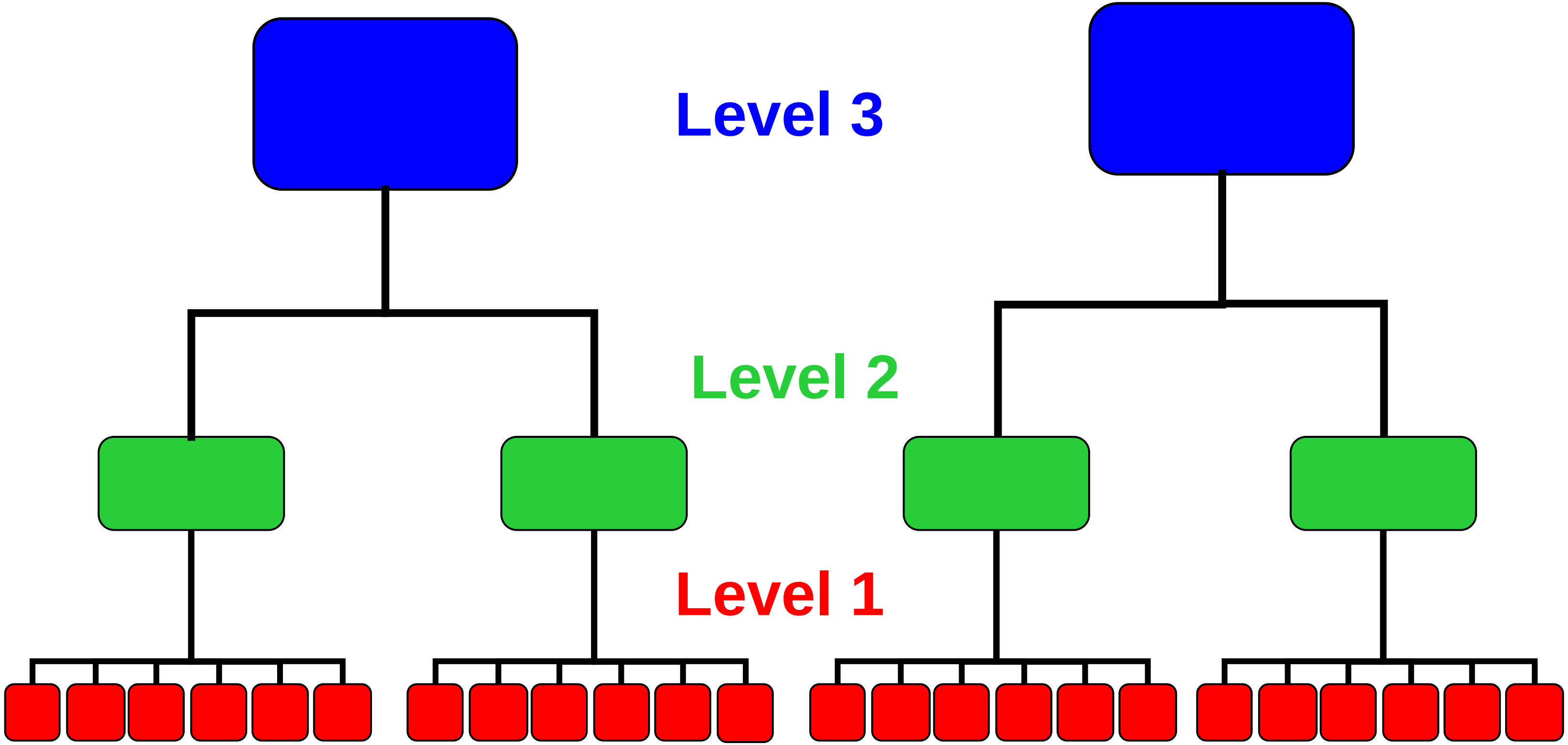


Hierarchical data: temporality

$$y_{ijk} = \beta_0 + \beta_1(age) + u_{0i.} + u_{1i.}(age) + e_{it}$$

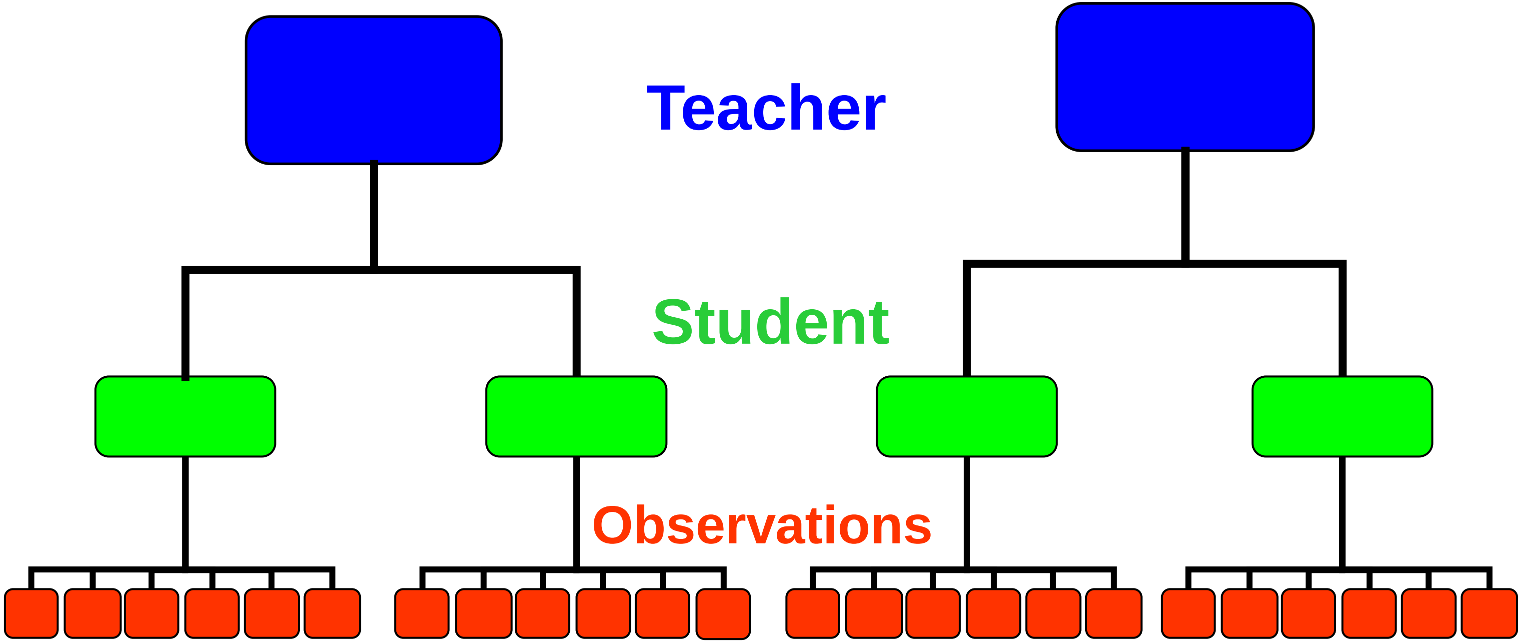


Hierarchical data: three-level models



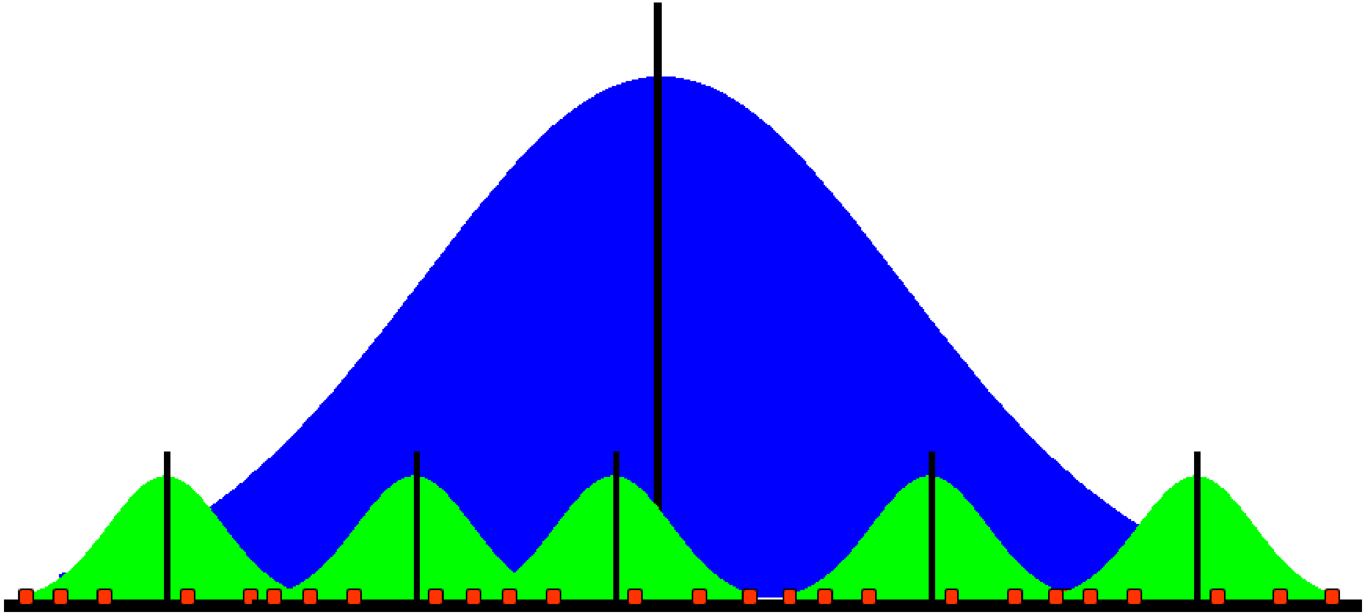
Hierarchical data: three-level models

- Repeated observations ***from students in the same teacher's classroom*** may be correlated



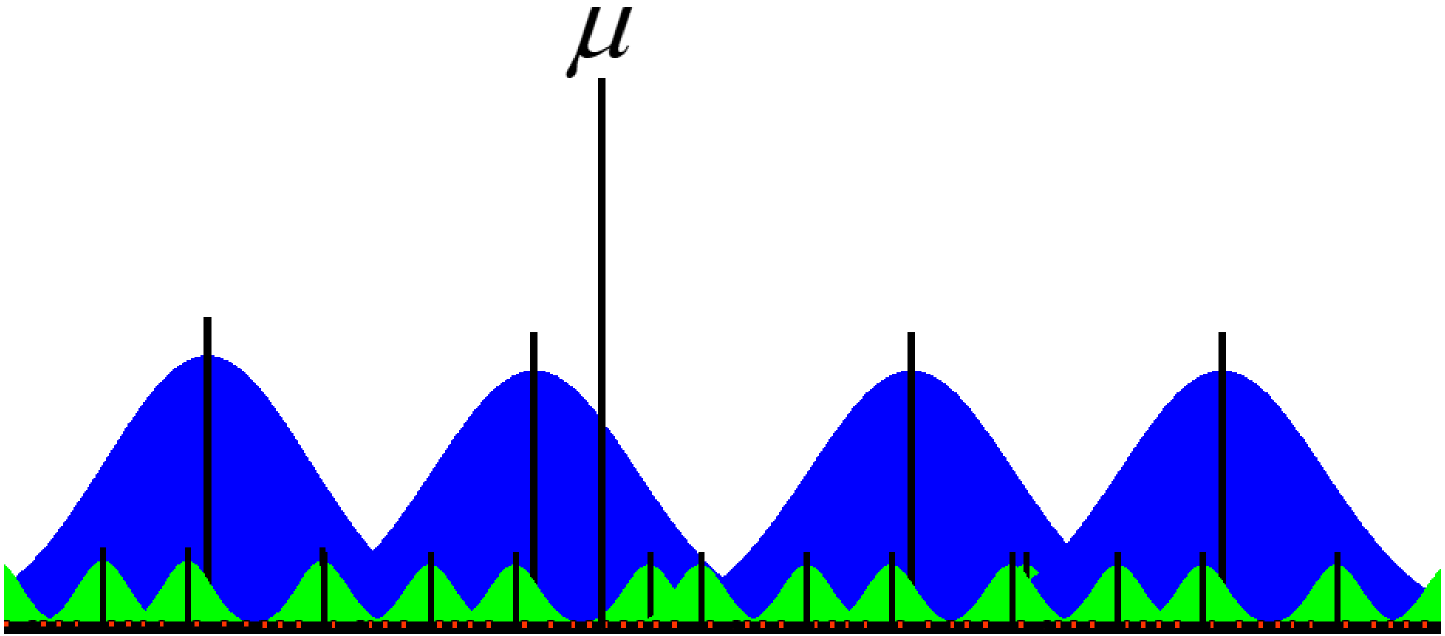
Hierarchical data: three-level models

- Students' means in the same teacher's classroom will vary about their teacher's mean

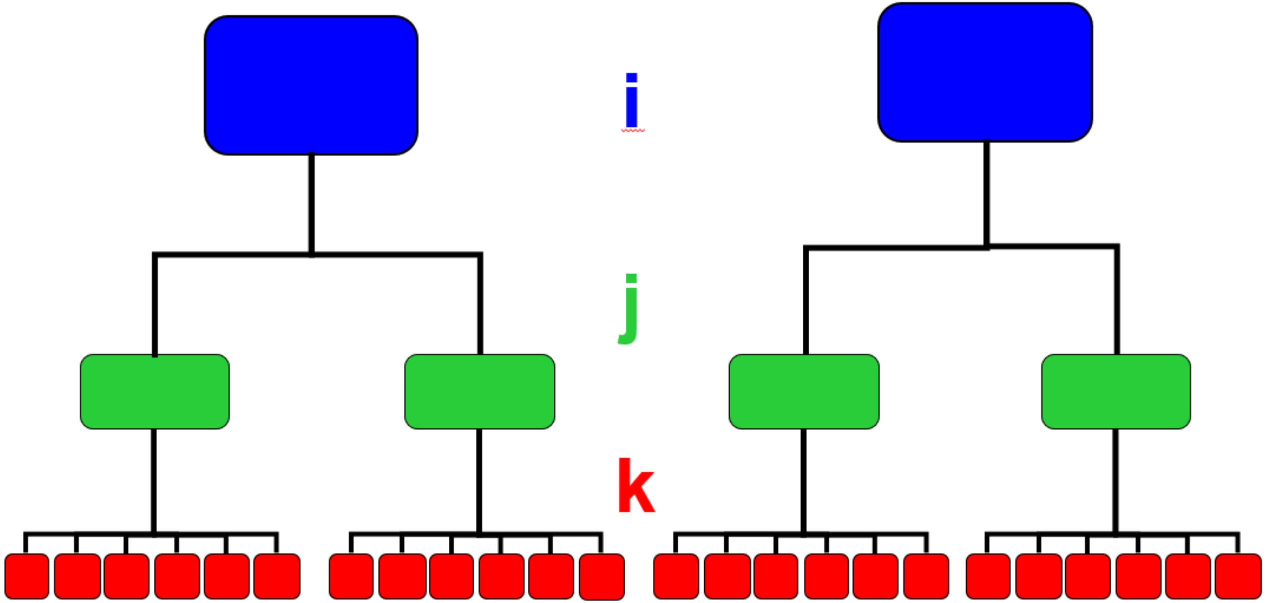


Hierarchical data: three-level models

- Teacher means will vary about the grand mean

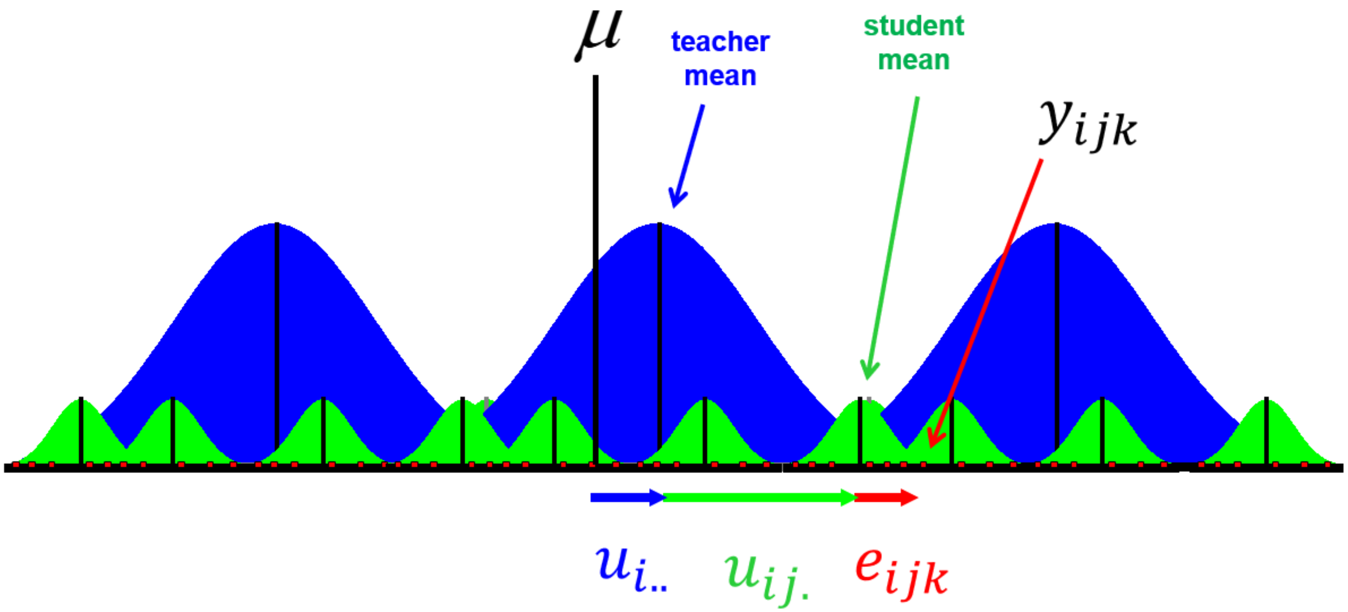


Hierarchical data: three-level models



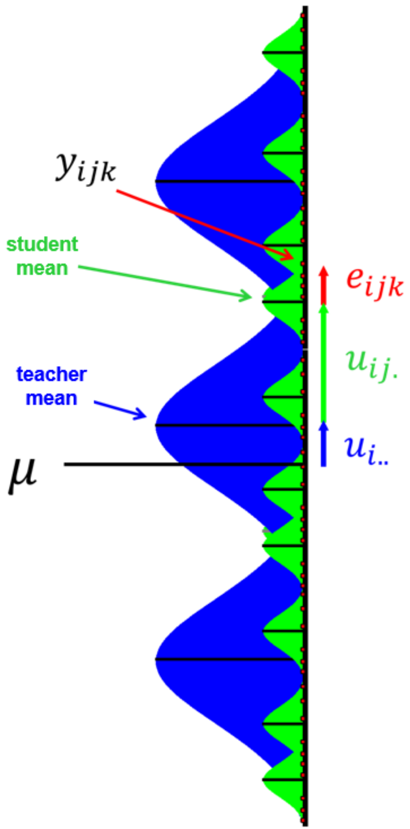
$$y_{ijk} = y_{teacher,student,observation}$$

Hierarchical data: three-level models



$$y_{ijk} = \mu + u_{i..} + u_{ij.} + e_{ijk}$$

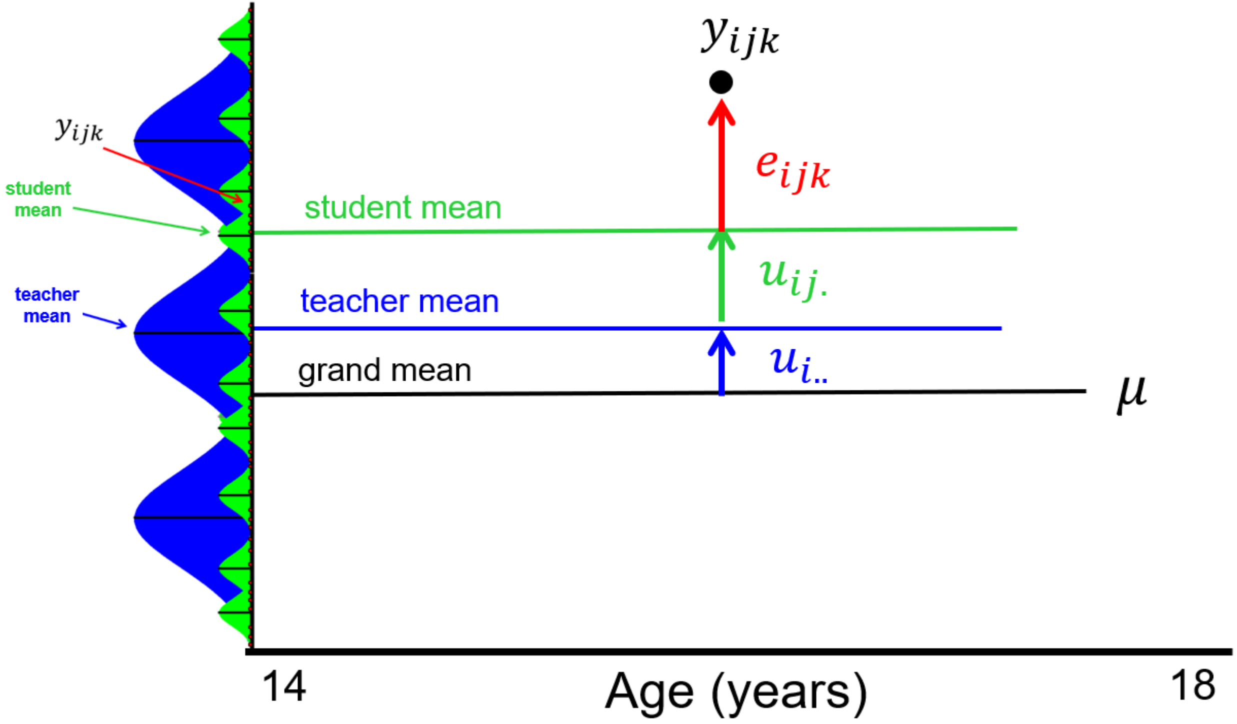
Hierarchical data: three-level models



$$y_{ijk} = \mu + u_{i..} + u_{ij.} + e_{ijk}$$

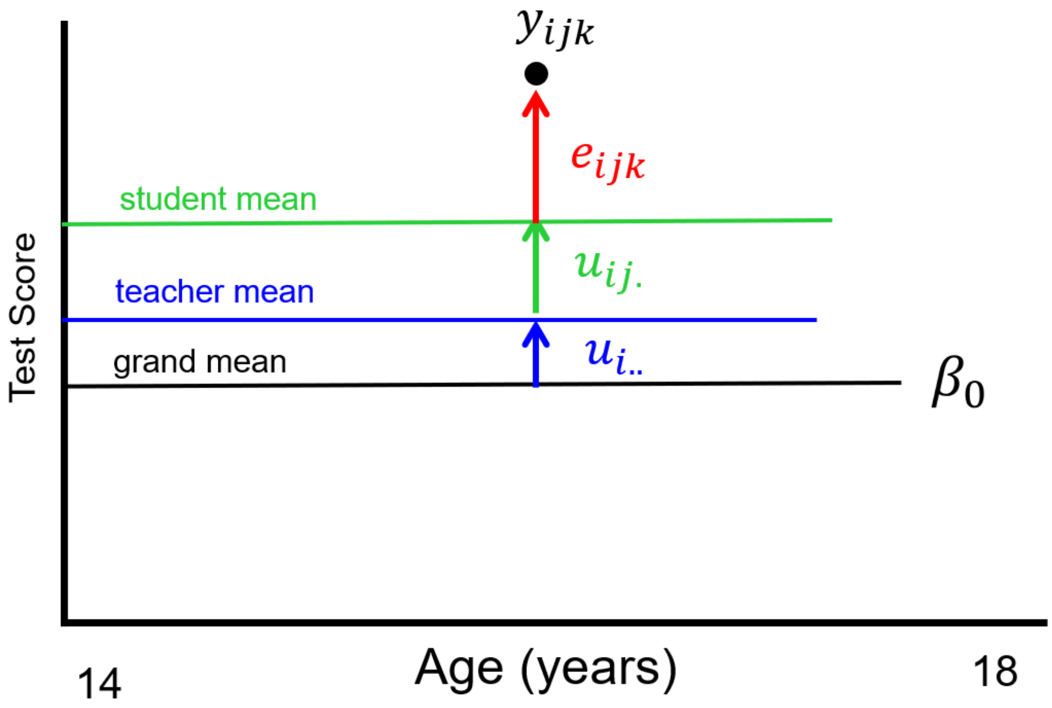
Hierarchical data: three-level models

$$y_{ijk} = \mu + u_{i..} + u_{ij.} + e_{ijk}$$



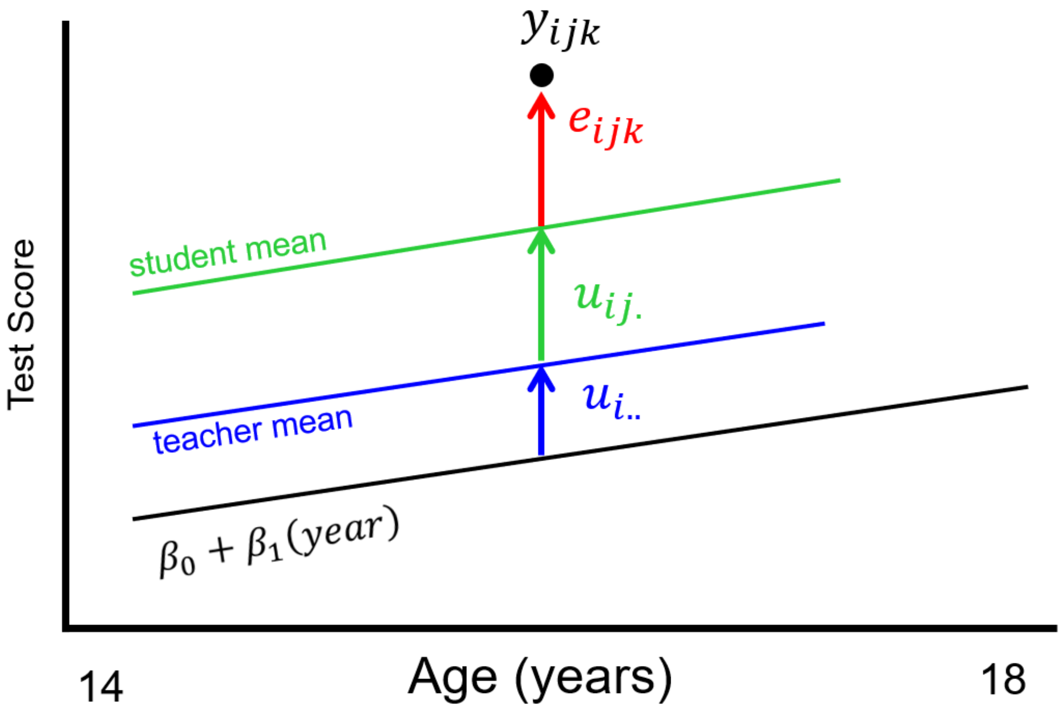
Hierarchical data: three-level models

$$y_{ijk} = \beta_0 + u_{i..} + u_{ij.} + e_{ijk}$$



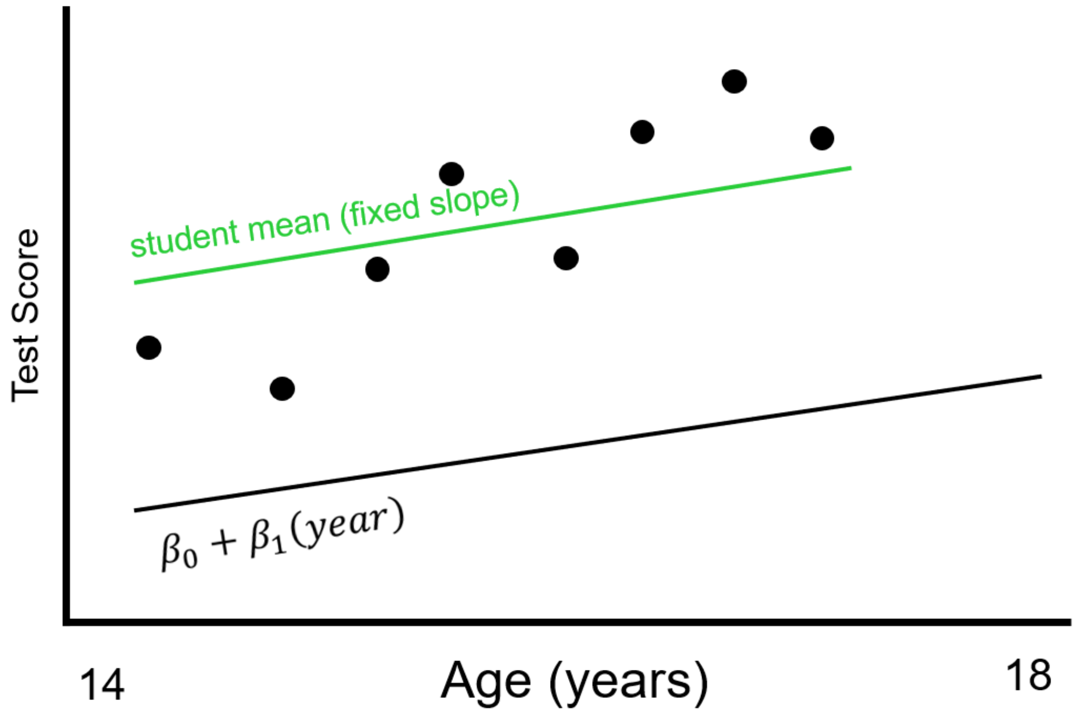
Hierarchical data: three-level models

$$y_{ijk} = \beta_0 + \beta_1(age) + u_{i..} + u_{ij.} + e_{ijk}$$



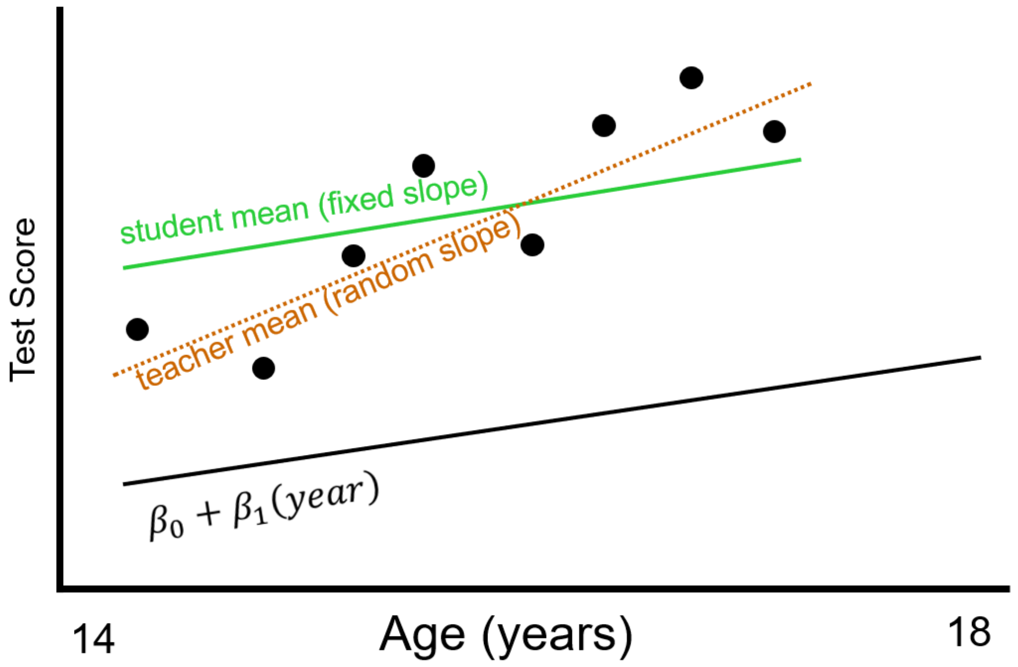
Hierarchical data: three-level models

$$y_{ijk} = \beta_0 + \beta_1(age) + \color{blue}{u_{i..}} + \color{green}{u_{ij.}} + \color{red}{e_{ijk}}$$



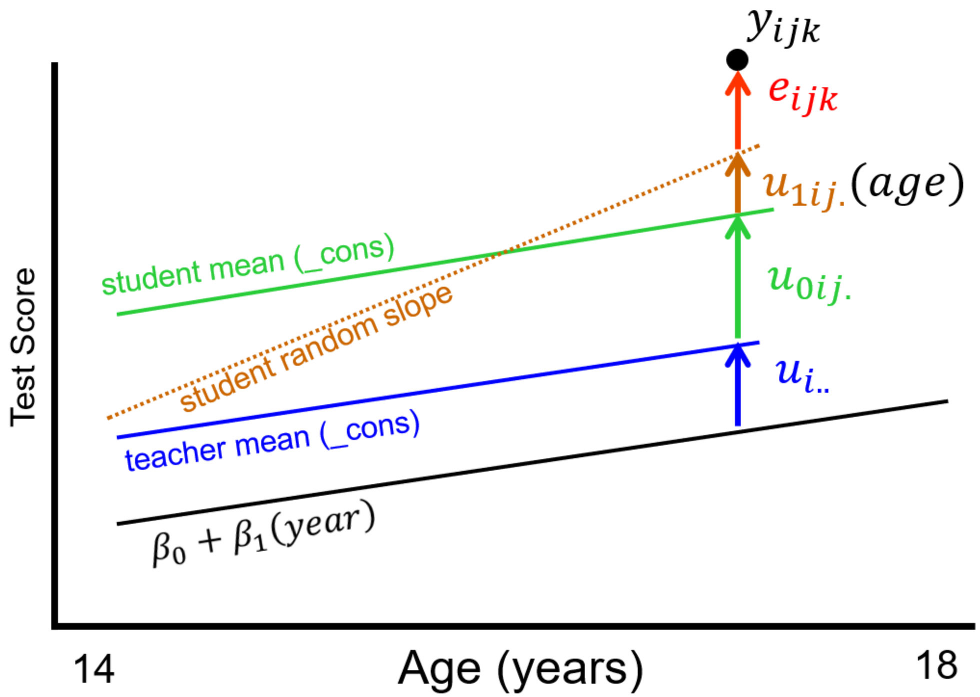
Hierarchical data: three-level models

$$y_{ijk} = \beta_0 + \beta_1(age) + u_{i..} + u_{0ij.} + u_{1ij.}(age) + e_{ijk}$$



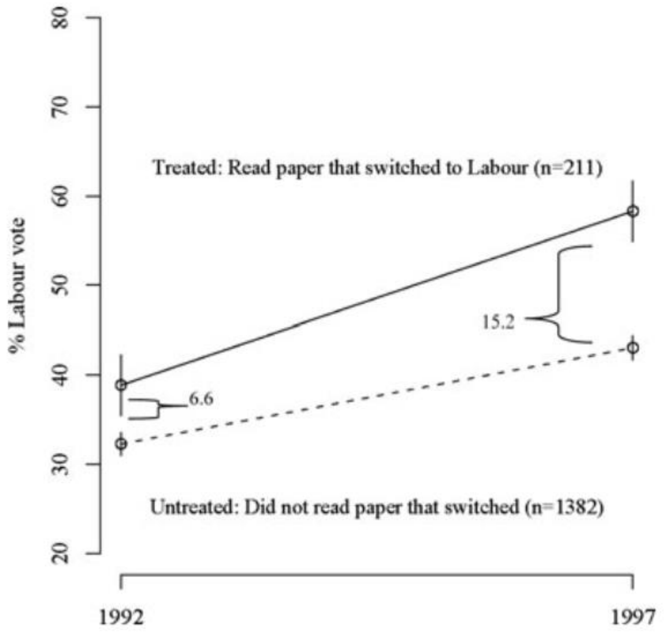
Hierarchical data: three-level models

$$y_{ijk} = \beta_0 + \beta_1(age) + u_{i..} + u_{0ij.} + u_{1ij.}(age) + e_{ijk}$$



Other "temporalities"

- We have already seen how *time* can enter a model. Example:
- Ladd, Jonathan McDonald, and Gabriel S. Lenz. 2009. "Exploiting a Rare Communication Shift to Document the Persuasive Power of the News Media." American Journal of Political Science 53(2): 394–410.



This figure shows that reading a paper that switched to Labour is associated with an (15.2 − 6.6 =) 8.6 percentage point shift to Labour between the 1992 and 1997 UK elections. Paper readership is measured in the 1996 wave, before the papers switched, or, if no 1996 interview was conducted, in an earlier wave. Confidence intervals show one standard error.

Other "temporalities"

- We have already seen how *time* can enter a model. Example:
- Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. 2017. "Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial." International Journal of Epidemiology 46(1): 348–55.

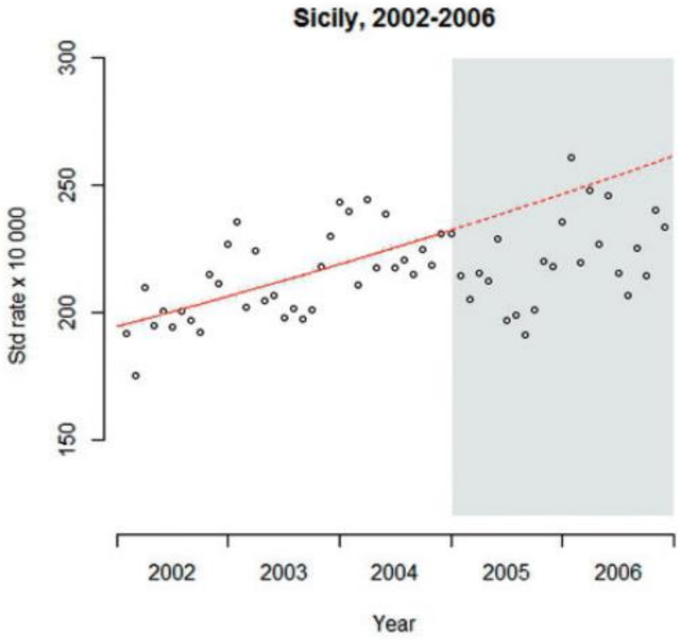
Year	Month	Time elapsed (T)	Smoking ban[a] (X)	ACEs (Y)	Std popn
2004	1	25	0	914	381656.3
2004	2	26	0	808	383680
2004	3	27	0	937	383504.2
2004	4	28	0	840	386462.9
2004	5	29	0	916	383783.1
2004	6	30	0	828	380836.8
2004	7	31	0	845	383483
2004	8	32	0	818	380906.2
2004	9	33	0	860	382926.8
2004	10	34	0	839	384052.4
2004	11	35	0	887	384449.6
2004	12	36	0	886	383428.4
2005	1	37	1	831	388153.2
2005	2	38	1	796	388373.2
2005	3	39	1	833	386470.1
2005	4	40	1	820	386033.2
2005	5	41	1	877	383686.4
2005	6	42	1	758	385509.3
2005	7	43	1	767	385901.9
2005	8	44	1	738	386516.6
2005	9	45	1	781	388436.5
2005	10	46	1	843	383255.2
2005	11	47	1	850	390148.7
2005	12	48	1	908	385874.9

ACEs, hospital admissions for acute coronary event; Std popn, age-standardized population in person-years.[16]

[a]Smoking ban: 0, smoking ban not in place; 1, smoking ban in place.

Other “temporalities”

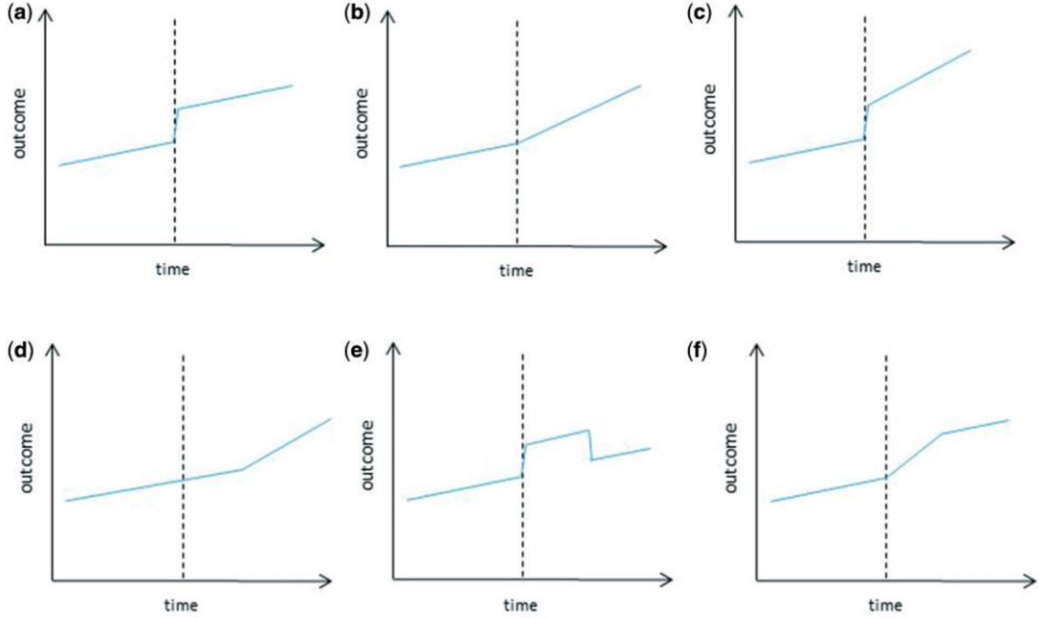
- We have already seen how *time* can enter a model. Example:
- Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. 2017. “Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial.” International Journal of Epidemiology 46(1): 348–55.



Standardized (Std) rate of ACE (acute coronary events) over time. White background, pre-intervention period; grey background, post-intervention period; continuous line, pre-intervention trend; dashed line, counterfactual scenario

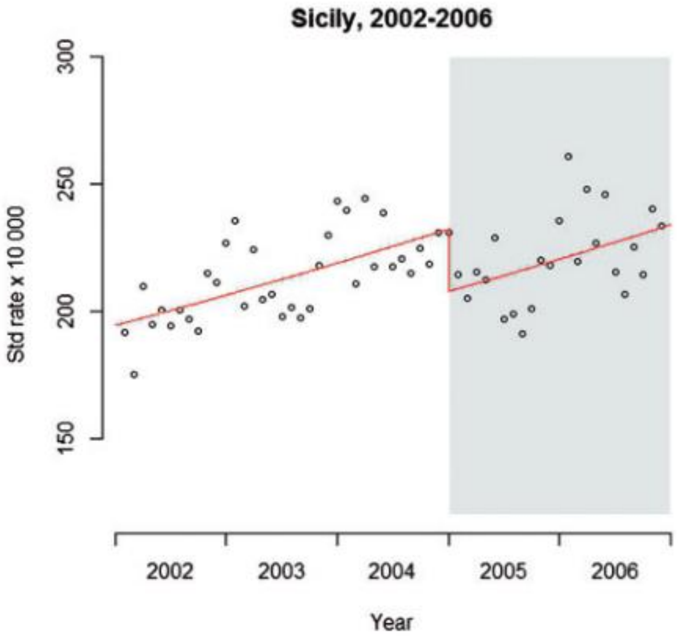
Other “temporalities”

- We have already seen how *time* can enter a model. Example:
- Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. 2017. “Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial.” International Journal of Epidemiology 46(1): 348–55.



Other “temporalities”

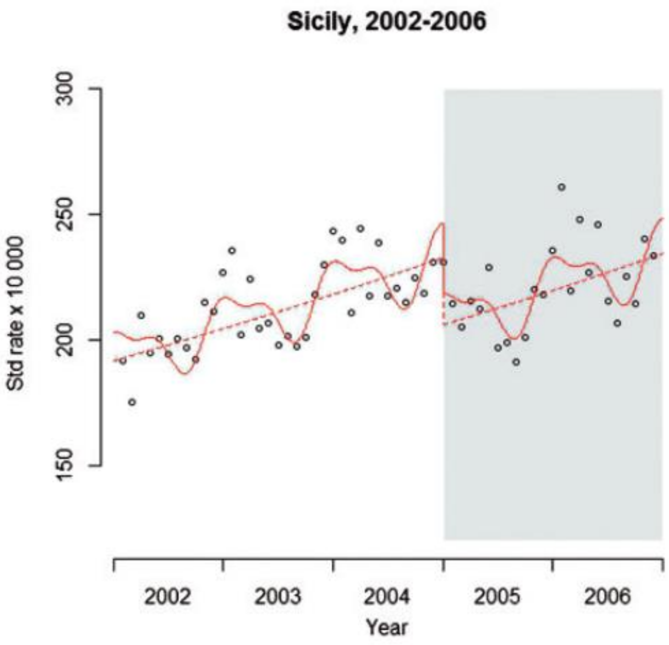
- We have already seen how *time* can enter a model. Example:
- Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. 2017. “Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial.” International Journal of Epidemiology 46(1): 348–55.



Interrupted time series with level change regression model. Line: predicted trend based on the unadjusted regression model

Other "temporalities"

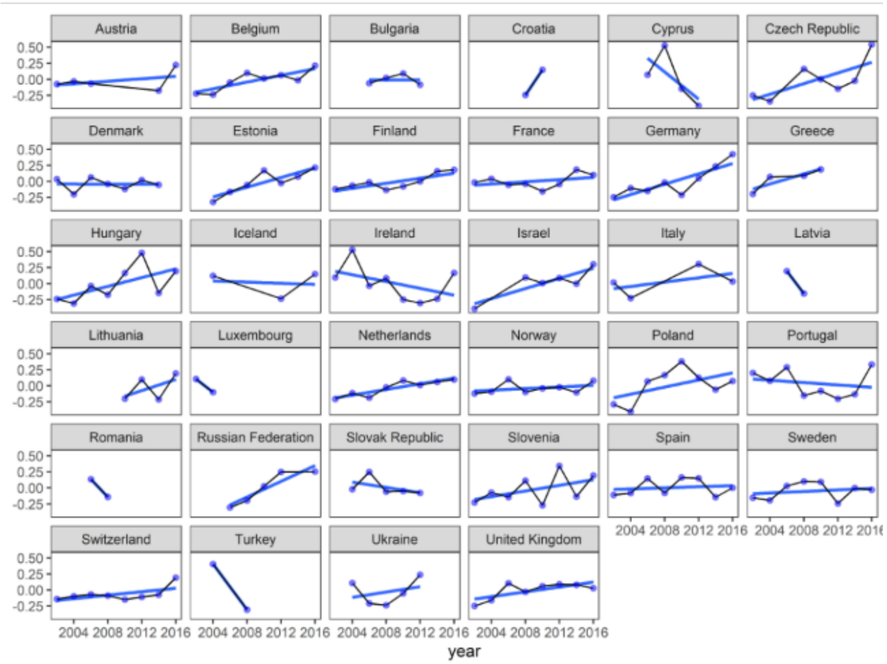
- We have already seen how *time* can enter a model. Example:
- Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. 2017. "Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial." International Journal of Epidemiology 46(1): 348–55.



Model adjusted for seasonality. Solid line: predicted trend based on the seasonally adjusted regression model. Dashed line: deseasonalized trend

Other "temporalities"

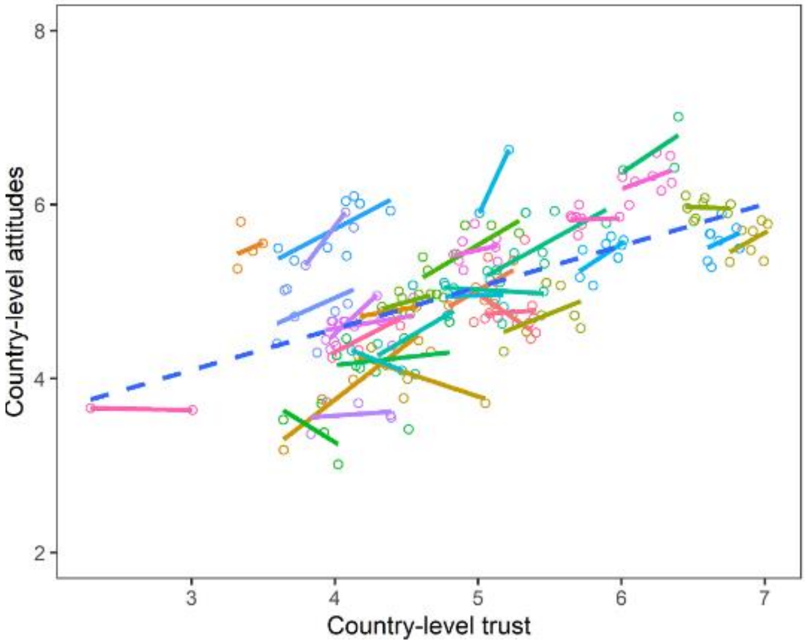
- We have already seen how *time* can enter a model. Example:
- Mitchell, Jeffrey. 2021. "Social Trust and Anti-Immigrant Attitudes in Europe: A Longitudinal Multi-Level Analysis." Frontiers in Sociology 6.



Within-country social trust over time

Other “temporalities”

- We have already seen how *time* can enter a model. Example:
- Mitchell, Jeffrey. 2021. “Social Trust and Anti-Immigrant Attitudes in Europe: A Longitudinal Multi-Level Analysis.” Frontiers in Sociology 6.



Scatterplot of country-year attitudes about immigrants and social trust

Modelling time hierarchies and time-series in R

- Several packages:
  - lme4
  - panelr
  - plm
  - pglm
- In the labs:
  - lme4
  - plm