UCLA School of Education & Information Studies
Department of Education

EDUC 231D
Advanced Quantitative Methods: Multilevel Analysis
Winter 2025

Use of Multilevel Models for Longitudinal Analysis

Lecture 11 Presentation Slides

February 13, 2025

Today's Topics

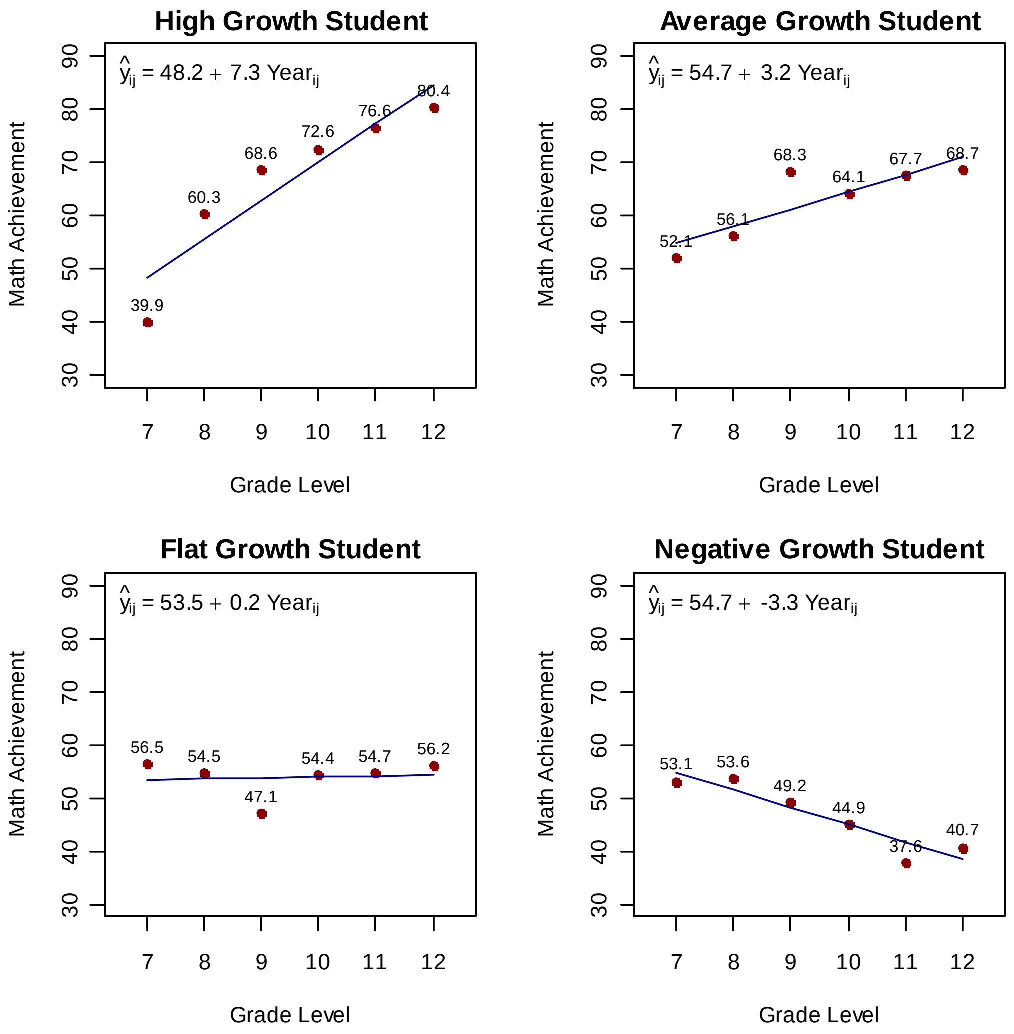
- Introduction to growth modeling
- Multilevel model to characterize change
- Multilevel model to study correlates of change

Introduction to Growth Modeling

The study of change

- Longitudinal analysis (or growth modeling) examines how repeated observations/outcomes collected on individuals change over time
- Interest in how growth trajectories can vary across individuals

LSAY data for 7th grade through 12th grade math test scores



The study of change

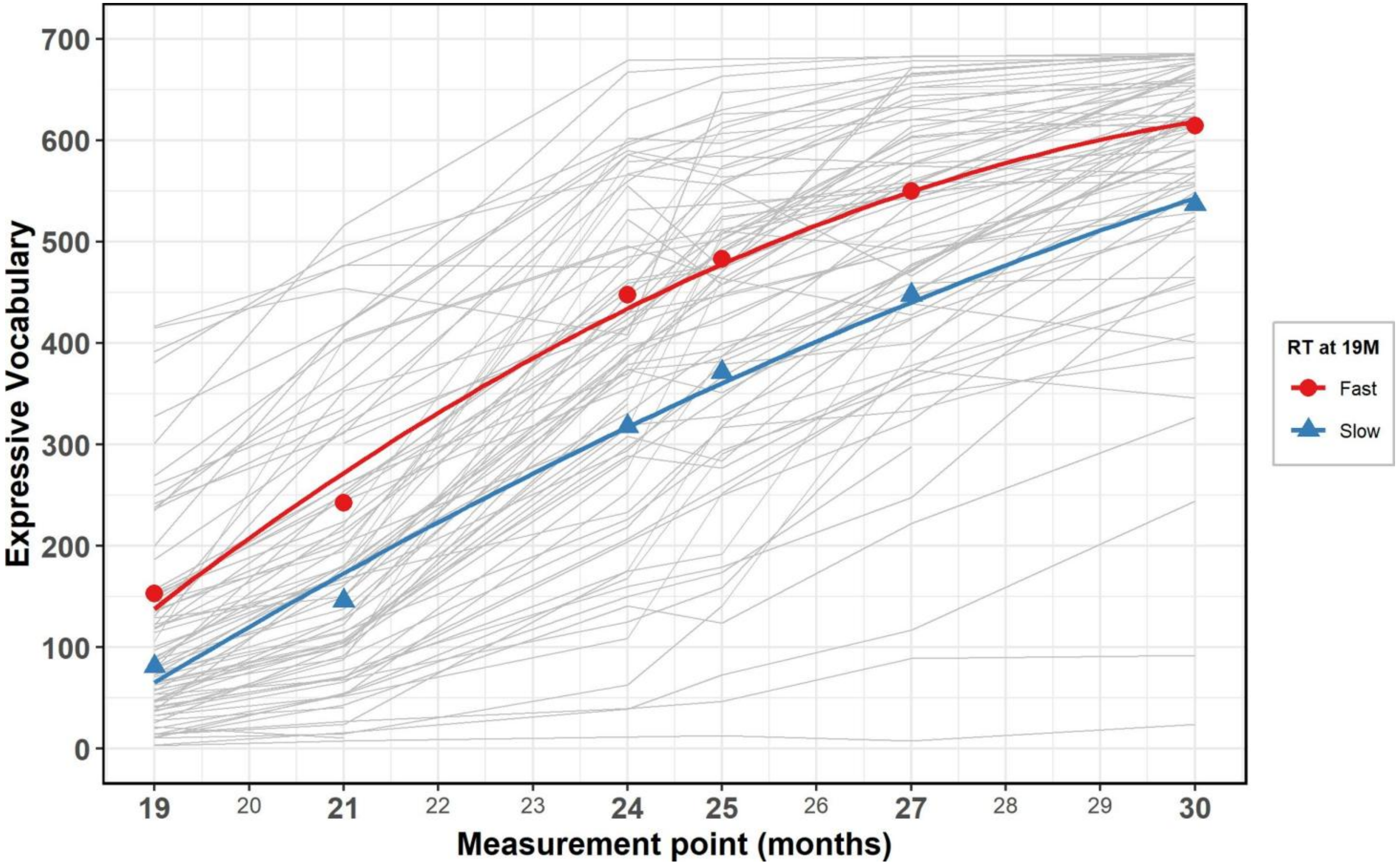


Image source: Peter, M. S., Durrant, S., Jessop, A., Bidgood, A., Pine, J. M., & Rowland, C. F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive Psychology*, *115*, 101238. https://doi.org/10.1016/j.cogpsych.2019.101238

The study of change

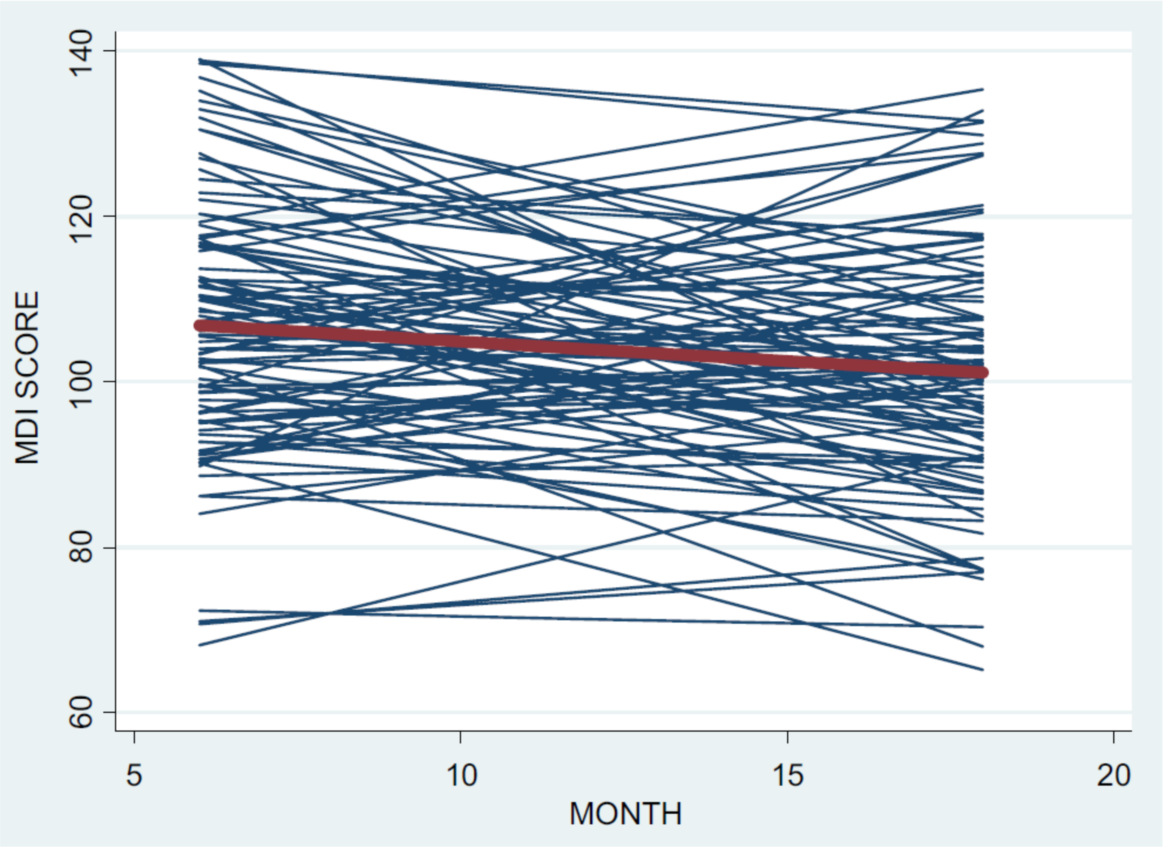
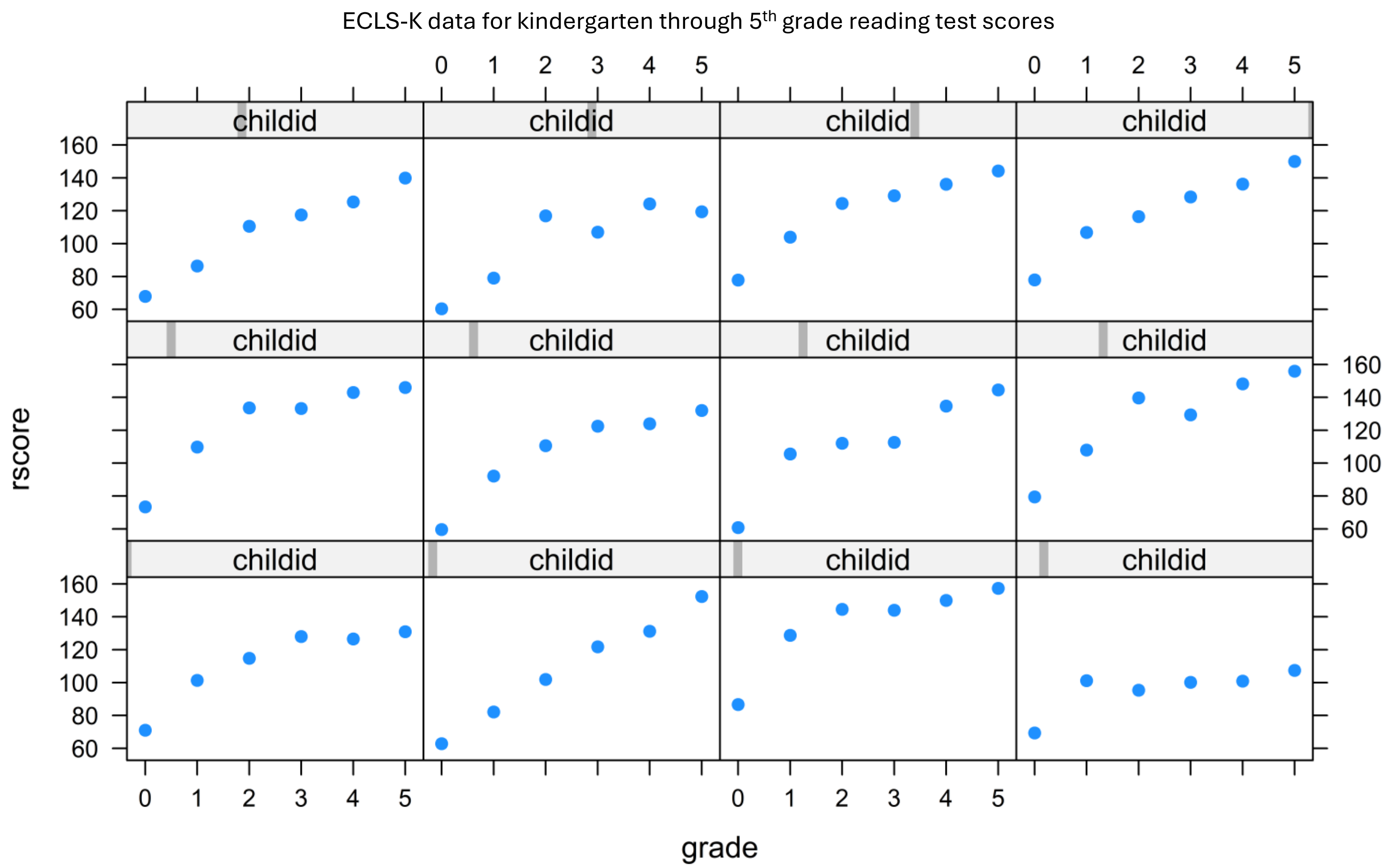


Image source: Gee, K. A. (2014). Multilevel Growth Modeling: An Introductory Approach to Analyzing Longitudinal Data for Evaluators. *American Journal of Evaluation*, *35*(4), 543–561. https://doi.org/10.1177/1098214014523823

The study of change



The study of change

- A long tradition of studying change by looking at the difference between two time points (e.g., pre-post tests or start-end points of time)
- Emphasis is typically on the overall average amount of change between the two time points: $\bar{Y}_2 - \bar{Y}_1$
  - Masks individual variation in change (though can look at individual change scores: $Y_{2i} - Y_{1i}$)
  - Not able to capture non-linearities in change over time
  - Not able to disentangle measurement error from true change (unless additional efforts are taken to account for standard errors of measurement)

The study of change

- Can overcome some limitations of two-time-point designs with data covering more than two time points and growth modeling
  - Characterize the shape/pattern of individual growth
  - Summarize what the growth process looks like on average
  - Estimate the extent to which individuals vary from the average growth pattern
  - Examine what factors/characteristics are associated with growth patterns

Data structure for growth modeling

- Data from multiple time points are nested within individuals
  - Time *variant* data = measures with values that can change over time
  - Time *invariant* data = measures with values that are fixed across time
- When data are collected on the same individuals over time, the data file can come in a “wide” or “long” format
- For growth modeling, we want to use a “long” format data file

Data structure for growth modeling

Long format = person-by-period data set
(multiple rows per person)

childid	female	bipoc	grade	rscore
10003426	0.00	1.00	0.00	59.57
10003426	0.00	1.00	1.00	92.12
10003426	0.00	1.00	2.00	110.61
10003426	0.00	1.00	3.00	122.42
10003426	0.00	1.00	4.00	123.91
10003426	0.00	1.00	5.00	132.05
10002116	0.00	0.00	0.00	86.64
10002116	0.00	0.00	1.00	128.71
10002116	0.00	0.00	2.00	144.51
10002116	0.00	0.00	3.00	144.01
10002116	0.00	0.00	4.00	149.99
10002116	0.00	0.00	5.00	157.31
10004918	0.00	0.00	0.00	79.41
10004918	0.00	0.00	1.00	107.91
10004918	0.00	0.00	2.00	139.62
10004918	0.00	0.00	3.00	129.31
10004918	0.00	0.00	4.00	148.18
10004918	0.00	0.00	5.00	155.87
10009310	1.00	0.00	0.00	77.81
10009310	1.00	0.00	1.00	103.92
10009310	1.00	0.00	2.00	124.44
10009310	1.00	0.00	3.00	129.16
10009310	1.00	0.00	4.00	136.11
10009310	1.00	0.00	5.00	144.24

Wide format = person-level data set
(single row for each person)

childid	female	bipoc	g0rscore	g1rscore	g2rscore	g3rscore	g4rscore	g5rscore
10003426	0.00	1.00	59.57	92.12	110.61	122.42	123.91	132.05
10002116	0.00	0.00	86.64	128.71	144.51	144.01	149.99	157.31
10004918	0.00	0.00	79.41	107.91	139.62	129.31	148.18	155.87
10009310	1.00	0.00	77.81	103.92	124.44	129.16	136.11	144.24
10004770	0.00	1.00	60.82	105.53	112.10	112.65	134.72	144.55
10001387	1.00	1.00	71.01	101.36	114.76	127.97	126.58	130.98
10002502	0.00	1.00	69.37	101.15	95.41	100.18	100.89	107.44
10001764	0.00	0.00	62.81	82.11	101.89	121.72	131.21	152.39
10003185	0.00	0.00	73.35	109.81	133.56	133.25	142.95	145.96
10006063	1.00	0.00	67.92	86.38	110.59	117.49	125.33	139.93
10008237	1.00	1.00	60.39	79.07	116.92	106.97	124.15	119.38
10013441	0.00	0.00	77.95	106.83	116.41	128.42	136.21	150.00

Multilevel Model to Characterize Change

Random-coefficient linear growth model

- Can use a simple random-coefficient linear model to examine:
  - Mean initial status (intercept) and rate of change (slope)
  - Between-person variation in initial status and rate of change
  - Relationship between initial status and rate of change

- The two-level model has two parts:
  - Level 1 can be thought of as the within-person or individual growth model
  - Level 2 can be thought of as the between-person or systematic growth model

Random-coefficient linear growth model

- (Level 1) Within-person model

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) + e_{ti} \,,$$

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

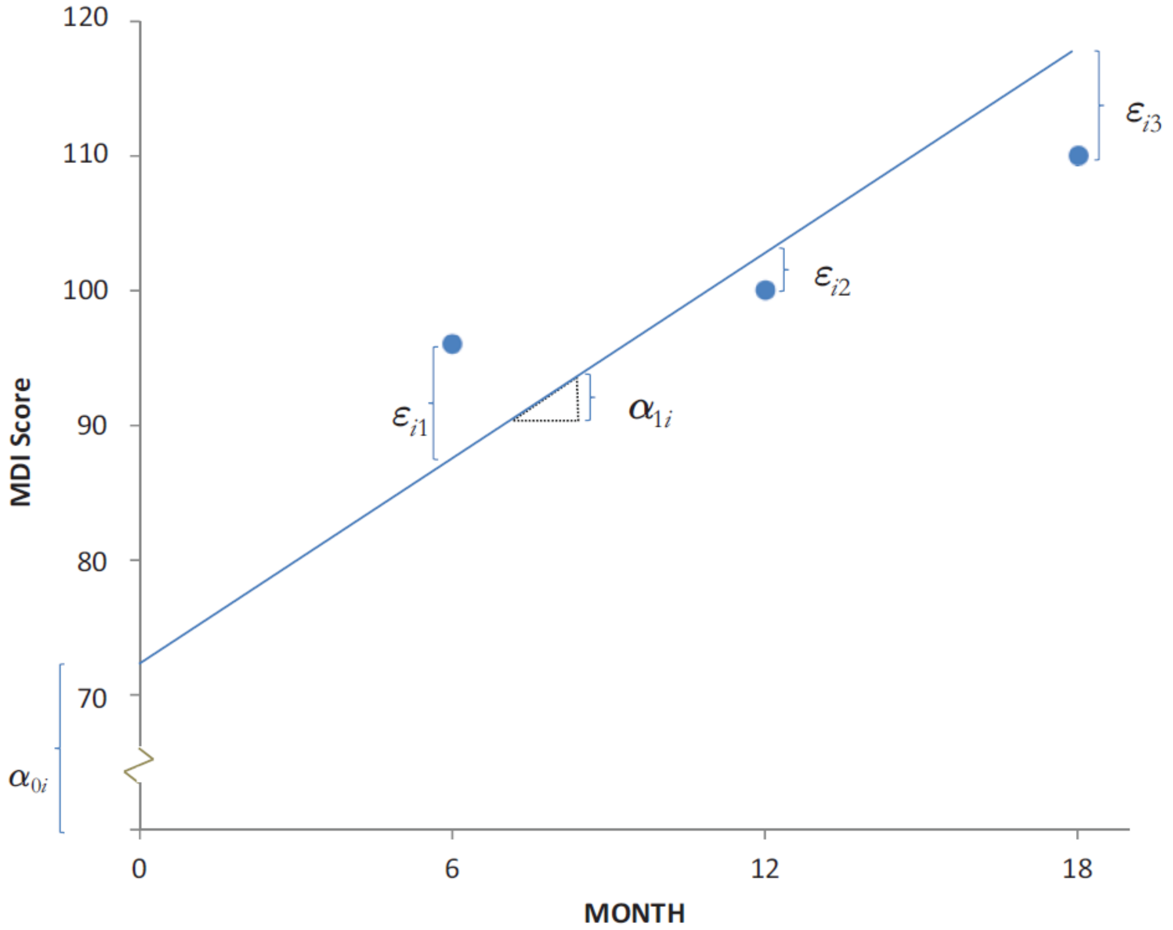


Image source: Gee, K. A. (2014). Multilevel Growth Modeling: An Introductory Approach to Analyzing Longitudinal Data for Evaluators. *American Journal of Evaluation*, *35*(4), 543–561. https://doi.org/10.1177/1098214014523823

Random-coefficient linear growth model

- (Level 1) Within-person model

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) + e_{ti} \ ,$$

π_{0i} : true initial status for person *i*

π_{1i} : true rate of change for person *i*

e_{ti} : deviation of observed score at time *t* for person *i* from the expected growth trajectory for person *i*

Random-coefficient linear growth model

- (Level 2) Between-person model

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$\begin{pmatrix} r_{0i} \\ r_{1i} \end{pmatrix} \sim MVN(\mathbf{0}, \mathbf{T})$$

$$\mathbf{T} = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}$$

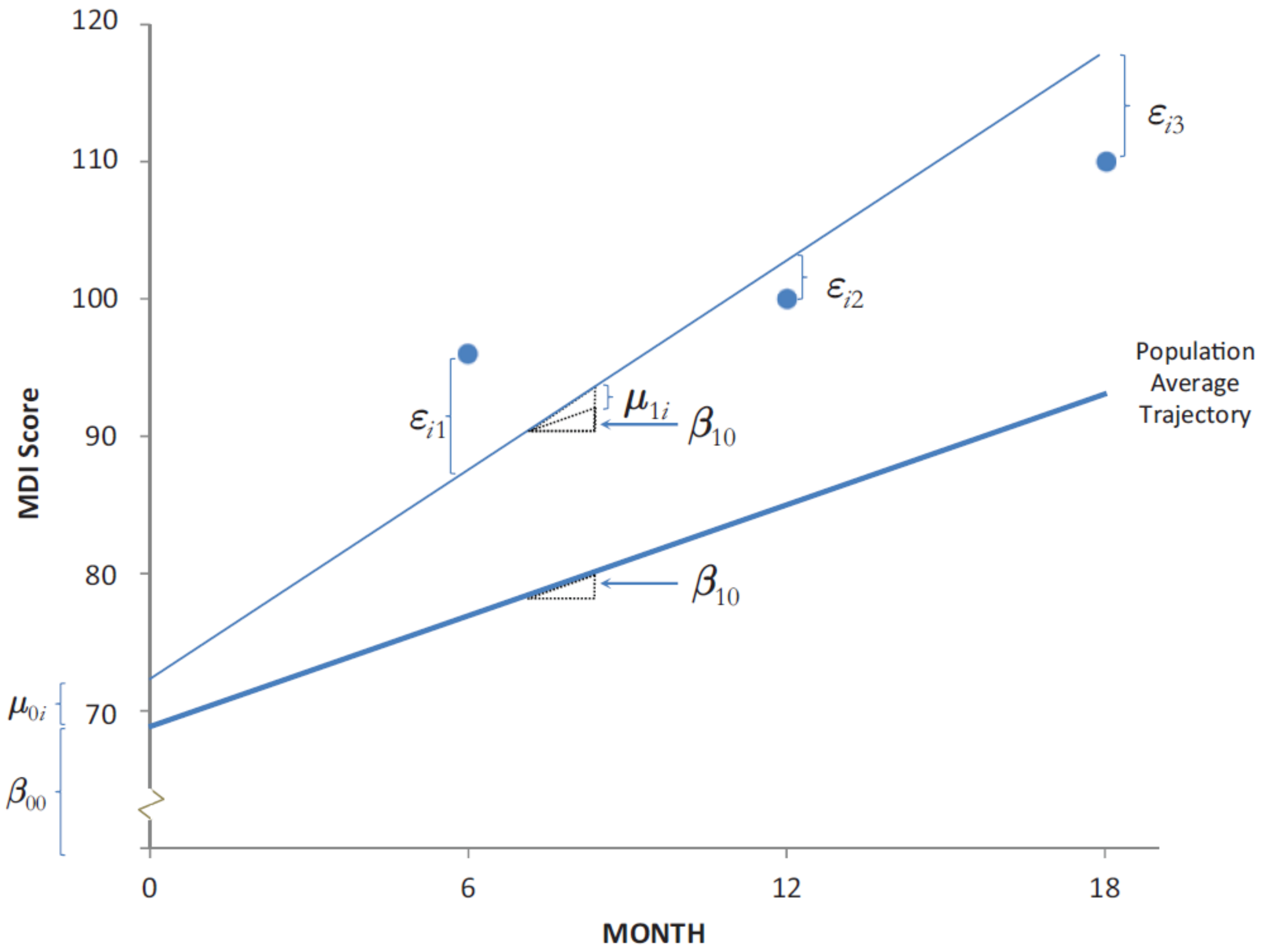


Image source: Gee, K. A. (2014). Multilevel Growth Modeling: An Introductory Approach to Analyzing Longitudinal Data for Evaluators. *American Journal of Evaluation*, *35*(4), 543–561. https://doi.org/10.1177/1098214014523823

Random-coefficient linear growth model

- (Level 2) Between-person model

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$\begin{pmatrix} r_{0i} \\ r_{1i} \end{pmatrix} \sim MVN(\mathbf{0}, \mathbf{T})$$

$$\mathbf{T} = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}$$

β_{00} : mean initial status for the population

β_{10} : mean rate of change for the population

r_{0i}: deviation of person *i*'s true initial status from mean initial status

r_{1i}: deviation of person *i*'s true rate of change from the mean rate

Random-coefficient linear growth model

- (Level 2) Between-person model

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$\begin{pmatrix} r_{0i} \\ r_{1i} \end{pmatrix} \sim MVN(\mathbf{0}, \mathbf{T})$$

$$\mathbf{T} = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}$$

τ_{00} : parameter variance in initial status across individuals

τ_{11} : parameter variance in rate of change across individuals

τ_{10} : parameter covariance between initial status and rate of change

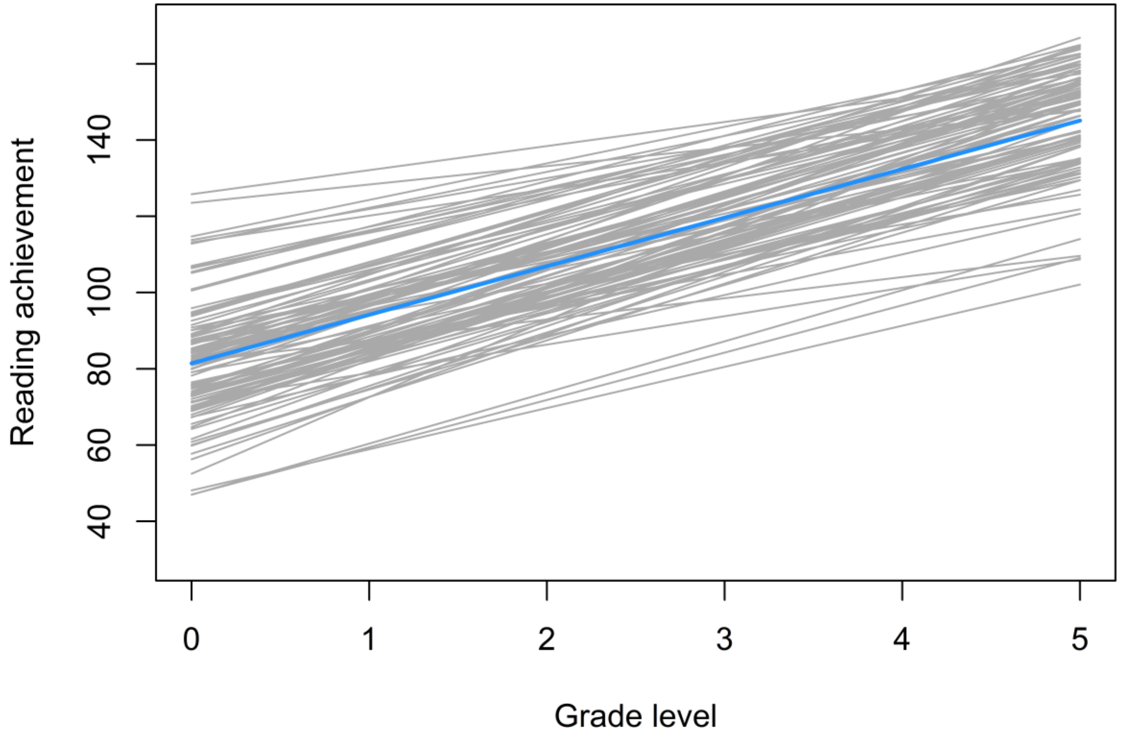
Linear growth model: ECLS-K example

- ECLS-K:2011 example
  - The data include 7,659 students with spring reading test scores from kindergarten through 5th grade
  - 6 time points per student
- Grade variable is our measure of "time," where 0 = spring of kindergarten

childid	female	bipoc	grade	rscore
10003426	0.00	1.00	0.00	59.57
10003426	0.00	1.00	1.00	92.12
10003426	0.00	1.00	2.00	110.61
10003426	0.00	1.00	3.00	122.42
10003426	0.00	1.00	4.00	123.91
10003426	0.00	1.00	5.00	132.05
10002116	0.00	0.00	0.00	86.64
10002116	0.00	0.00	1.00	128.71
10002116	0.00	0.00	2.00	144.51
10002116	0.00	0.00	3.00	144.01
10002116	0.00	0.00	4.00	149.99
10002116	0.00	0.00	5.00	157.31
10004918	0.00	0.00	0.00	79.41
10004918	0.00	0.00	1.00	107.91
10004918	0.00	0.00	2.00	139.62
10004918	0.00	0.00	3.00	129.31
10004918	0.00	0.00	4.00	148.18
10004918	0.00	0.00	5.00	155.87
10009310	1.00	0.00	0.00	77.81
10009310	1.00	0.00	1.00	103.92
10009310	1.00	0.00	2.00	124.44
10009310	1.00	0.00	3.00	129.16
10009310	1.00	0.00	4.00	136.11
10009310	1.00	0.00	5.00	144.24

Linear growth model: ECLS-K example

- OLS-estimated trajectories for random sample of 100 students



childid	female	bipoc	grade	rscore
10003426	0.00	1.00	0.00	59.57
10003426	0.00	1.00	1.00	92.12
10003426	0.00	1.00	2.00	110.61
10003426	0.00	1.00	3.00	122.42
10003426	0.00	1.00	4.00	123.91
10003426	0.00	1.00	5.00	132.05
10002116	0.00	0.00	0.00	86.64
10002116	0.00	0.00	1.00	128.71
10002116	0.00	0.00	2.00	144.51
10002116	0.00	0.00	3.00	144.01
10002116	0.00	0.00	4.00	149.99
10002116	0.00	0.00	5.00	157.31
10004918	0.00	0.00	0.00	79.41
10004918	0.00	0.00	1.00	107.91
10004918	0.00	0.00	2.00	139.62
10004918	0.00	0.00	3.00	129.31
10004918	0.00	0.00	4.00	148.18
10004918	0.00	0.00	5.00	155.87
10009310	1.00	0.00	0.00	77.81
10009310	1.00	0.00	1.00	103.92
10009310	1.00	0.00	2.00	124.44
10009310	1.00	0.00	3.00	129.16
10009310	1.00	0.00	4.00	136.11
10009310	1.00	0.00	5.00	144.24

Linear growth model: ECLS-K example

- Model estimation in R

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Grade_{ti}) + e_{ti}$$

$$\pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

```
m1 <- lmer(rscore ~ 1 + grade + (1 + grade | childid),
           data = e11xl,
           control = lmerControl(optimizer ="Nelder_Mead"))
```

Linear growth model: ECLS-K example

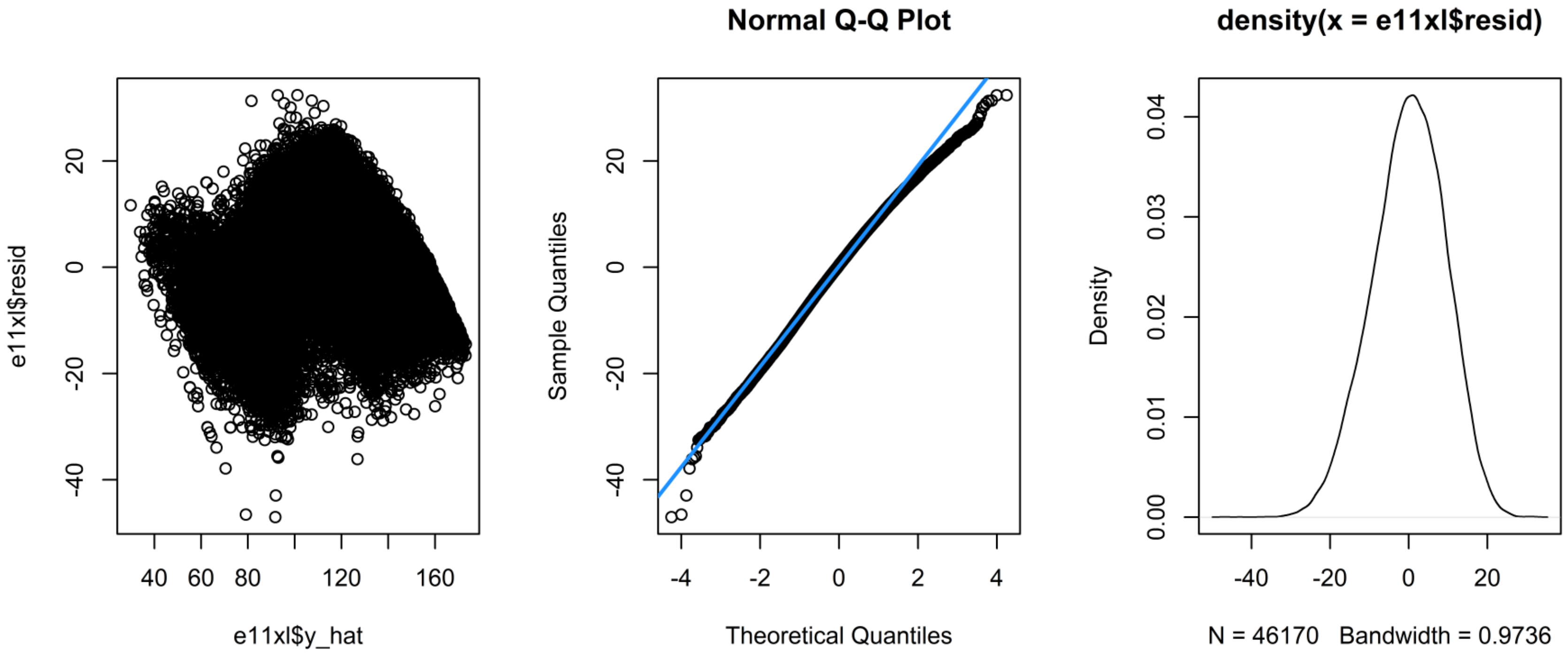
```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: rscore ~ 1 + grade + (1 + grade | childid)
   Data: e11xl
Control: lmerControl(optimizer = "Nelder_Mead")

REML criterion at convergence: 364678.1

Scaled residuals: 
    Min      1Q  Median      3Q     Max 
-4.6329 -0.6058  0.0379  0.6500  3.1896 

Random effects:
 Groups   Name        Variance Std.Dev. Corr 
 childid  (Intercept) 189.2605 13.7572       
          grade         0.5586  0.7474  -0.18
 Residual             103.1002 10.1538       
Number of obs: 46170, groups:  childid, 7695

Fixed effects:
             Estimate Std. Error        df t value Pr(>|t|)    
(Intercept) 7.916e+01  1.778e-01 7.694e+03   445.2   <2e-16 ***
grade       1.280e+01  2.895e-02 7.694e+03   442.1   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Small group discussion

- In groups of 3-4, take 10 minutes to answer the following questions based on the model output on the previous slide:
  - What's the estimated grand-mean reading score when students are in kindergarten?
  - What's the estimated grand-mean rate of reading growth from one year to the next?
  - To what extent do reading scores in kindergarten vary between students?
  - To what extent does the rate of reading growth vary between students?
  - What's the expected reading score in spring of 3rd grade for a student with average reading achievement in kindergarten and a rate of growth 1 standard deviation <u>below</u> the grand-mean rate of growth?
  - What's the expected reading score in spring of 3rd grade for a student with average reading achievement in kindergarten and a rate of growth 1 standard deviation <u>above</u> the grand-mean rate of growth?
  - Do students with higher-than-average reading scores in kindergarten tend to have faster or slower rates of growth compared to students with lower-than-average reading scores in kindergarten? What model results help you make that determination?

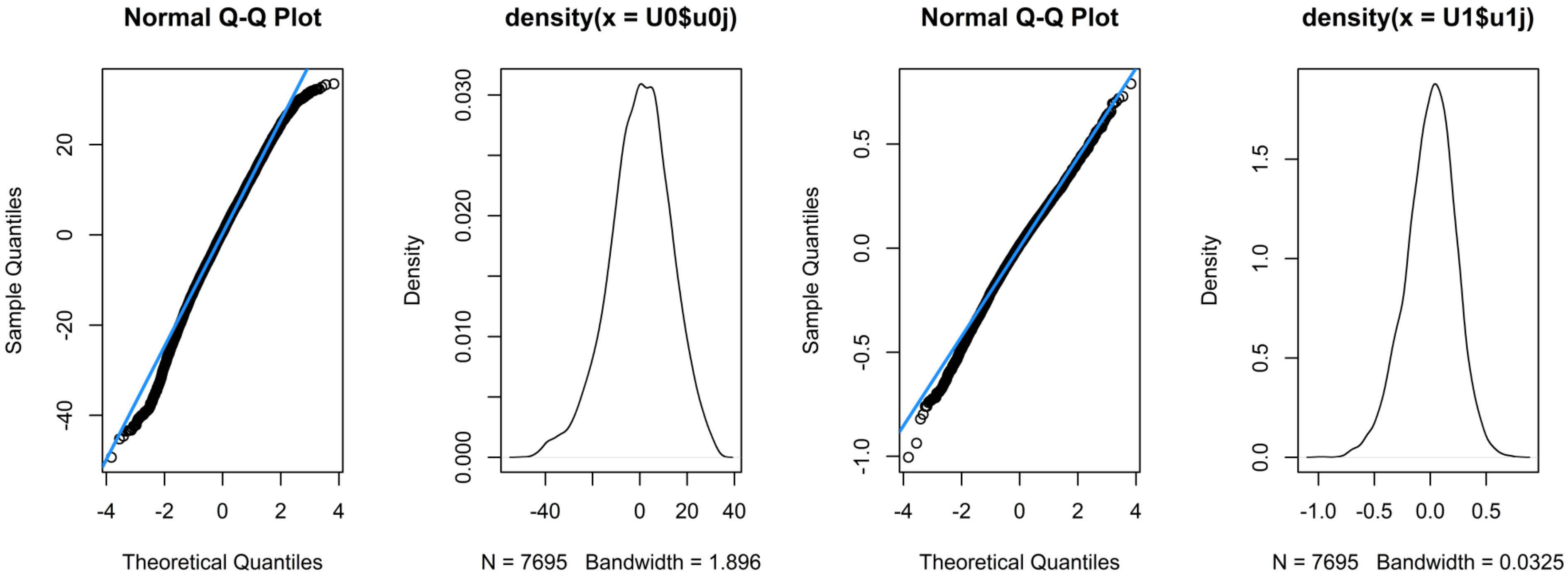
Linear growth model: ECLS-K example

- Check level-1 residuals to investigate model assumptions
  - Normality
  - Homoscedasticity



Linear growth model: ECLS-K example

- Check level-2 residuals to investigate model assumptions



Multilevel Model to Study Correlates of Change

Intercept- and slopes-as-outcomes model

- Can add time invariant measures in the level 2 model to test relationships with initial status and rates of change:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Time_{ti}) + e_{ti}\ , \qquad e_{ti} \sim N(0, \sigma^2\ )$$

$$\pi_{0i} = \beta_{00} + \beta_{01}(X_i) + r_{0i}$$

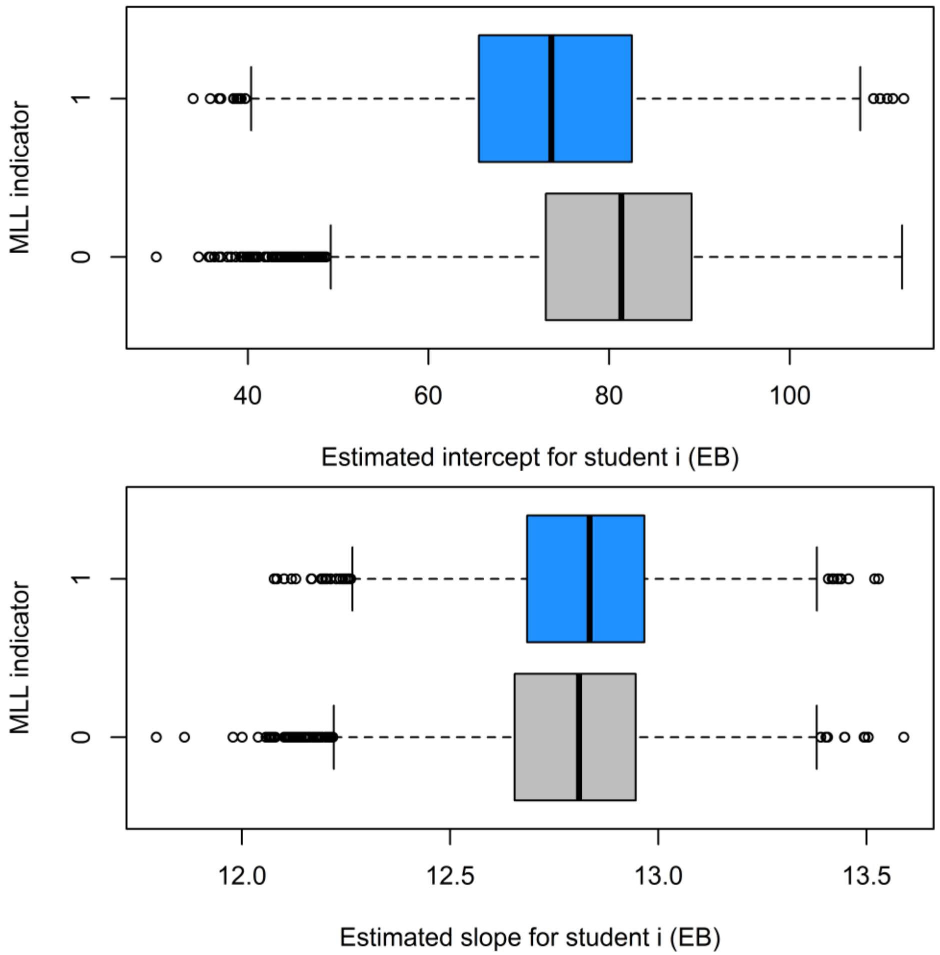
$$\pi_{1i} = \beta_{10} + \beta_{11}(X_i) + r_{1i}$$

$$\begin{pmatrix} r_{0i} \\ r_{1i} \end{pmatrix} \sim MVN(\mathbf{0}, \mathbf{T}), \qquad \mathbf{T} = \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix}$$

Linear growth model: ECLS-K example

- How does English reading achievement in kindergarten and the rate of English reading growth during elementary school for multilanguage learners (MLL) compare to that of English-only students?

	English only (N=6023)	MLL (N=1658)	Overall (N=7681)
Intercept (π^*_{0i})			
Mean (SD)	80.7 (12.6)	73.7 (13.2)	79.2 (13.0)
Median [Min, Max]	81.3 [29.9, 112]	73.6 [33.9, 113]	79.7 [29.9, 113]
Slope (π^*_{1i})			
Mean (SD)	12.8 (0.221)	12.8 (0.221)	12.8 (0.221)
Median [Min, Max]	12.8 [11.8, 13.6]	12.8 [12.1, 13.5]	12.8 [11.8, 13.6]



Linear growth model: ECLS-K example

- Model estimation in R

$$Y_{ti} = \pi_{0i} + \pi_{1i}(Grade_{ti}) + e_{ti}$$

$$\pi_{0i} = \beta_{00} + \beta_{01}(MLL_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(MLL_i) + r_{1i}$$

```
m2 <- lmer(rscore ~ 1 + grade + childmll + childmll*grade
           + (1 + grade | childid), data = e11xl)
```

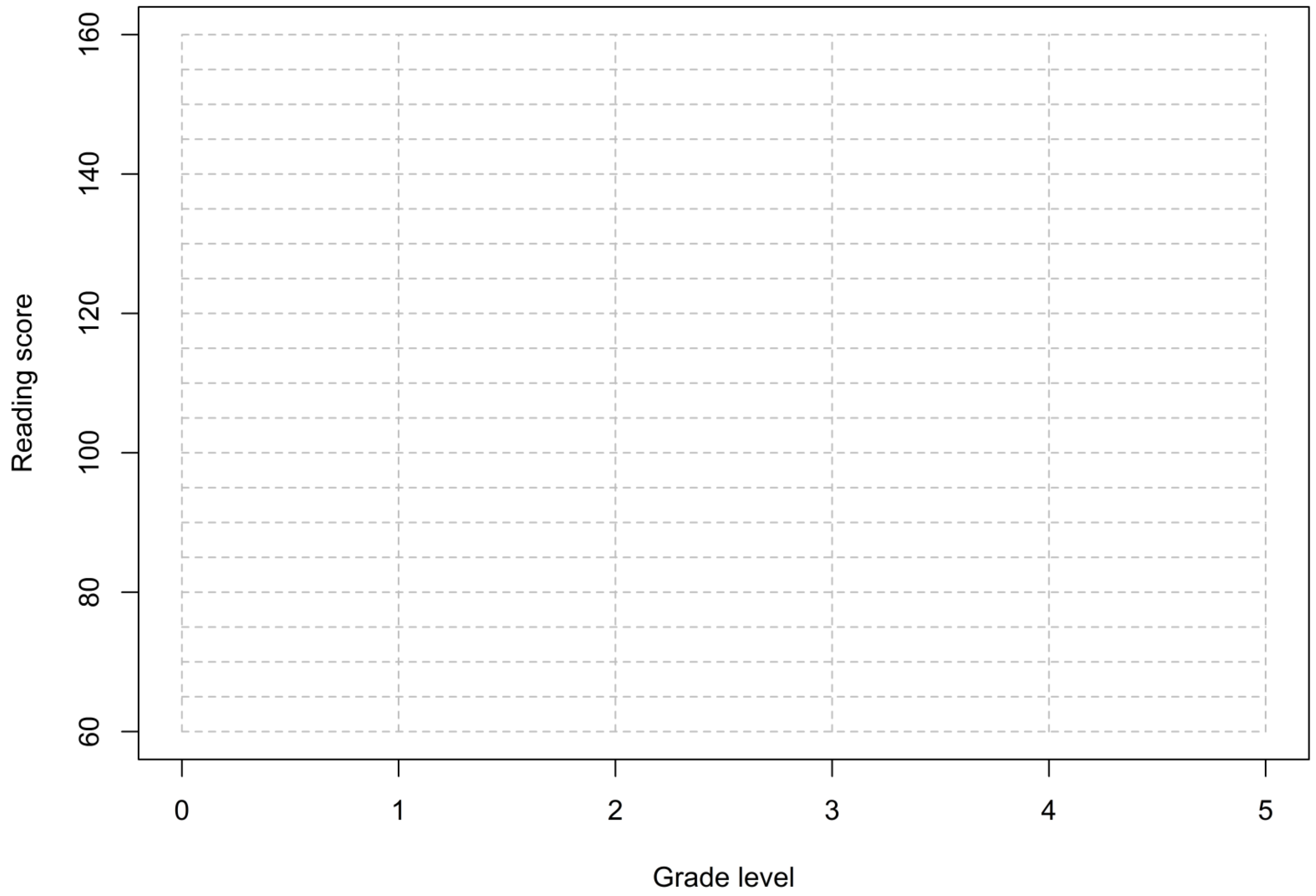
Linear growth model: ECLS-K example

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: rscore ~ 1 + grade + childmll + childmll * grade + (1 + grade |      childid)
   Data: e11xl

REML criterion at convergence: 363634.9

Scaled residuals: 
   Min      1Q Median      3Q     Max 
-4.649 -0.604  0.038  0.649  3.185 

Random effects:
 Groups   Name        Variance Std.Dev. Corr 
 childid  (Intercept) 178.3533 13.3549       
          grade         0.5437  0.7373  -0.16
 Residual             103.1517 10.1564       
Number of obs: 46086, groups:  childid, 7681

Fixed effects:
                 Estimate Std. Error         df t value Pr(>|t|)    
(Intercept)      80.91543    0.19643 7679.00577 411.940  < 2e-16 ***
grade            12.75763    0.03269 7679.11755 390.210  < 2e-16 ***
childmll         -7.99738    0.42278 7679.00428 -18.916  < 2e-16 ***
grade:childmll    0.18846    0.07037 7679.11719   2.678  0.00742 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Small group discussion

- In groups of 3-4, take 10 minutes to answer the following questions based on the model output on the previous slide:
  - What’s the estimated grand-mean reading score for a MLL in kindergarten?
  - What’s the estimated grand-mean reading score for an English only student in kindergarten?
  - Would you say that reading performance in kindergarten differs based on MLL status? Why or why not?
  - What’s the estimated grand-mean rate of reading growth for a MLL student?
  - What’s the estimated grand-mean rate of reading growth for an English-only student?
  - Would you say that rate of growth differs based on MLL status? Why or why not?
  - Sketch out the estimated linear growth trajectory for the average MLL students and the average English-only student. (You can use the empty graph on the next slide.)
  - Based on the model results and your sketch, what implications do you draw about MLL students learning to read in elementary school?

Linear growth model: ECLS-K example



Linear growth model: ECLS-K example