Unit 3, Lecture 2
Quadratic Growth Models

Submitting Project A

1. Go on to Canvas
2. Click on "People"
3. Click on "ProjectA" (a "Group")
4. Select a random group:

```
> sample( 17, 1 )
```

5. Sign up you and your partner for that group (if it is not taken). Otherwise, re-run your R code and try again.
6. Do group-submission for Project A.

Today's Goals

- ★ Continue our case-study on science knowledge in HeadStart to…
  - ★ See how we can use covariates to predict growth
  - ★ Measure how much of growth we are explaining

- ★ Look at an alternate grown curve, the quadratic growth curve
  - ★ This allows us to capture nonlinear growth (more plausible in many contexts)
  - ★ We will need to think carefully about what our random coefficients mean given the curvature
  - ★ We also need to handle three random effects per student!

- ★ Talk about how statistics and modeling can intersect with policy debate

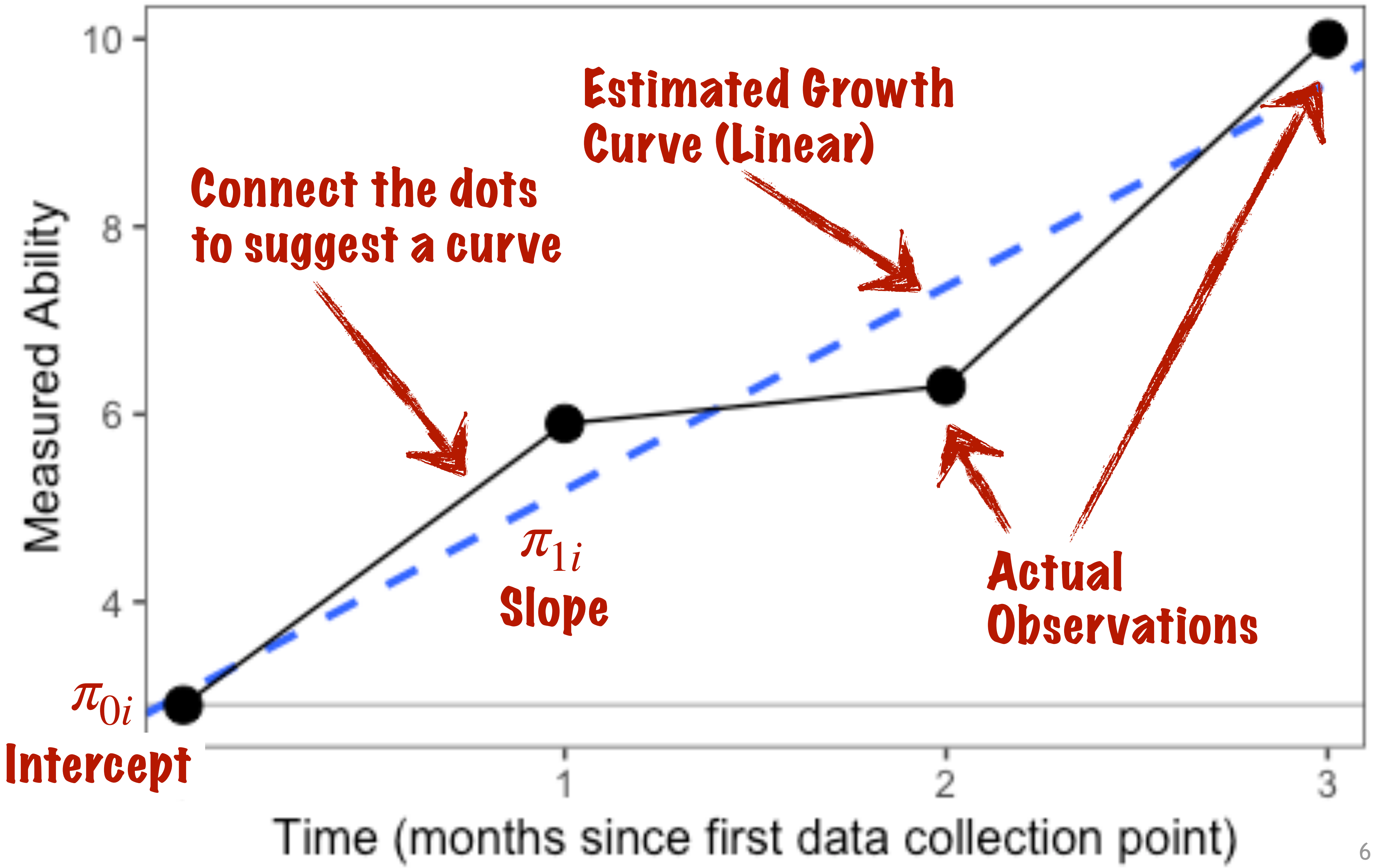
Including covariates in Linear Growth Models

Natural Sciences Knowledge (HeadStart)

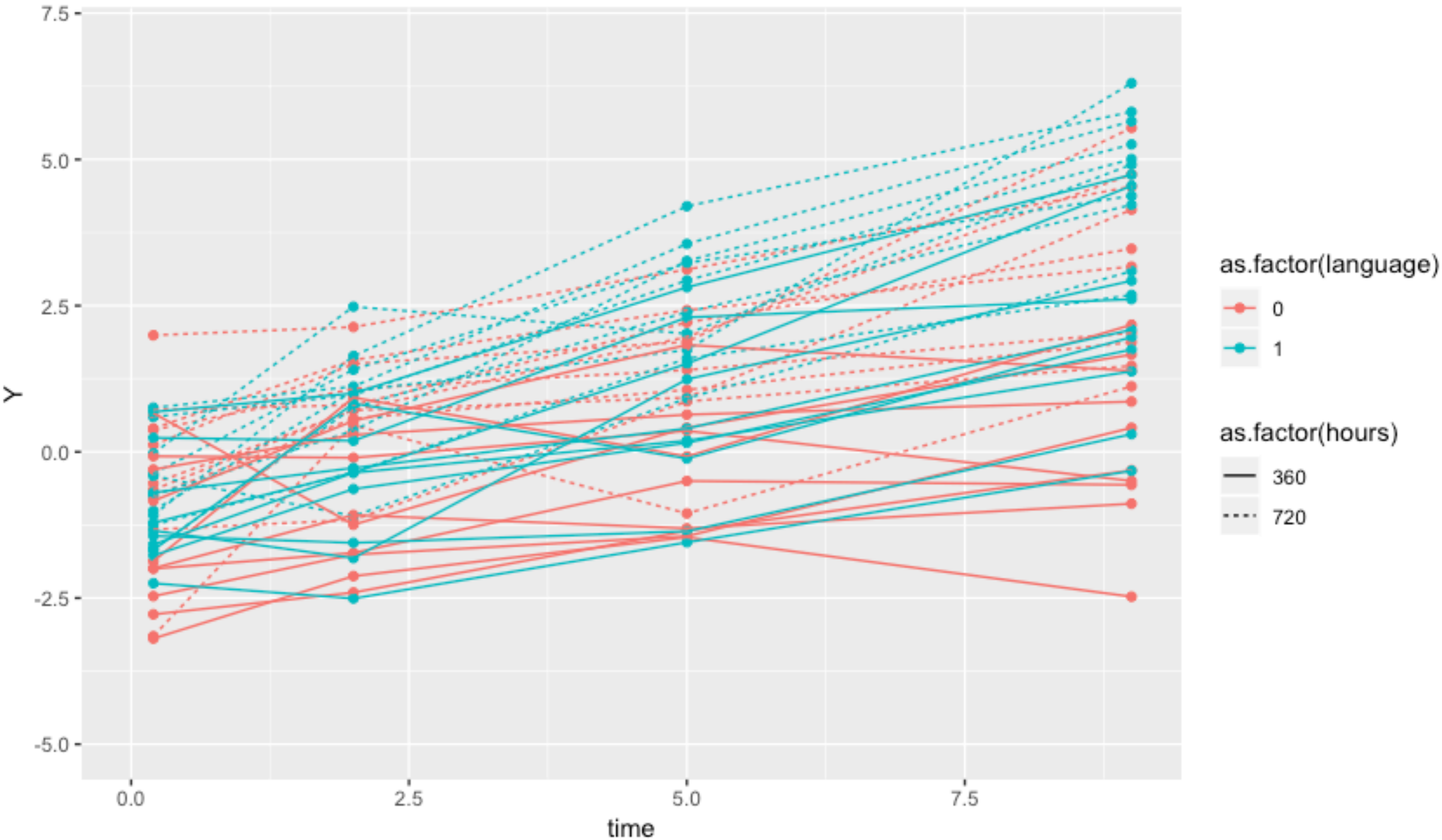
Data described on R&B pg 164

- ★ 143 children enrolled in Head Start
- ★ Outcome is an IRT scaling of a set of test items, a "logit metric" (and so is a measure of student knowledge)
- ★ Data are lost; using tables and discussion from R&B chapter 6
- ★ Each child (supposed to be) tested 4 times along year
- ★ Age a_{ti} defined as *time from first data-collection point* (in months)
- ★ Also collected home language (Spanish or English) and number of hours of direct instruction

A Hypothetical Child



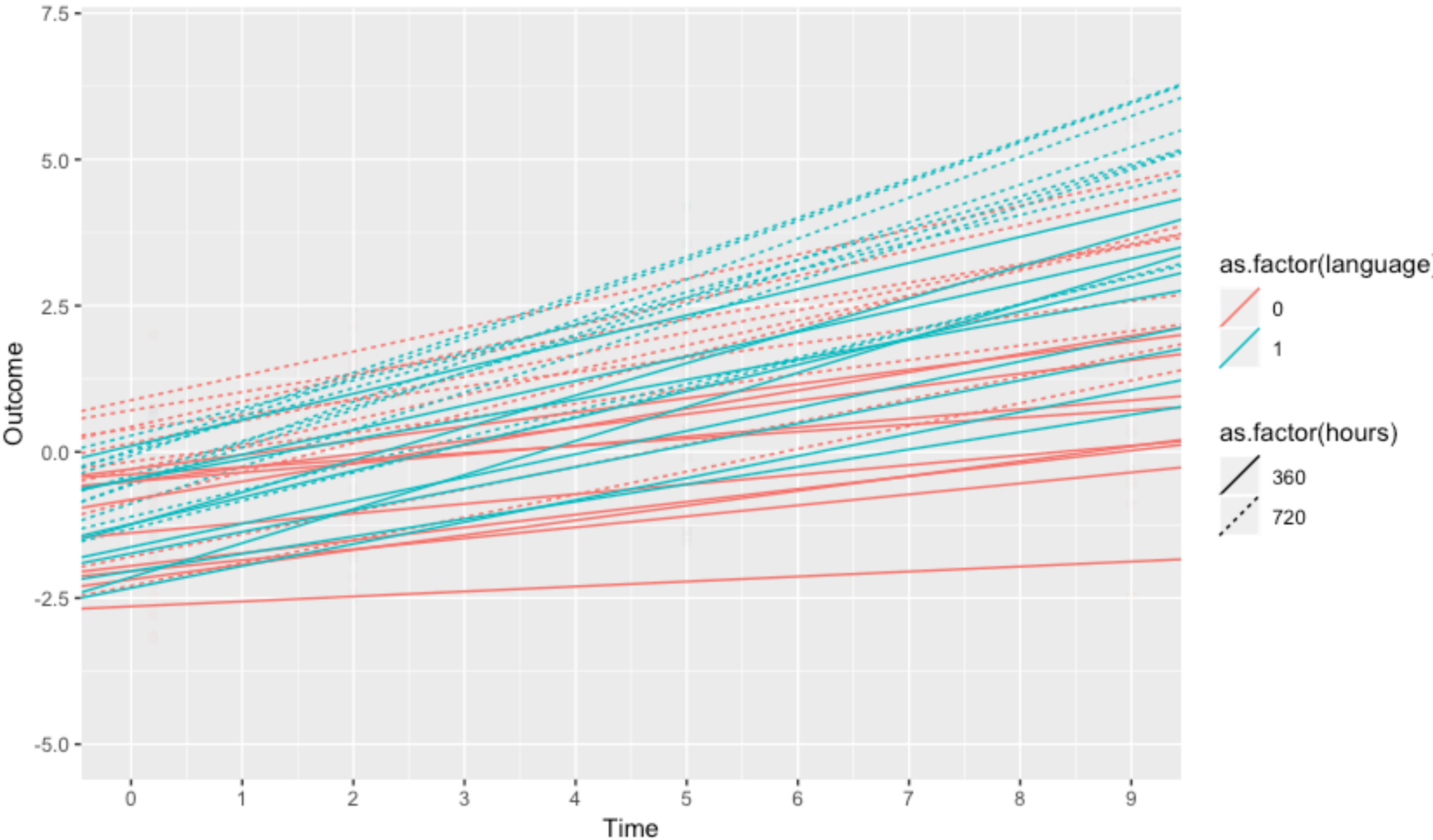
Nominal raw data for a bunch of children



Data unfortunately simulated

Each child has their own trajectory

These trajectories may be systematically varying by language or hours of instruction



Question to motivate what happens next:

How do we model the connection between this variations and covariates?

Linear Growth Model with Child Predictors

$$Y_{ti} = \pi_{0i} + \pi_{1i} a_{ti} + e_{ti}$$

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

Xi is our child-level predictor. In this equation it predicts initial ability

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

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

$$(r_{0i}, r_{1i}) \sim N(\mathbf{0}, \Sigma)$$

Here it is predicting impact on rate of growth.

★ We can easily add in level-2 covariates to predict (model) the various aspects of the curve. These are our demographics, for example.

★ Similarly (not pictured above), adding level-1 predictors is easy (these would be *time-varying* predictors).

A model incorporating 2 predictors

Fixed Effect	*Coefficient*	*se*	*t Ratio*
Model for initial status, π_{0i}			
BASE, β_{00}	0.895	0.267	3.35
LANGUAGE, β_{01}	−0.463	0.304	−1.52
HOURS, β_{02}	1.523×10^{-3}	0.853×10^{-3}	1.79
Model for growth rate, π_{1i}			
BASE, β_{10}	0.029	0.039	0.74
LANGUAGE, β_{11}	0.187	0.045	4.20
HOURS, β_{12}	4.735×10^{-4}	1.252×10^{-4}	3.78

★ Language: 1 = non-English, 0 = English

★ Hours: Continuous measure of number of hours of direct instruction received by child in that program year

R&B pg 168

What is the model that was fit?

Thinking about the model

★ The first testing occasion (coded as time $a_{1i} = 0$) for kid i occurred between 6 and 14 weeks into the program year

How might this have shown up in our estimated coefficients?

★ The coefficient for hours was estimated as 0.00474. Is this meaningful?

- Consider two kids, one in a 40hr/month care and one in 80hr/month care (approx the min and max in the sample) for a 9 month program.
  What is the difference in their expected knowledge by the end?

Comparing Models

Model	*Initial Status* Var(π_{0i})	*Growth Rate* Var(π_{1i})
Unconditional[a]	1.689	0.041
Conditional on LANGUAGE and HOURS[b]	0.761	0.010
Proportion of variance explained	54.9	75.0

a. From Table 6.1.
b. These are residual variances based on the model estimated in Table 6.2.

- ★ 1 - Var(Cond)/Var(Uncond) = Proportion explained
- ★ We see that our two variables do a great job at explaining variation in initial knowledge and, more so, rate of learning.

R&B pg 169

Quick Comment on Model Building

- ★ We started (last lecture) with an **unconditional model** where we just tried to assess whether there was random variation (yeah, there was!).
- ★ We also saw that the variation was substantial enough, relative to the scale of *y* and number of time points measured, to potentially capture with predictors.
- ★ We then fit models with predictors and interpreted that model.
- ★ We also assessed how much of the original variation was explained by these predictors.

Quadratic Growth

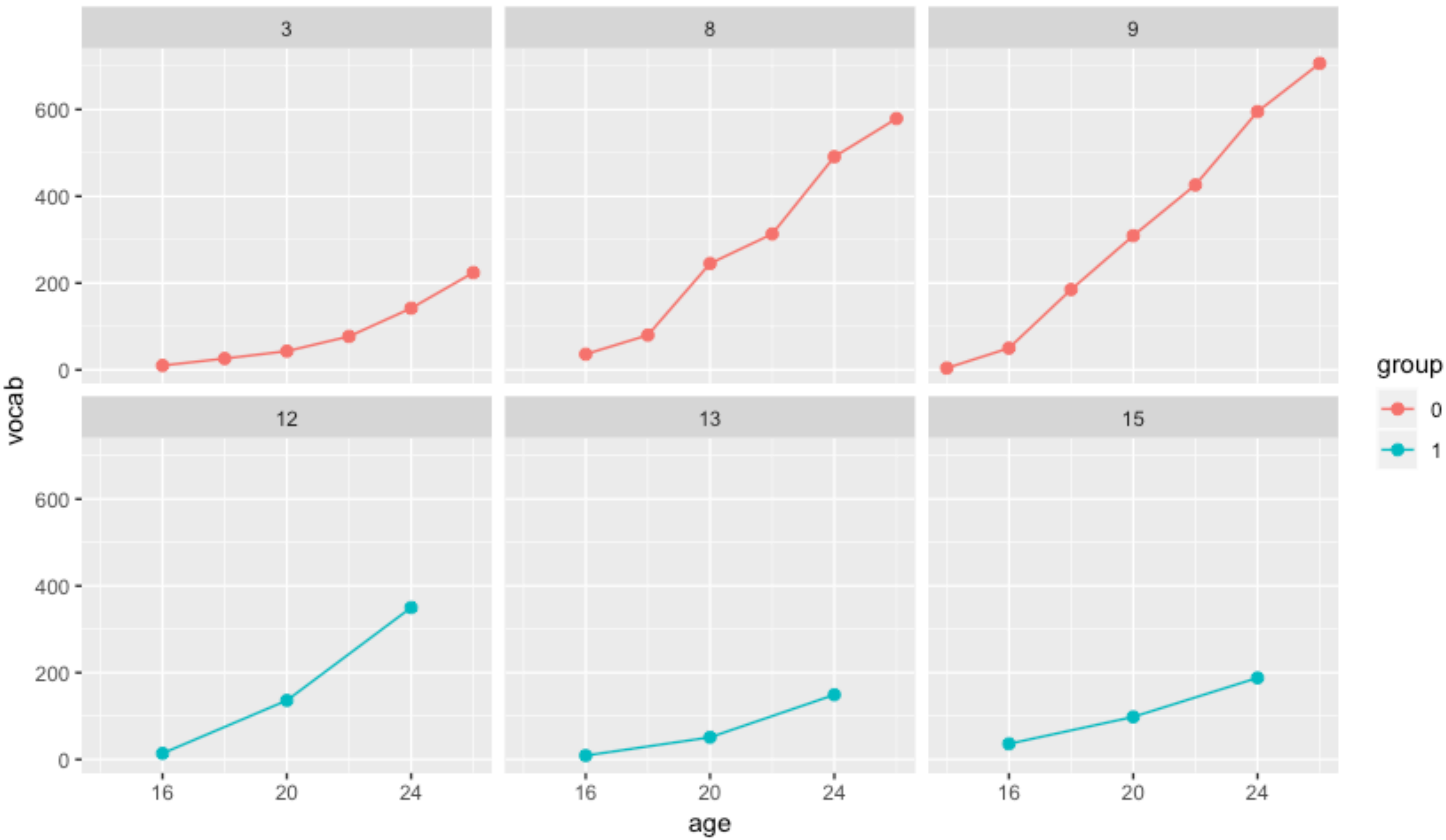
Example from Raudenbush & Bryk, Chapter 6, pgs 169-176

Maternal Speech on Children's Vocabulary

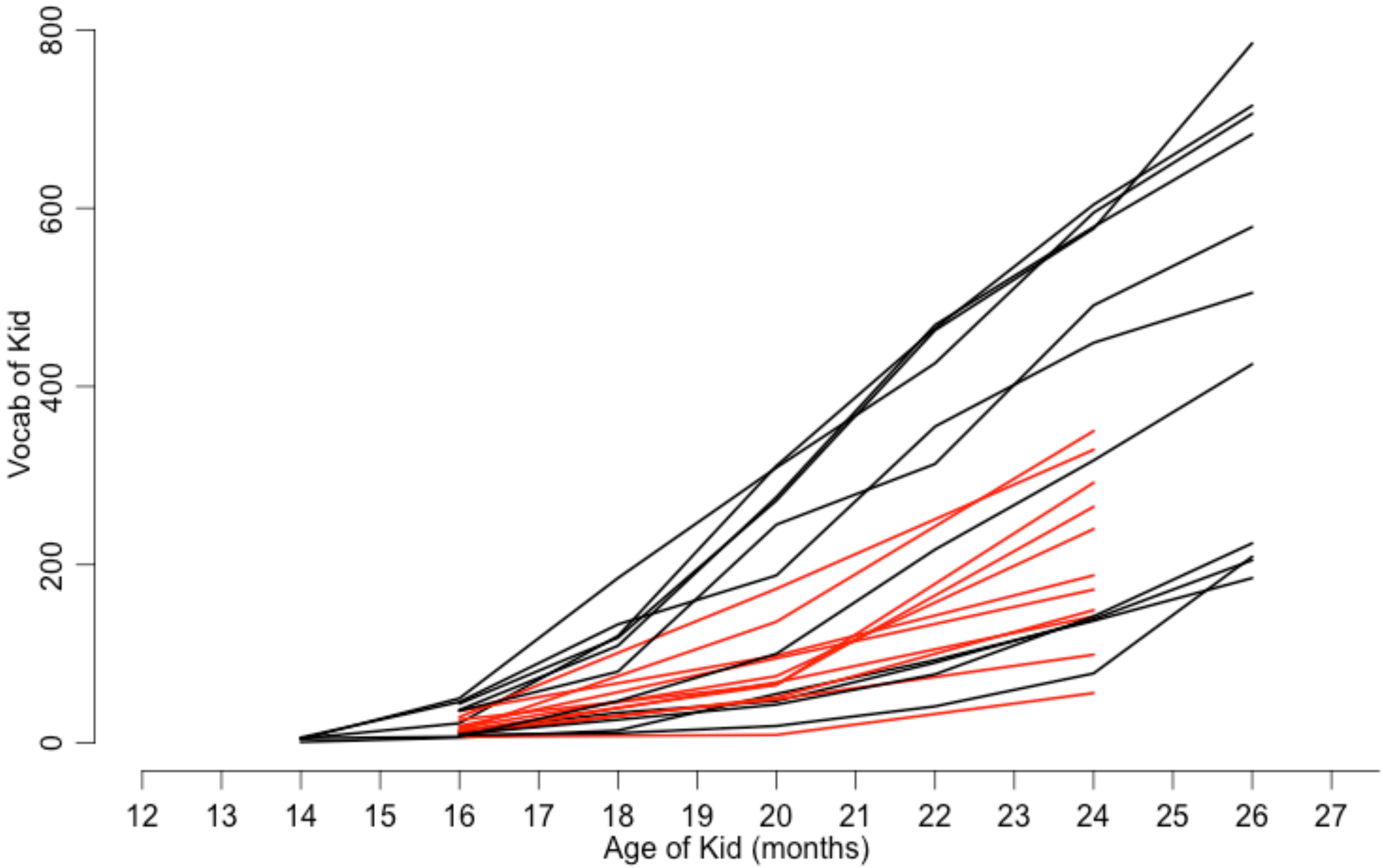
- ★ 22 children observed in the home on differing number of occasions throughout about 1.5 years
- ★ At 16 months, amount of maternal speech was recorded
- ★ Actually two studies with different patterns of observation
  - Study 1 (11 kids) has more home visits (observations)

Research Question:
Describe the association between maternal speech and child's vocabulary

6 of the kids across time



All the raw data (22 children)

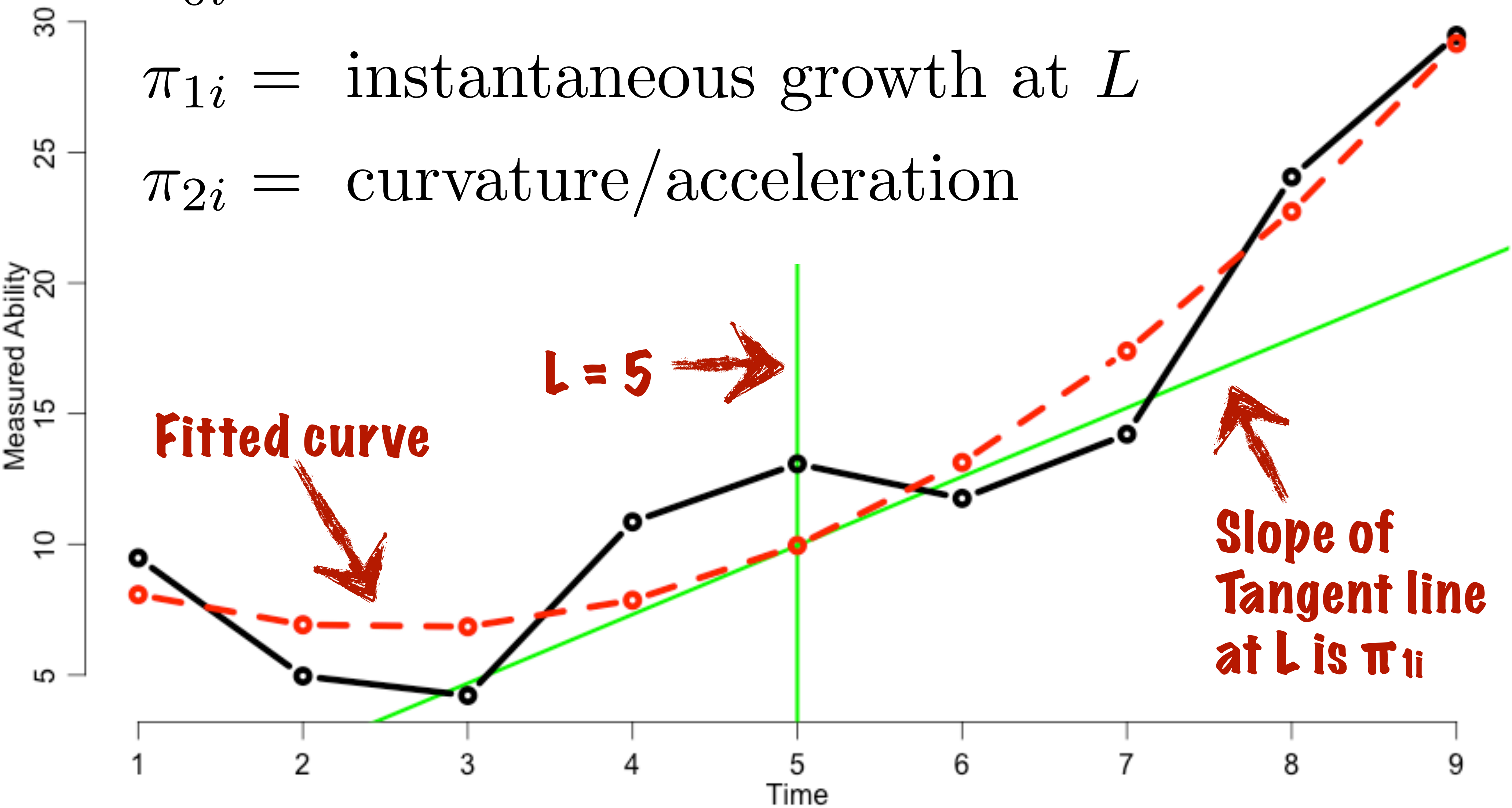


Anatomy of a Quadratic Growth Curve

$$\pi_{0i} = \text{ status at } L$$

$$\pi_{1i} = \text{ instantaneous growth at } L$$

$$\pi_{2i} = \text{ curvature/acceleration}$$



Unconditional Quadratic Growth Model

$$Y_{ti} = \pi_{0i} + \pi_{1i}(a_{ti} - L) + \pi_{2i}(a_{ti} - L)^2 + \varepsilon_{ti}$$

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

$$\pi_{0i} = \gamma_{00} + u_{0i}$$

$$\pi_{1i} = \gamma_{10} + u_{0i}$$

$$\pi_{2i} = \gamma_{20} + u_{0i}$$

$$(u_{0i}, u_{1i}, u_{2i}) \sim N(0, \Sigma)$$

How many parameters?

★ *Three* random effects, all correlated.

★ We have a fixed centering constant *L*. *(This is a constant picked by you, not a parameter.)*

Our Covariance Matrix Expanded

$$\sum = \begin{bmatrix} \tau_{00} & \text{Symmetric} & \\ \tau_{10} & \tau_{11} & \\ \tau_{20} & \tau_{21} & \tau_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \text{Var}(\pi_{0i}) & & \text{Symmetric} \\ \text{Cov}(\pi_{1i}, \pi_{0i}) & \text{Var}(\pi_{1i}) & \\ \text{Cov}(\pi_{2i}, \pi_{0i}) & \text{Cov}(\pi_{2i}, \pi_{1i}) & \text{Var}(\pi_{2i}) \end{bmatrix}.$$

Picking Model, Setting L

Individual plots from before suggest a curved fit. Quadratic is a first reasonable choice.

L is an important centering decision.

In our case, L=12 months because this is age where words first begin.

(However, picking L in the center of your data is usually preferable without good reason not to.)

We start by looking only at the 11 kids with mostly full data.

Fitting our quadratic unconditional model

We first recenter and make our quadratic predictor

```
> dat$age12 = dat$age - 12
> dat$age12sq = dat$age12 ^ 2
> M1 = lmer( vocab ~ 1 + age12 + age12sq + (1+age12 +
                              age12sq|pers), data=dat.g0 )
> display( M1 )
            coef.est coef.se
(Intercept) -45.05    27.52
age12        12.14     8.57
age12sq       1.84     0.26

Error terms:
 Groups   Name        Std.Dev. Corr
 pers     (Intercept) 73.76
          age12       24.62    -0.99
          age12sq      0.27     0.45 -0.54
 Residual             26.89
---
number of obs: 71, groups: pers, 11
AIC = 737.1, DIC = 738
deviance = 727.3
```

γ_{00} γ_{10} γ_{20}

σ

$\sqrt{\tau_{00}}$ $\sqrt{\tau_{11}}$ $\sqrt{\tau_{22}}$

! Results in R&B book different. Different data?

$$Y_{ti} = \pi_{0i} + \pi_{1i}(a_{ti} - L) + \pi_{2i}(a_{ti} - L)^2 + \varepsilon_{ti}$$

Predicting Vocabulary

Using the model, predict vocabulary for a typical kid at

- ★ 20 months
- ★ 12 months

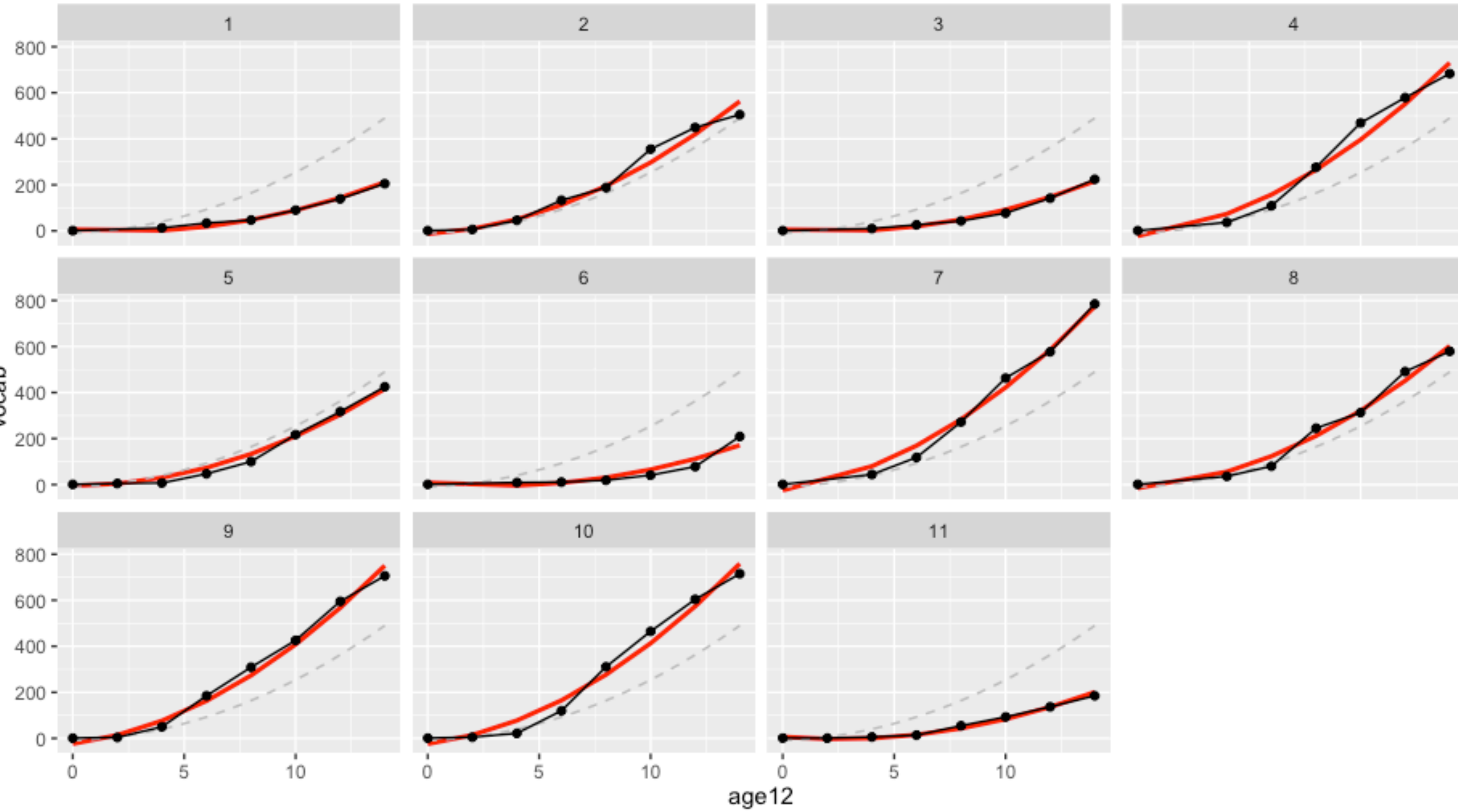
pollev.com/yayMLM

Getting more complicated:

- ★ How can we predict a distribution of vocabularies across kids at 20 months?

```
            coef.est  coef.se   Error terms:
(Intercept) -45.05    27.52      Groups   Name        Std.Dev.
age12        12.14     8.57      pers     (Intercept) 73.76
age12sq       1.84     0.26               age12       24.62
                                          age12sq      0.27
                                 Residual             26.89
```

Individual Curves for Each Kid



Getting these individual curves

```
> ranef( M1 )$pers
   (Intercept) age12   age12sq
1           81 -26.1  0.06196
2          -22   9.2 -0.22034
3           80 -26.6  0.13738
4          -70  23.1 -0.09955
5           22  -9.1  0.19027
6           93 -31.8  0.22082
7          -82  25.0  0.05896
8          -33  10.3  0.00099
9          -76  27.0 -0.28605
10         -78  26.1 -0.14186
11          83 -27.1  0.07743
```

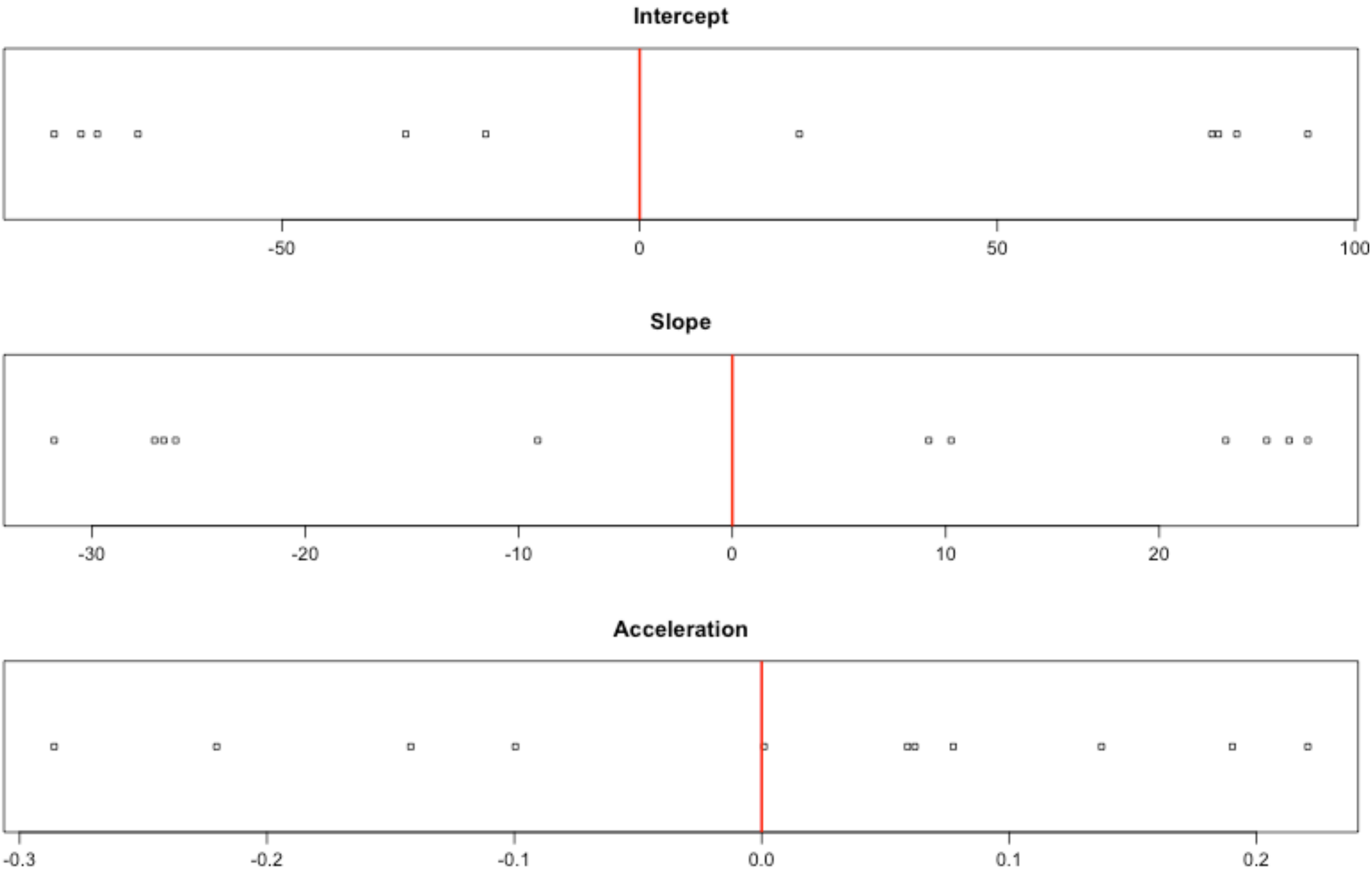
Each row is a single kid.

We have individual coefficients that vary for each kid.

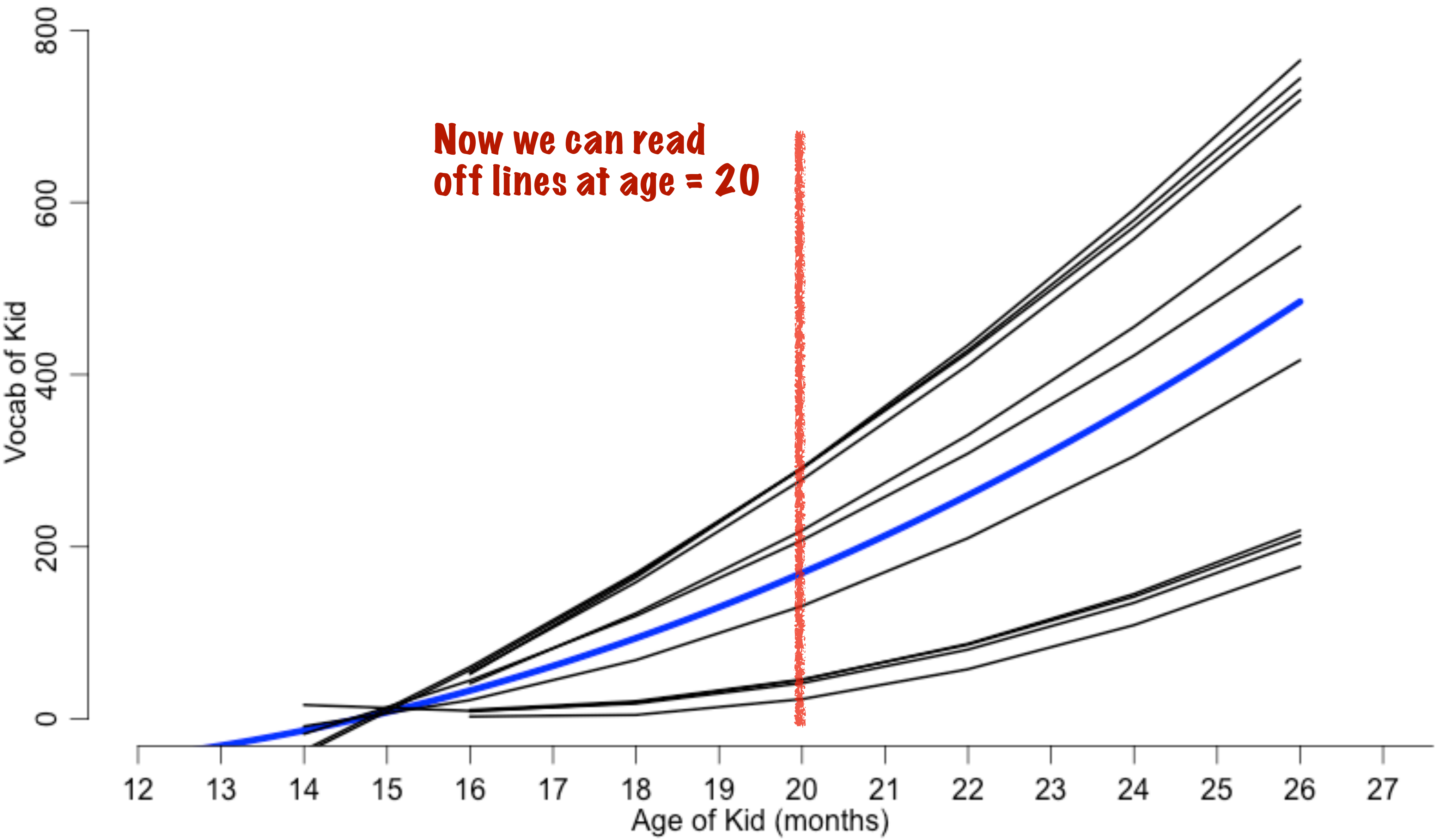
```
> coefs = coef( M1 )$pers
> coefs
   (Intercept)       age12  age12sq
1     35.84194 -13.930663 1.897603
2    -66.57032  21.343986 1.615293
3     34.95373 -14.488695 1.973026
4   -115.17942  35.260103 1.736078
```

This takes the fixed effects into account as well. We can read off our quadratic curves

Individual random effects



Predicted Median Curve vs Individually Estimated Growth Curves (from Empirical Bayes)



Interpretation: Changing Growth Rate

$$\frac{d}{da} Y_{ti} = \pi_{0i} + \pi_{1i}(a - L) + \pi_{2i}(a - L)^2 + e_{ti}$$
$$= \pi_{1i} + 2\pi_{2i}(a - L)$$

★ Rate of growth at age *a* for kid *i* is the derivative of our curve at *a*.

★ So, what do we have for our rate of growth at 12 months? 14 months?

Evaluation, modeling and policy

Let's Stop Talking About
the '30 Million Word Gap'

Talking with children matters:
Defending the 30 million word gap

Model building

How can we get a nice, simple model given our data?

Our Original Unconditional Model (repeat slide)

```
> dat$age12 = dat$age - 12
> dat$age12sq = dat$age12 ^ 2
> M1 = lmer( vocab ~ 1 + age12 + age12sq + (1+age12 +
                           age12sq|pers), data=dat.g0 )
> display( M1 )
lmer(formula = vocab ~ 1 + age12 + age12sq + (1 + age12 +
age12sq |
    pers), data = dat.g0)
            coef.est coef.se
(Intercept) -45.05    27.52
age12        12.14     8.57
age12sq       1.84     0.26

Error terms:
 Groups   Name        Std.Dev. Corr
 pers     (Intercept) 73.76
          age12       24.62    -0.99
          age12sq      0.27     0.45 -0.54
 Residual             26.89
---
number of obs: 71, groups: pers, 11
AIC = 737.1, DIC = 738
deviance = 727.3
```

Wildly off our theory

Something to worry about

A lot of model for not much data

Theory tells us:

★ Initial vocab at 12 mo should be generally very small (i.e., around 0)

Model:

★ Our curves seem “overdispersed”: possibly missing a variable?

★ Intercept is very negative when it shouldn’t be

Next step: Drop our intercept! (Fix it to zero.)

Model Refinement #1: Quadratic Growth with no intercept

```
> M2 = lmer( vocab ~ 0 + age12 + age12sq +
(0+age12+age12sq|pers), data=dat.g0 )
> display( M2 )
lmer(formula = vocab ~ 0 + age12 + age12sq + (0 + age12
+ age12sq |
    pers), data = dat.g0)
        coef.est coef.se
age12   0.73     2.52
age12sq 2.46     0.30

Error terms:
 Groups   Name    Std.Dev. Corr
 pers     age12    6.09
          age12sq  0.87    1.00
 Residual         32.80
---
number of obs: 71, groups: pers, 11
AIC = 758, DIC = 752
deviance = 749.0
```

SD has gone up from 27. Our constrained model is not hugging the data as tightly, but the difference is not huge

Model Refinement #2
Set **expected** rate of growth at 12mo to 0

```
> M3 = lmer( vocab ~ 0 + age12sq + (0+age12+age12sq|
pers), data=dat.g0 )
> display( M3 )
lmer(formula = vocab ~ 0 + age12sq + (0 + age12 +
age12sq | pers),
    data = dat.g0)
coef.est  coef.se
    2.43     0.29

Error terms:
 Groups   Name     Std.Dev. Corr
 pers     age12     5.71
          age12sq   0.87    1.00
 Residual          32.68
---
number of obs: 71, groups: pers, 11
AIC = 759.8, DIC = 748
deviance = 749.1
```

Now we have dropped our linear term in our fixed effect, but left it as a random effect. (Hence expected is 0, but not individual.)

lr test of simpler vs more complete models

```
> anova( M1, M2, M3 )
refitting model(s) with ML (instead of REML)
Data: dat.g0
Models:
M3: vocab ~ 0 + age12sq + (0 + age12 + age12sq | pers)
M2: vocab ~ 0 + age12 + age12sq + (0 + age12 + age12sq | pers)
M1: vocab ~ 1 + age12 + age12sq + (1 + age12 + age12sq | pers)
   Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)
M3  5 759 770   -375      749
M2  6 761 775   -375      749  0.09      1    0.76722
M1 10 747 770   -364      727 21.72      4    0.00023 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Check out refitting automatically

Note how these models are NESTED
(each is an expansion of the one above it)

anova() commands does lrtest and refits models automatically

Looking back

Recap

Check-In
http://cs179.org/lec32

On Level 2 Predictors

- We can include predictors to help understand why some grow faster than others.
- We can measure how much variation in growth these predictors explain.

On Quadradic Growth Models

- Quadratic growth allows for a moderately flexible *curved* growth curve
- We can constrain aspects of a growth curve model to get a simpler model that is still not linear
- The interpretation of the intercept, slope, and acceleration depend heavily on the choice of centering variable L.

Appendix:
Including covariates in our quadratic growth model

Including covariates: Our full model on all 22 kids

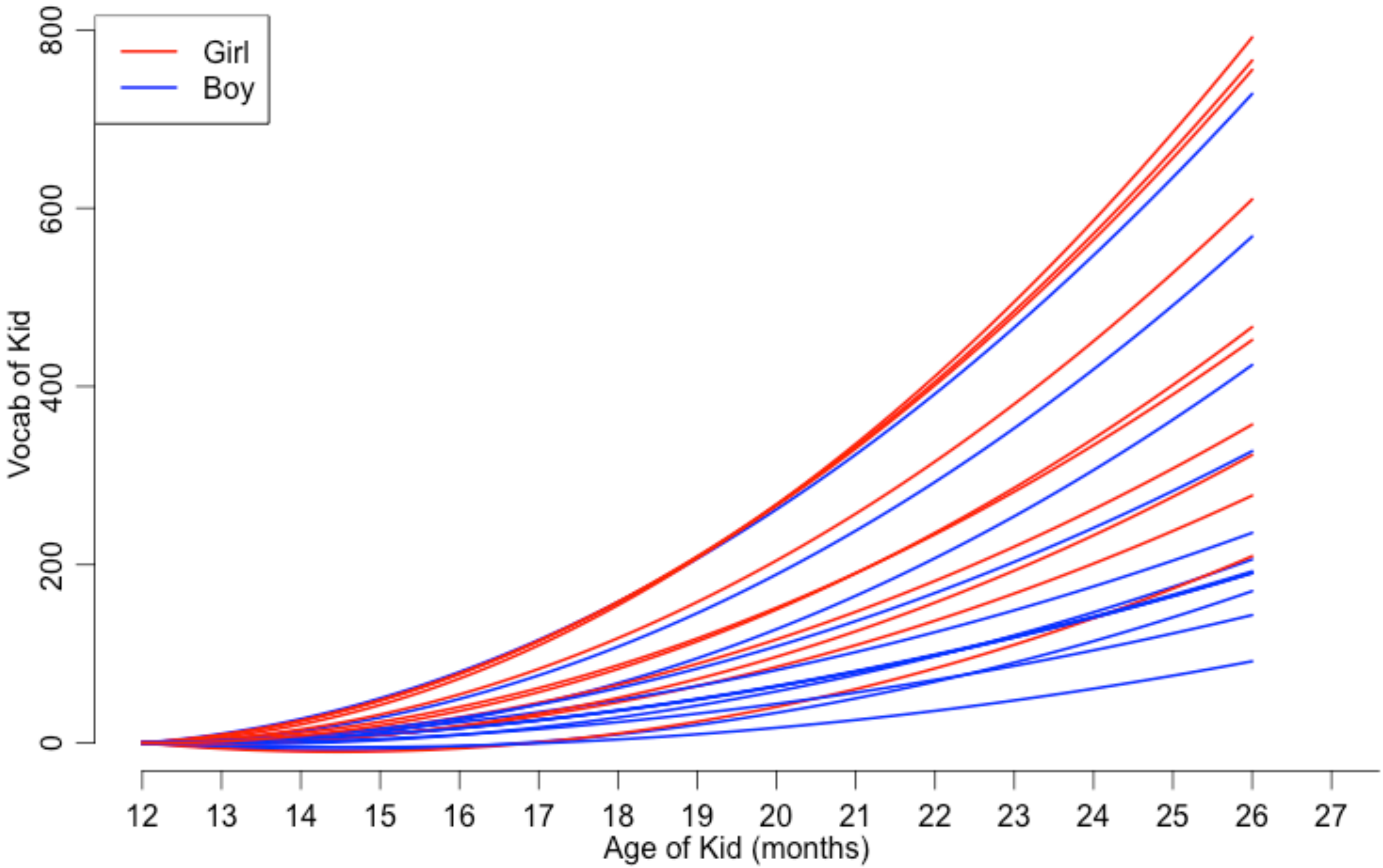
```
> M4 = lmer( vocab ~ 0 + age12sq + age12sq:group +
age12sq:sex + age12sq:logmom + age12sq:sex:group +
age12sq:logmom:group + (0 + age12 + age12sq|pers), data=dat )
> display( M4 )

                     coef.est coef.se
age12sq              -5.43     2.98
age12sq:group         1.64     5.44
age12sq:sex           0.37     0.46
age12sq:logmom        0.97     0.38
age12sq:group:sex     0.53     0.68
age12sq:group:logmom -0.35     0.69

Error terms:
 Groups   Name    Std.Dev. Corr
 pers     age12    3.98
          age12sq  0.52    1.00
 Residual         30.23
---
number of obs: 104, groups: pers, 22
AIC = 1098.7, DIC = 1075
deviance = 1076.8
```

Can we explain our acceleration?

Gender Differences



Variance Reduction due to gender?

```
> M3.full = lmer( vocab ~ 0 + age12sq + age12sq:group +
              (0 + age12 + age12sq|pers), data=dat )

> M4 = lmer( vocab ~ 0 + age12sq + age12sq:group +
              age12sq:sex + age12sq:logmom +
              age12sq:sex:group + age12sq:logmom:group +
             (0 + age12 + age12sq|pers), data=dat )


> diag( VarCorr( M3.full )$pers )
  age12 age12sq
  15.70    0.53

> diag( VarCorr( M4 )$pers )
  age12 age12sq
  15.85    0.27
```

? How much did we explain the variance in growth rate?