Dynamic Predictors

Dynamic Predictor

- Any predictor *other than a time predictor* that changes over time
- LMER model can include both time and dynamic predictors
  - Important that they are not redundant
- When dynamic predictor is not a proxy for time
  - Treat the predictor as dynamic or static?
    - Static predictors only account for between-subject variability
    - Dynamic predictors can account for between-subject *and* within-subject variability
    - Between-subjects variance exits if the overall level varies among the subjects
    - Within-subjects variance exits if there is variability over time

- Consider MPLS data set
  - Risk originally measured as dynamic predictor
  - Analyzed thus far as a static predictor
  - Theoretical justification to treat as static
    - Subjects experiencing even a single year in either of the two disadvantaged groups (poverty, HHM) should be denoted the particular status throughout the duration of the study
    - Hypothesized that between-subject differences are most important, and that within-subject effects are trifling for the time period considered
  - Empirical justification has also been found
    - Membership in the risk groups tends to be steady for students in the MPLS school district, and the within-subjects variance is considered to be negligible

- When the researcher is convinced that important within-subject variability will be accounted for by a dynamic predictor (or to investigate the possibility) number of models that can be considered
- Primary concern is whether a time predictor will be included along with the dynamic predictor
- Choice of model depends on the research question(s) to be addressed by the analysis.
    - Including only the dynamic predictor (excluding time predictors)
    - Including the dynamic predictor and time predictor(s) as main effects (single effects)
    - Including a dynamic predictor by time predictor interaction along with the main effects
-

- Dynamic predictors can be quantitative or categorical
- Categorical dynamic predictor is introduced in these notes
  - Generalization to a quantitative predictor is hopefully straight-forward
- Dynamic predictor is based on a hypothetical study of the effect of financial incentives on student achievement
  - Evidence suggests that financial incentives improve academic achievement and related factors, such as school attendance
  - Important question in the study of incentives is what occurs when incentives are discontinued

- Suppose a researcher carries out the following study
  - Over 4 years (grades), each student is assigned two instances of financial incentive, and two instances of no incentive.
  - To control for order effects, the order of the conditions is randomized.
  - At the beginning of each year of the study, the students are told they will either receive a cash payment at the end of the year to motivate improved performance, or they are informed they will not receive a cash payment at the end of the year.
  - In the data frame, the dynamic predictor is coded as 1 = received cash payment, and 0 = did not receive cash payment

```
   subid incent.5 incent.6 incent.7 incent.8
1      1        0        1        0        1
2      2        0        1        1        0
3      3        0        1        1        0
4      4        1        0        0        1
5      5        1        1        0        0
6      6        1        0        1        0
7      7        0        1        1        0
8      8        1        0        0        1
9      9        0        0        1        1
10    10        0        1        0        1
11    11        1        1        0        0
12    12        0        0        1        1
13    13        0        1        0        1
14    14        1        1        0        0
15    15        0        0        1        1
16    16        1        0        1        0
17    17        1        0        0        1
18    18        0        1        0        1
19    19        1        1        0        0
20    20        0        1        1        0
21    21        1        0        1        0
22    22        0        1        1        0
```

Read in `Minneapolis-Long2.csv` data

```
> mpls.l <- read.csv(file = "Minneapolis-Long2.csv")
```

Create `dadv` and `ethW` predictors

```
> mpls.l <- read.csv(file = "Minneapolis-Long2.csv")
> mpls.l$grade5 <- mpls.l$grade - 5
> mpls.l$dadv <- as.factor(
    ifelse(mpls.l$risk == "ADV", "ADV", "DADV"))
> mpls.l$ethW <- as.factor(
    ifelse(mpls.l$eth == "Whi", "W", "NW"))
```

Read in `Financial-Incentive.txt` data

```
> incentive <- read.table( "Financial-Incentive.txt",
    header = TRUE)
```

Reshape incentive data to long format

```
> incentive.l <- reshape(
   incentive,
   varying = c("incent.5", "incent.6", "incent.7", "incent.8"),
   timevar = "grade",
   times = 5:8,
   idvar = "subid",
   direction = "long"
   )
```

Merge Minneapolis and Incentive data

```
> MPLS <- merge(mpls.l, incentive.l, by = c("subid", "grade"))
```

Order and browse data

```
> MPLS <- arrange(MPLS, subid)
> head(MPLS, n = 8)

  subid grade X risk gen eth ell sped  att read grade5 dadv ethW incent
1     1     5 1  HHM   F Afr   0     N 0.94  172      0 DADV   NW      0
2     1     6 2  HHM   F Afr   0     N 0.94  185      1 DADV   NW      1
3     1     7 3  HHM   F Afr   0     N 0.94  179      2 DADV   NW      0
4     1     8 4  HHM   F Afr   0     N 0.94  194      3 DADV   NW      1
5     2     5 5  HHM   F Afr   0     N 0.91  200      0 DADV   NW      0
6     2     6 6  HHM   F Afr   0     N 0.91  210      1 DADV   NW      1
7     2     7 7  HHM   F Afr   0     N 0.91  209      2 DADV   NW      1
8     3     5 9  HHM   M Afr   0     N 0.97  191      0 DADV   NW      0
```

Incentive Variable

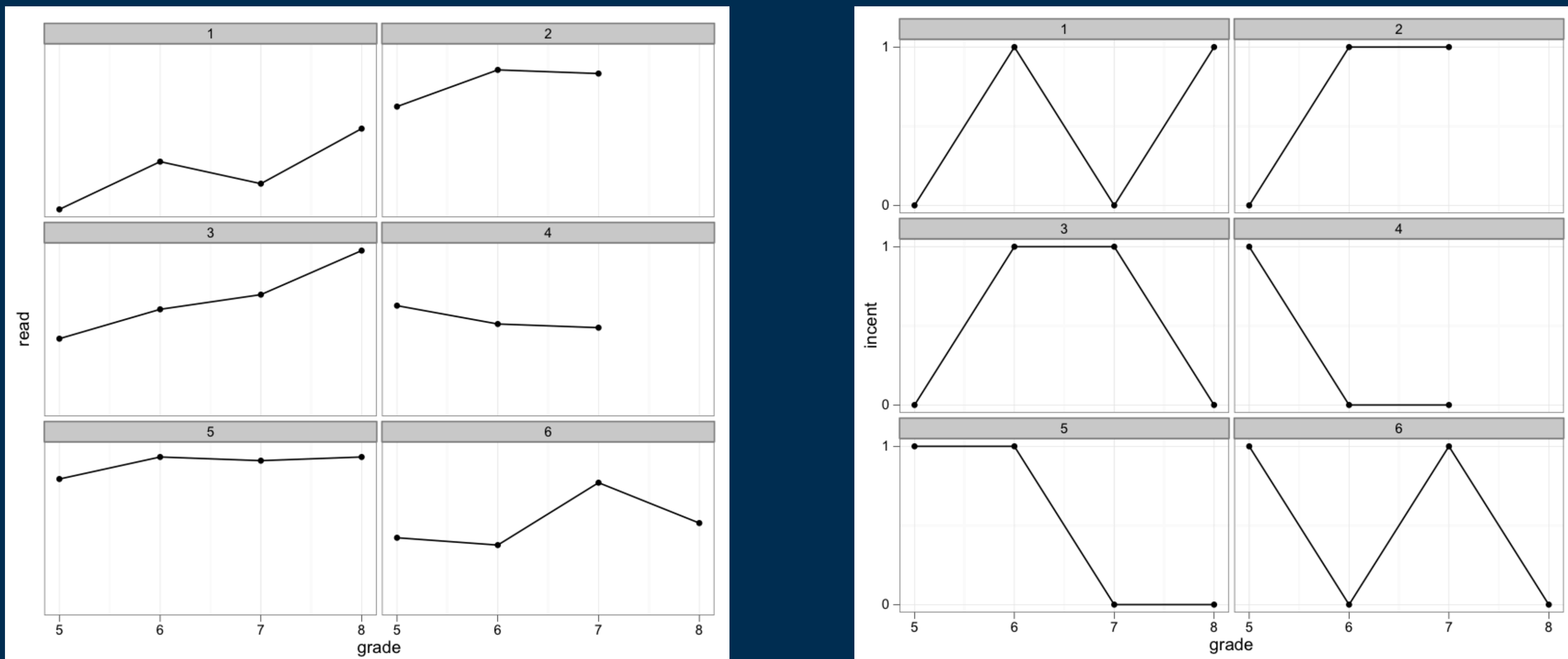
- Output shows that the incentive variable values change over time
- Pattern of 0s and 1s is not the same for the two subjects depicted in the output
- To better understand the incentive variable and the research questions related to reading, consider a graph of `read` and `incent` as a function of `grade`
  - We plot only the first 6 subjects for brevity

```r
## Select the first 6 subjects
> plotdata <- na.omit(MPLS[MPLS$subid < 7, ])

## Create the read plot
> ggplot(data = plotdata, aes(x = grade, y = read,
      group = subid)) +
    geom_line() +
    geom_point() +
    facet_wrap(~ subid, ncol = 2) +
    scale_x_continuous(breaks = 5:8) +
    scale_y_continuous(breaks = 0:1) +
    theme_bw()

## Create the incent plot
> ggplot(data = plotdata, aes(x = grade, y = incent, group = subid)) +
    geom_line() +
    geom_point() +
    facet_wrap(~ subid, ncol = 2) +
    scale_x_continuous(breaks = 5:8) +
    scale_y_continuous(breaks = 0:1) +
    theme_bw()
```

Incentive Variable



- There are some missing data values
  - See panels for subjects 2 and 4
- Due to the nature of the research design, each subject has a nonlinear trend for financial incentive

Dynamic Predictor as a Single Effect

- Dynamic predictor model to be used depends on the research question
  - Simplest model includes the dynamic predictor as a main effect (no time predictor) in the LMER model
  - Expected value of this model is

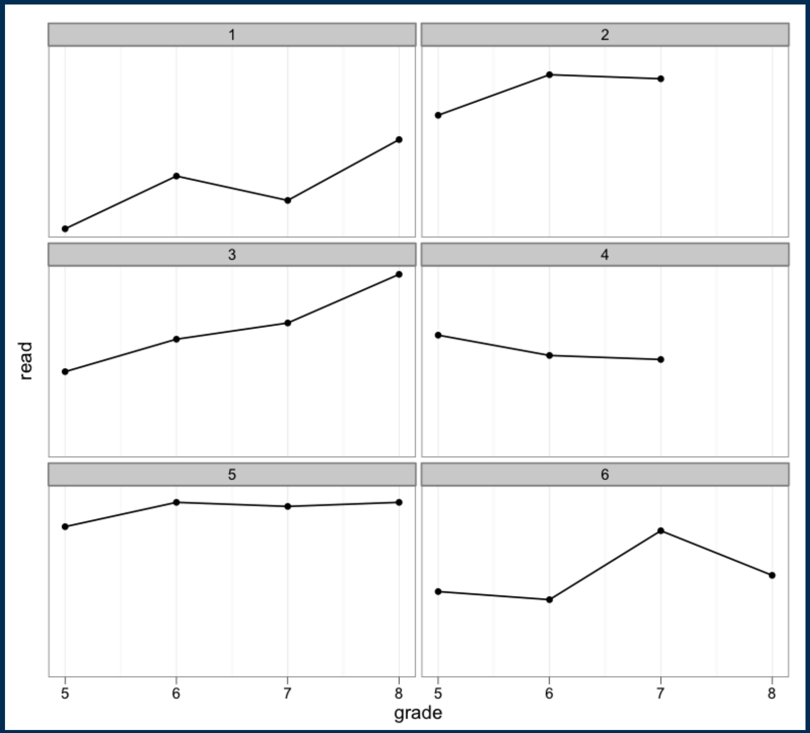
$$E(y_{ij}) = \beta_0 + \beta_1(\texttt{incent}_{ij})$$

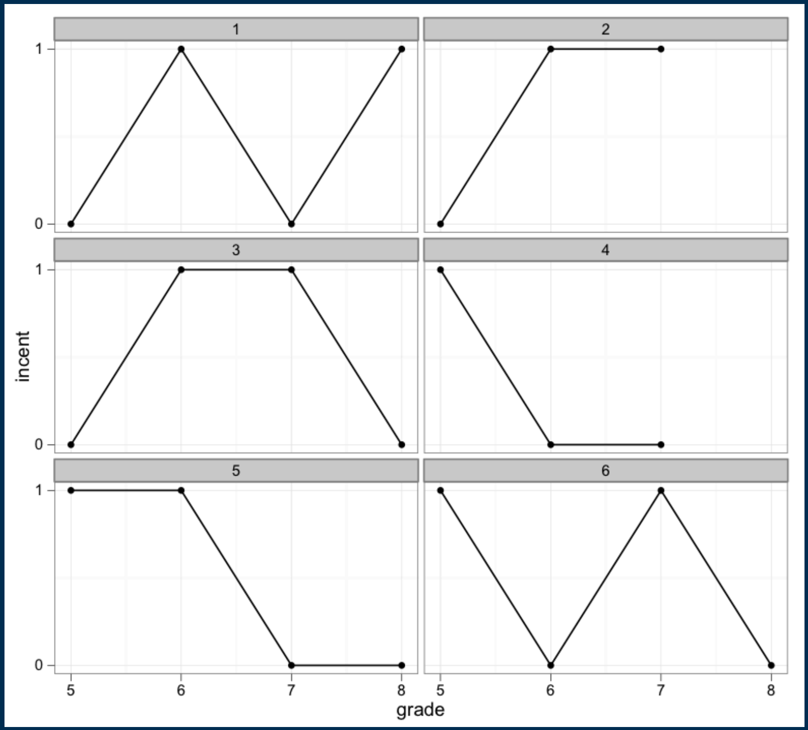
- At each time point, mean reading is predicted by financial incentive (measured at the same time point)
- Goal of this model is to study how the two variables covary over time (*longitudinal covariation*)

- To gain additional insight, consider expressing change in the expected response from wave $j - 1$ to wave j
    - Expected value of this difference is

$$E(y_{ij}) - E(y_{ij-1}) = \beta_0 + \beta_1(\texttt{incent}_{ij}) - [\beta_0 + \beta_1(\texttt{incent}_{ij-1})]$$
$$= \beta_1(\texttt{incent}_{ij} - \texttt{incent}_{ij-1})$$

- β_1 indexes the relationship between the change in the mean response and the change in the dynamic predictor
- In terms of the earlier figure the model allows the researcher to investigate if the changes in the right-hand graphs are associated with the average changes in the corresponding left-hand graphs
    - Whether change in financial incentive is associated with mean change in reading scores





- For at least some subjects, change in the two variables might be related
- For the first subject, a change from 0 to 1 in `incent` is associated with an increase in `read`, and a change from 1 to 0 in `incent` is associated with a decrease in `read`
- Similar relationships appear for some of the other subjects

- If β_1 is positive, then the direction of change in the response is the same as in the predictor
  - This can be either an increase or decrease, and the change can be linear or nonlinear
- If β_1 is negative, then the direction of change in the response is the opposite of the predictor, as when the response increases, but the predictor decreases.
- Dynamic predictors are included in `lmer()` just like any other predictor
- Similar to time predictors, random effects can be associated with dynamic predictors
  - Inclusion of a random effect for a dynamic predictor is desirable when one wants to allow for individual differences in the longitudinal covariation of the two variables

- Consider estimating the model using a single random effect
- For comparison, an intercept-only model is also estimated

```
## Intercept-only
> dyn.0 <- lmer(read ~ 1 + (1 | subid), data = MPLS,
    REML = FALSE)

## Dynamic predictor
> dyn.1 <- lmer(read ~ incent + (1 | subid), data = MPLS,
    REML = FALSE)

## Comparison of fit
> anova(dyn.0, dyn.1)

Data: MPLS
Models:
dyn.0: read ~ 1 + (1 | subid)
dyn.1: read ~ incent + (1 | subid)
      Df    AIC    BIC  logLik Chisq Chi Df Pr(>Chisq)
dyn.0  3 630.53 637.68 -312.26
dyn.1  4 627.55 637.07 -309.77 4.985      1    0.02557 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Consider quadratic model with two random effects

```
> print(aictab(list(dyn.0, dyn.1), c("Intercept", "Dynamic")),
    LL = FALSE)

Model selection based on AICc :

          K   AICc Delta_AICc AICcWt Cum.Wt
Dynamic   4 628.08       0.00    0.8    0.8
Intercept 3 630.85       2.77    0.2    1.0
```

- Output indicates LRT-p is relatively small ($p < 0.03$)
- Weight of evidence of the model with `incent` is relatively large ($W_2 = 0.8$)

- Equation makes no statement about the nature of the longitudinal trajectories of reading or financial incentive
- Time is expressed completely through the dynamic predictor
- In the case of `incent`, this is perhaps a bit awkward, as the dynamic predictor is a binary variable, allowing only a coarse type of change curve
- The effect of duration on the response variable shows up completely through its association with the dynamic predictor
- In general, the shape of the trajectories of the variables can be inferred from graphs
    - But there are no terms in the model that are informative in this regard

Dynamic Predictor with a Time Variable

- When the shape of the response trajectory is of interest, then a time predictor can be included along with the dynamic predictor
- In addition to main effects, the interaction among the time predictor and dynamic predictor can be included in the model
    - This interaction model is similar to what was encountered with time predictors and static predictors in past units
    - Fundamental difference is that the effects have a within-subjects interpretation rather than a between-subjects interpretation
- Consider the interaction model using `grade` and `incent` as predictors

$$y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})(\texttt{grade}_{ij}) + \beta_2(\texttt{incent}_{ij}) + \beta_3(\texttt{grade}_{ij})(\texttt{incent}_{ij}) + \epsilon_{ij}$$

Dynamic Predictor with a Time Variable

- Similar to static predictor models, the interaction allows one to examine if linear change in reading depends on financial incentive
- Because `incent` is a binary predictor, this is made explicit by substituting the values of 0 and 1 in the model

For `incent` = 0

$$
\begin{aligned}
E(y_{ij}) &= \beta_0 + \beta_1(\texttt{grade}_{ij}) + \beta_2(0) + \beta_3(\texttt{grade}_{ij})(0) \\
&= \beta_0 + \beta_1(\texttt{grade}_{ij})
\end{aligned}
$$

For `incent` = 1

$$
\begin{aligned}
E(y_{ij}) &= \beta_0 + \beta_1(\texttt{grade}_{ij}) + \beta_2(1) + \beta_3(\texttt{grade}_{ij})(1) \\
&= (\beta_0 + \beta_2) + (\beta_1 + \beta_3)(\texttt{grade}_{ij})
\end{aligned}
$$

- If financial incentive were a static predictor, then the equations would reflect between-group differences
  - Students in the no incentive group would have a mean predicted curve with intercept β_0 and slope β_1
  - Students in the incentive group would have a mean predicted curve with intercept $(\beta_0 + \beta_2)$ and slope $(\beta_1 + \beta_3)$
  - The between-subjects interpretation focuses on the contrast between different groups of students
  - Fundamental difference is that the effects have a within-subjects interpretation rather than a between-subjects interpretation

- Financial incentive is a dynamic predictor, and here there is a within-subjects interpretation
  - This means that as the same students change from no incentive ($\texttt{incent}_{ij} = 0$) to incentive ($\texttt{incent}_{ij} = 1$) the linear curve of reading changes its intercept by β_2 and its slope by β_3
  - The within-subjects interpretation focuses on the average change in reading trajectory, for changes in incentive levels experienced by all the students.
- The `lmer()` syntax is identical to the case when the predictor is static

```
> dyn.int <- lmer(read ~ grade * incent + (grade | subid),
    data = MPLS, REML = FALSE)
> print(dyn.int, corr = FALSE)

Linear mixed model fit by maximum likelihood
Formula: read ~ grade * incent + (grade | subid)
   Data: MPLS
   AIC   BIC logLik deviance REMLdev
 560.7 579.7 -272.3    544.7   535.7
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subid    (Intercept) 861.2544 29.3471
          grade         8.4000  2.8983  -0.816
 Residual               9.8854  3.1441
Number of obs: 80, groups: subid, 22

Fixed effects:
             Estimate Std. Error t value
(Intercept)  177.8611     7.2279  24.608
grade          4.4689     0.8372   5.338
incent        -0.7439     6.0771  -0.122
grade:incent   0.8810     0.9472   0.930
```

- Consider the estimates from the fixed effects table in the output, substituted in the fitted value equations

$$\hat{y}_{ij} = 177.86 + 4.47(\texttt{grade}_{ij})$$ For no financial incentive (`incent` = 0)

$$\hat{y}_{ij} = 177.12 + 5.35(\texttt{grade}_{ij})$$ For financial incentive (`incent` = 1)

- For the sample, the reading slope increases by 0.88, when the incentive is introduced
- Likewise, the reading slope decreases by 0.88, when the incentive is withdrawn
- The *t*-ratio for the interaction is rather small (t = 0.93), meaning the change in the slope may not be statistically reliable

```
## Main effects model
> dyn.main <- lmer(read ~ grade + incent + (grade | subid),
    data = MPLS, REML = FALSE)

## Compare models (LRT)
> anova(dyn.main, dyn.int)
Data: MPLS
Models:
dyn.main: read ~ grade + incent + (grade | subid)
dyn.int: read ~ grade * incent + (grade | subid)
         Df    AIC    BIC  logLik  Chisq Chi Df Pr(>Chisq)
dyn.main  7 559.48 576.15 -272.74
dyn.int   8 560.66 579.72 -272.33 0.8184      1     0.3657

## Compare models (AICc)
> print(aictab(list(dyn.main, dyn.int), c("Main Effects",
"Interaction")), LL = FALSE)

Model selection based on AICc :

             K   AICc Delta_AICc AICcWt Cum.Wt
Main Effects 7 561.03       0.00    0.7    0.7
Interaction  8 562.69       1.65    0.3    1.0
```

- According to the LRT and the AICc, the interaction is not statistically reliable
- The output for the main effects model is printed

```
## Model results
> print(dyn.main, cor = FALSE)

Linear mixed model fit by maximum likelihood
Formula: read ~ grade + incent + (grade | subid)
   Data: MPLS
   AIC   BIC logLik deviance REMLdev
 559.5 576.2 -272.7    545.5   538.3
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subid    (Intercept) 846.6802 29.0978
          grade         7.9918  2.8270  -0.816
 Residual              10.3160  3.2118
Number of obs: 80, groups: subid, 22

Fixed effects:
            Estimate Std. Error t value
(Intercept) 175.2881     6.6435  26.385
grade         4.8740     0.7062   6.902
incent        4.8391     0.8663   5.586
```

- Similar to a main effects model with a static predictor, the linear curves for the levels of financial incentive are parallel, with an estimated vertical shift of $\beta_2 = 4.84$
- Since financial incentive is a dynamic predictor, there is a within-subjects interpretation
    - As financial incentive increases, there is a predicted intercept increase of $\beta_2 = 4.84$, and as financial incentive decreases, there is a predicted decrease of the same amount
    - The *t*-ratio for `incent` is almost as large as that for `grade`, meaning the financial incentive effect is about as strong as the duration effect

- To this point, the discussion has focused on a categorical dynamic predictor, but the same ideas generalize to quantitative dynamic predictors
    - Conditioning on a quantitative dynamic predictor is similar to a categorical predictor, except the number of level values can be numerous, and perhaps unique
    - Thus, with quantitative dynamic predictors, it is often impractical to write conditional equations
    - As an alternative, one can estimate the parameters and produce fitted equations using, say, the empirical quartiles of the quantitative dynamic predictor
- Static predictors can be introduced along with dynamic predictors and time predictors
- The models can be relatively complex, especially if an interaction among the time predictor, dynamic predictor, and static predictor is included

$$
\begin{aligned}
y_{ij} = {} & (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})(\texttt{grade}_{ij}) + \beta_2(\texttt{incent}_{ij}) \\
& + \beta_3(\texttt{grade}_{ij} \cdot \texttt{incent}_{ij}) \\
& + \beta_4(\texttt{dadv}_i) + \beta_5(\texttt{dadv}_i \cdot \texttt{grade}_{ij}) \\
& + \beta_6(\texttt{dadv}_i \cdot \texttt{incent}_{ij}) \\
& + \beta_7(\texttt{dadv}_i \cdot \texttt{grade}_{ij} \cdot \texttt{incent}_{ij}) \\
& + \epsilon_{ij}
\end{aligned}
$$

- The triple interaction in the last line of the equation warrants further comment

- One means of interpreting the triple interaction is to express the equation conditional on dadv

For dadv = 0

$$E(y_{ij}) = \begin{cases} \beta_0 + \beta_1(\texttt{grade}_{ij}) & \text{if } \texttt{incent}_{ij} = 0; \\ (\beta_0 + \beta_2) + (\beta_1 + \beta_3)(\texttt{grade}_{ij}) & \text{if } \texttt{incent}_{ij} = 1 \end{cases}$$

The equations show that as the subjects in the advantaged group increase from no incentive to incentive or vice versa, there is a change in the intercept of the reading curve of β_2, and a change in the slope of β_3. This is a within-subjects interpretation, as focus is on the change of incentive status for the same subjects over time.

- One means of interpreting the triple interaction is to express the equation conditional on dadv

For dadv = 1

$$E(y_{ij}) = \begin{cases} (\beta_0 + \beta_4) + (\beta_1 + \beta_5)(\mathtt{grade}_{ij}) & \text{if } \mathtt{incent}_{ij} = 0; \\ (\beta_0 + \beta_2 + \beta_4 + \beta_6) + (\beta_1 + \beta_3 + \beta_5 + \beta_7)(\mathtt{grade}_{ij}) & \text{if } \mathtt{incent}_{ij} = 1 \end{cases}$$

The equations show that as the subjects in the advantaged group increase from no incentive to incentive or vice versa, there is a change in the intercept of the reading curve of $(\beta_2 + \beta_6)$ and a change in the slope of $(\beta_3 + \beta_7)$. Again, this is a within-subjects interpretation.

- Statements regarding between-group differences are also possible
  - The equations indicate the risk groups have different parameters for levels of incentive
  - For no incentive, the difference in the mean intercept of the risk groups is β_4
  - Similar between-groups comparisons for the slope can also be made
- Finally, it should be noted that a limited number of the possible dynamic predictor models were presented
  - Alternatives might be considered when required by the research questions or subject matter
  - For example, a class of models popular in some areas of research, such as business and economics, are lagged models that essentially use earlier dynamic scores to predict later response scores