Replication

In this notebook, we replicate the **Impacts on Target Outcomes** section of the chosen paper. We restrict our view to the key outcomes of rates of attendance, truancy, and suspension, along with academic achievement as measured by reading and math scores on standardized examinations.

```r
# data loading utilities
source("load.R")
```

Randomization Check

Load in baseline covariates about the schools and random assignments.

```r
school_df <- load_school_info()

baseline_covariates <- c(
  "FRL", "PUPTCH", "PCT_BL", "PCT_HS", "PCT_WH",
  "SMALLSCHOOL", "LARGESCHOOL", "GRDRNG_ELEM", "GRDRNG_HS"
  )
```

```r
library(modelsummary)
```

```
## 'modelsummary' 2.0.0 now uses 'tinytable' as its default table-drawing
##   backend. Learn more at: https://vincentarelbundock.github.io/tinytable/
##
## Revert to 'kableExtra' for one session:
##
##   options(modelsummary_factory_default = 'kableExtra')
##   options(modelsummary_factory_latex = 'kableExtra')
##   options(modelsummary_factory_html = 'kableExtra')
##
## Silence this message forever:
##
##   config_modelsummary(startup_message = FALSE)
```

```r
to_display <- school_df %>%
  select(all_of(baseline_covariates)) %>%
  rename(
      "Proportion Eligible for Free/Reduced Lunch" = FRL,
      "Student-Teacher Ratio" = PUPTCH,
      "% Black" = PCT_BL,
      "% Hispanic" = PCT_HS,
      "% White" = PCT_WH,
      "Small 6th and 7th Grades (less than 200 students)" = SMALLSCHOOL,
      "Large 6th and 7th Grades (more than 700 students)" = LARGESCHOOL,
      "Also an Elementary School" = GRDRNG_ELEM,
      "Also a High School" = GRDRNG_HS,
  )

# Create the table object
make_tbl <- function(output) {
```

Table 1: Summary Statistics

	Mean	Std. Dev.
Proportion Eligible for Free/Reduced Lunch	0.596	0.277
Student-Teacher Ratio	16.008	3.303
% Black	7.318	9.763
% Hispanic	47.564	28.692
% White	37.901	29.858
Small 6th and 7th Grades (less than 200 students)	0.304	0.465
Large 6th and 7th Grades (more than 700 students)	0.196	0.401
Also an Elementary School	0.261	0.444
Also a High School	0.087	0.285

```r
  datasummary_balance(
    ~ 1,
    data = to_display,
    title = "Summary Statistics",
    fmt = 3,    # number of decimal places
    # notes = "Note: Standard deviations in parentheses.",
    output = output
  )
}

make_tbl("kableExtra")
```

```r
# make_tbl("doc/baseline_covariates.tex")
```

We'll perform a block balanced randomization check.

```r
baseline_covariates <- c(
  "FRL", "PUPTCH", "PCT_BL", "PCT_HS", "PCT_WH",
  "SMALLSCHOOL", "LARGESCHOOL", "GRDRNG_ELEM", "GRDRNG_HS"
  )
formula <- as.formula(paste(
  "TREATMENT ~", paste(baseline_covariates, collapse=" + "), " + factor(COHORT)"
  ))

g <- lm(formula, data=school_df)

f_stat <- summary(g)$fstatistic
p_val <- pf(f_stat["value"], f_stat["numdf"], f_stat["dendf"], lower.tail=FALSE)

print(sprintf(
    "Joint F-test of covariate balance: F(%d, %d) = %.3f (p = %.3f)",
    f_stat["numdf"],
    f_stat["dendf"],
    f_stat["value"],
    p_val
))
```

```
## [1] "Joint F-test of covariate balance: F(10, 35) = 0.531 (p = 0.856)"
```

Great! We can be reasonably confident that the randomization procedure was performed correctly then, and that there are no strange inaccuracies from the matching process.

School-Level: Attendance, Truancy, and Suspension

Load in the data.

```
# attendance, truancy, suspensions
school_level_outcomes <- load_school_outcomes()
```

View a few random rows for clarity.

Attendance is measured as the total number of student days attended, divided by the total number of days school was in session. It's unclear exactly what this means, but we'll assume that this is the average attendance rate across all students in the school since all values are < 1.

Truancy is defined as the total number of student unexcused absences, divided by the total number of days school was in session. Again unclear, but we'll assume that this is the average truancy rate across all students in the school.

Suspension rates are defined as the number of suspensions divided by the total number of students in the school.

```
set.seed(0527)
# Rename columns for compact display:
col_mapping <- c(
  COHORT = "COHORT",
  Z = "TREATMENT",
  PID = "PID",
  # Attendance years 0-2
  A0 = "ATTEND_Y0",
  A1 = "ATTEND_Y1",
  A2 = "ATTEND_Y2",
  # Truancy years 0-2
  T0 = "TRUANCY_Y0",
  T1 = "TRUANCY_Y1",
  T2 = "TRUANCY_Y2",
  # Suspensions years 1-2
  S_IN1 = "SUS_IN_SCH_YR1",
  S_IN2 = "SUS_IN_SCH_YR2",
  S_OUT1 = "SUS_OUT_SCH_YR1",
  S_OUT2 = "SUS_OUT_SCH_YR2"
)
df_to_dply <- school_level_outcomes %>%
  select(all_of(unname(col_mapping))) %>%
  rename(!!!col_mapping)

knitr::kable(
  df_to_dply %>% sample_n(5),
  digits=3,
  caption="School-Level Outcome Information"
  )
```

Table 2: School-Level Outcome Information

COHORT	Z	PID	A0	A1	A2	T0	T1	T2	S_IN1	S_IN2	S_OUT1	S_OUT2
1	1	pair5	0.938	0.933	0.925	0.011	0.010	0.013	34.954	26.911	20.517	15.764
2	0	pair18	0.934	0.935	0.929	0.018	0.017	0.020	34.984	24.256	26.133	27.231
1	0	pair2	0.969	0.928	0.943	0.005	0.045	0.027	2.765	9.009	7.373	8.108
2	0	pair11	0.910	0.909	0.904	0.041	0.031	0.034	13.153	29.348	38.298	44.783
1	0	pair1	0.944	0.936	0.935	0.013	0.018	0.018	10.454	10.870	3.156	5.254

Attendance

First, we replicate their analysis on attendance. The authors run two separate models for each outcome variable, one for 1-year effects and one for 2-year effects.

```
attendance <- load_school_level_outcome("attendance")

formula <- as.formula(paste(
  "ATTEND_Y1 ~",
  "TREATMENT + ATTEND_Y0 + ",
  # postpend _Y1 since each covariate is measured in both years
  paste(paste0(baseline_covariates, "_Y1"), collapse=' + '),
  "+",
  paste(paste0('PAIR', 2:24), collapse=" + ")
))

g1 <- lm(formula, attendance)

summary(g1)
```

```
##
## Call:
## lm(formula = formula, data = attendance)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.022062 -0.003561  0.000000  0.003561  0.022062
##
## Coefficients: (1 not defined because of singularities)
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     5.680e-01  2.427e-01   2.340   0.0374 *
## TREATMENT       5.949e-03  5.227e-03   1.138   0.2773
## ATTEND_Y0       3.571e-01  1.860e-01   1.920   0.0790 .
## FRL_Y1          3.693e-04  5.752e-02   0.006   0.9950
## PUPTCH_Y1       4.315e-05  1.543e-03   0.028   0.9782
## PCT_BL_Y1      -6.526e-04  1.165e-03  -0.560   0.5857
## PCT_HS_Y1       1.248e-04  8.101e-04   0.154   0.8801
## PCT_WH_Y1       3.349e-04  1.204e-03   0.278   0.7856
## SMALLSCHOOL_Y1  5.990e-03  1.127e-02   0.532   0.6047
## LARGESCHOOL_Y1  7.544e-04  1.252e-02   0.060   0.9529
## GRDRNG_ELEM_Y1  1.665e-02  1.253e-02   1.329   0.2085
## GRDRNG_HS_Y1   -2.547e-02  1.427e-02  -1.785   0.0996 .
## PAIR2           5.994e-03  2.646e-02   0.227   0.8246
```

```
## PAIR3              6.939e-03  1.516e-02   0.458   0.6553
## PAIR4                     NA         NA      NA        NA
## PAIR5             -2.505e-03  1.508e-02  -0.166   0.8708
## PAIR6              3.422e-03  2.741e-02   0.125   0.9027
## PAIR7             -2.721e-03  1.976e-02  -0.138   0.8928
## PAIR8             -4.721e-03  2.813e-02  -0.168   0.8695
## PAIR9              1.060e-04  2.618e-02   0.004   0.9968
## PAIR10            -5.699e-03  1.827e-02  -0.312   0.7604
## PAIR11            -2.470e-03  2.098e-02  -0.118   0.9082
## PAIR12             2.954e-02  2.965e-02   0.996   0.3388
## PAIR13            -8.577e-04  2.870e-02  -0.030   0.9766
## PAIR14            -2.136e-02  3.063e-02  -0.698   0.4987
## PAIR15             1.419e-02  2.691e-02   0.527   0.6076
## PAIR16             8.032e-03  2.313e-02   0.347   0.7344
## PAIR17             2.406e-03  2.671e-02   0.090   0.9297
## PAIR18             9.776e-03  2.775e-02   0.352   0.7307
## PAIR19            -4.001e-03  1.431e-02  -0.280   0.7845
## PAIR20            -7.528e-03  1.375e-02  -0.548   0.5940
## PAIR21            -2.228e-04  2.509e-02  -0.009   0.9931
## PAIR22             5.220e-03  1.760e-02   0.297   0.7719
## PAIR23            -8.499e-03  2.475e-02  -0.343   0.7372
## PAIR24             3.015e-02  2.818e-02   1.070   0.3056
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01256 on 12 degrees of freedom
## Multiple R-squared:  0.9273, Adjusted R-squared:  0.7272
## F-statistic: 4.636 on 33 and 12 DF,  p-value: 0.00352
```

Nothing significant achieved, matching the paper. We do obtain a different coefficient value, however, interestingly. Our $\beta_1 = 0.006$ and in the paper they obtain 0.01, though this could be due to rounding to two significant figures in the display.

```
formula <- as.formula(paste(
  "ATTEND_Y2 ~",
  "TREATMENT + ATTEND_Y0 + ",
  # postpend _Y1 since each covariate is measured in both years
  paste(paste0(baseline_covariates, "_Y1"), collapse=' + '),
  "+",
  paste(paste0('PAIR', 2:24), collapse=" + ")
))

g2 <- lm(formula, attendance)

summary(g2)
```

```
##
## Call:
## lm(formula = formula, data = attendance)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.013550 -0.003965  0.000000  0.003965  0.013550
```

```
##
## Coefficients: (2 not defined because of singularities)
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6696744  0.2693538   2.486   0.0302 *
## TREATMENT        0.0020089  0.0050250   0.400   0.6970
## ATTEND_Y0        0.4241224  0.2037740   2.081   0.0616 .
## FRL_Y1          -0.0854956  0.0616042  -1.388   0.1927
## PUPTCH_Y1        0.0013926  0.0014945   0.932   0.3714
## PCT_BL_Y1       -0.0012286  0.0011287  -1.089   0.2996
## PCT_HS_Y1       -0.0010006  0.0007930  -1.262   0.2331
## PCT_WH_Y1       -0.0014632  0.0012106  -1.209   0.2521
## SMALLSCHOOL_Y1   0.0138142  0.0108923   1.268   0.2309
## LARGESCHOOL_Y1  -0.0155274  0.0120348  -1.290   0.2234
## GRDRNG_ELEM_Y1  -0.0140820  0.0123388  -1.141   0.2780
## GRDRNG_HS_Y1    -0.0157607  0.0137173  -1.149   0.2749
## PAIR2           -0.0030258  0.0256109  -0.118   0.9081
## PAIR3            0.0081262  0.0146036   0.556   0.5890
## PAIR4                   NA         NA      NA       NA
## PAIR5           -0.0021100  0.0145273  -0.145   0.8871
## PAIR6            0.0114808  0.0264381   0.434   0.6725
## PAIR7            0.0001714  0.0193958   0.009   0.9931
## PAIR8            0.0056130  0.0270522   0.207   0.8394
## PAIR9            0.0065397  0.0253063   0.258   0.8009
## PAIR10          -0.0005878  0.0178023  -0.033   0.9743
## PAIR11           0.0012257  0.0201867   0.061   0.9527
## PAIR12           0.0244700  0.0285356   0.858   0.4094
## PAIR13           0.0168388  0.0276693   0.609   0.5552
## PAIR14                  NA         NA      NA       NA
## PAIR15          -0.0013405  0.0261690  -0.051   0.9601
## PAIR16          -0.0278480  0.0234018  -1.190   0.2591
## PAIR17           0.0262343  0.0258300   1.016   0.3316
## PAIR18          -0.0012384  0.0270564  -0.046   0.9643
## PAIR19          -0.0081095  0.0139072  -0.583   0.5716
## PAIR20          -0.0067912  0.0134128  -0.506   0.6226
## PAIR21           0.0024413  0.0244538   0.100   0.9223
## PAIR22          -0.0112745  0.0180292  -0.625   0.5445
## PAIR23           0.0115555  0.0239779   0.482   0.6393
## PAIR24          -0.0172436  0.0281312  -0.613   0.5524
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01207 on 11 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.9051, Adjusted R-squared:  0.6291
## F-statistic:  3.28 on 32 and 11 DF,  p-value: 0.02028
```

Again, we match the paper and find no significant effects.

Truancy

```
truancy <- load_school_level_outcome("truancy")
```

6

```r
formula <- as.formula(paste(
  "TRUANCY_Y1 ~",
  "TREATMENT + TRUANCY_Y0 + ",
  # postpend _Y1 since each covariate is measured in both years
  paste(paste0(baseline_covariates, "_Y1"), collapse=' + '),
  "+",
  paste(paste0('PAIR', 2:24), collapse=" + ")
))

g1 <- lm(formula, truancy)

summary(g1)
```

```
##
## Call:
## lm(formula = formula, data = truancy)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.011238 -0.002671  0.000000  0.002671  0.011238
##
## Coefficients: (1 not defined because of singularities)
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.1404010  0.0822409   1.707  0.11350
## TREATMENT       -0.0045298  0.0038365  -1.181  0.26059
## TRUANCY_Y0       0.5720078  0.1618754   3.534  0.00412 **
## FRL_Y1          -0.0313506  0.0431729  -0.726  0.48166
## PUPTCH_Y1        0.0008889  0.0012056   0.737  0.47509
## PCT_BL_Y1       -0.0002556  0.0008509  -0.300  0.76902
## PCT_HS_Y1       -0.0007178  0.0006035  -1.189  0.25725
## PCT_WH_Y1       -0.0015537  0.0008685  -1.789  0.09889 .
## SMALLSCHOOL_Y1  -0.0066112  0.0087304  -0.757  0.46350
## LARGESCHOOL_Y1  -0.0028575  0.0095453  -0.299  0.76979
## GRDRNG_ELEM_Y1  -0.0188637  0.0095626  -1.973  0.07202 .
## GRDRNG_HS_Y1    -0.0009131  0.0105886  -0.086  0.93270
## PAIR2           -0.0180992  0.0207971  -0.870  0.40122
## PAIR3            0.0003043  0.0118341   0.026  0.97991
## PAIR4                   NA         NA      NA       NA
## PAIR5           -0.0057333  0.0116967  -0.490  0.63286
## PAIR6           -0.0257217  0.0206272  -1.247  0.23620
## PAIR7           -0.0106844  0.0151410  -0.706  0.49387
## PAIR8           -0.0219274  0.0204750  -1.071  0.30527
## PAIR9           -0.0328363  0.0193694  -1.695  0.11579
## PAIR10          -0.0114448  0.0141755  -0.807  0.43517
## PAIR11          -0.0175014  0.0160562  -1.090  0.29711
## PAIR12          -0.0387673  0.0226065  -1.715  0.11205
## PAIR13          -0.0201080  0.0208462  -0.965  0.35379
## PAIR14          -0.0109927  0.0204759  -0.537  0.60118
## PAIR15          -0.0327989  0.0204853  -1.601  0.13534
## PAIR16          -0.0235714  0.0164502  -1.433  0.17742
## PAIR17          -0.0311791  0.0202744  -1.538  0.15002
## PAIR18          -0.0319782  0.0209843  -1.524  0.15344
## PAIR19          -0.0027070  0.0110190  -0.246  0.81009
```

```
## PAIR20        -0.0041192  0.0100517  -0.410   0.68918
## PAIR21        -0.0251130  0.0186105  -1.349   0.20212
## PAIR22        -0.0086354  0.0131927  -0.655   0.52510
## PAIR23         0.0041032  0.0186213   0.220   0.82930
## PAIR24        -0.0368645  0.0208667  -1.767   0.10268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.009657 on 12 degrees of freedom
## Multiple R-squared:  0.9531, Adjusted R-squared:  0.824
## F-statistic: 7.383 on 33 and 12 DF,  p-value: 0.0003625
```

```r
truancy <- load_school_level_outcome("truancy")

formula <- as.formula(paste(
  "TRUANCY_Y2 ~",
  "TREATMENT + TRUANCY_Y0 + ",
  # postpend _Y1 since each covariate is measured in both years
  paste(paste0(baseline_covariates, "_Y1"), collapse=' + '),
  "+",
  paste(paste0('PAIR', 2:24), collapse=" + ")
))

g2 <- lm(formula, truancy)

summary(g2)
```

```
##
## Call:
## lm(formula = formula, data = truancy)
##
## Residuals:
##       Min       1Q   Median       3Q      Max
## -0.012828 -0.003225  0.000000  0.003225  0.012828
##
## Coefficients: (2 not defined because of singularities)
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0365698  0.0946945   0.386  0.70672
## TREATMENT     -0.0047347  0.0043971  -1.077  0.30463
## TRUANCY_Y0     0.7082963  0.2003988   3.534  0.00468 **
## FRL_Y1         0.0498007  0.0523532   0.951  0.36191
## PUPTCH_Y1      0.0002333  0.0013466   0.173  0.86561
## PCT_BL_Y1     -0.0002244  0.0009471  -0.237  0.81708
## PCT_HS_Y1     -0.0002906  0.0006763  -0.430  0.67573
## PCT_WH_Y1     -0.0005234  0.0009844  -0.532  0.60552
## SMALLSCHOOL_Y1 -0.0079299  0.0097145  -0.816  0.43165
## LARGESCHOOL_Y1  0.0106822  0.0106194   1.006  0.33607
## GRDRNG_ELEM_Y1 -0.0005133  0.0110628  -0.046  0.96382
## GRDRNG_HS_Y1    0.0047042  0.0118068   0.398  0.69794
## PAIR2         -0.0162002  0.0231360  -0.700  0.49834
## PAIR3         -0.0066559  0.0133655  -0.498  0.62829
## PAIR4                 NA         NA      NA       NA
## PAIR5         -0.0086587  0.0131719  -0.657  0.52446
## PAIR6         -0.0436540  0.0229387  -1.903  0.08351 .
```

```
## PAIR7           -0.0073188  0.0169893  -0.431  0.67494
## PAIR8           -0.0277307  0.0229978  -1.206  0.25319
## PAIR9           -0.0507350  0.0215467  -2.355  0.03817 *
## PAIR10          -0.0149727  0.0157837  -0.949  0.36319
## PAIR11          -0.0339501  0.0179439  -1.892  0.08509 .
## PAIR12          -0.0497827  0.0251782  -1.977  0.07362 .
## PAIR13          -0.0290535  0.0232515  -1.250  0.23740
## PAIR14                  NA         NA      NA       NA
## PAIR15          -0.0243880  0.0228056  -1.069  0.30780
## PAIR16           0.0008725  0.0186639   0.047  0.96355
## PAIR17          -0.0459519  0.0225475  -2.038  0.06633 .
## PAIR18          -0.0284768  0.0234267  -1.216  0.24959
## PAIR19           0.0072871  0.0123540   0.590  0.56720
## PAIR20          -0.0046088  0.0111867  -0.412  0.68826
## PAIR21          -0.0233955  0.0206974  -1.130  0.28237
## PAIR22           0.0108623  0.0150184   0.723  0.48461
## PAIR23          -0.0160208  0.0207258  -0.773  0.45582
## PAIR24          -0.0077895  0.0233518  -0.334  0.74497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01074 on 11 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.9391, Adjusted R-squared:  0.7618
## F-statistic: 5.297 on 32 and 11 DF,  p-value: 0.002783
```

Curiosity: Correlation between Truancy and Attendance

```r
par(mfrow=c(1, 2))
xlim <- c(0.8, 1.0)
ylim <- c(0.0, 0.1)

plot(
  school_level_outcomes$ATTEND_Y1,
  school_level_outcomes$TRUANCY_Y1,
  xlab='Attendance, Year 1',
  ylab='Truancy, Year 1',
  xlim=xlim, ylim=ylim
  )
plot(
  school_level_outcomes$ATTEND_Y2,
  school_level_outcomes$TRUANCY_Y2,
  xlab='Attendance, Year 2',
  ylab='Truancy, Year 2',
  xlim=xlim, ylim=ylim
  )

mtext(
  "Correlation between Attendance and Truancy Outcomes",
      side = 3,
      line = - 2,
      outer = TRUE
```

```
        )
```

Correlation between Attendance and Truancy Outcomes
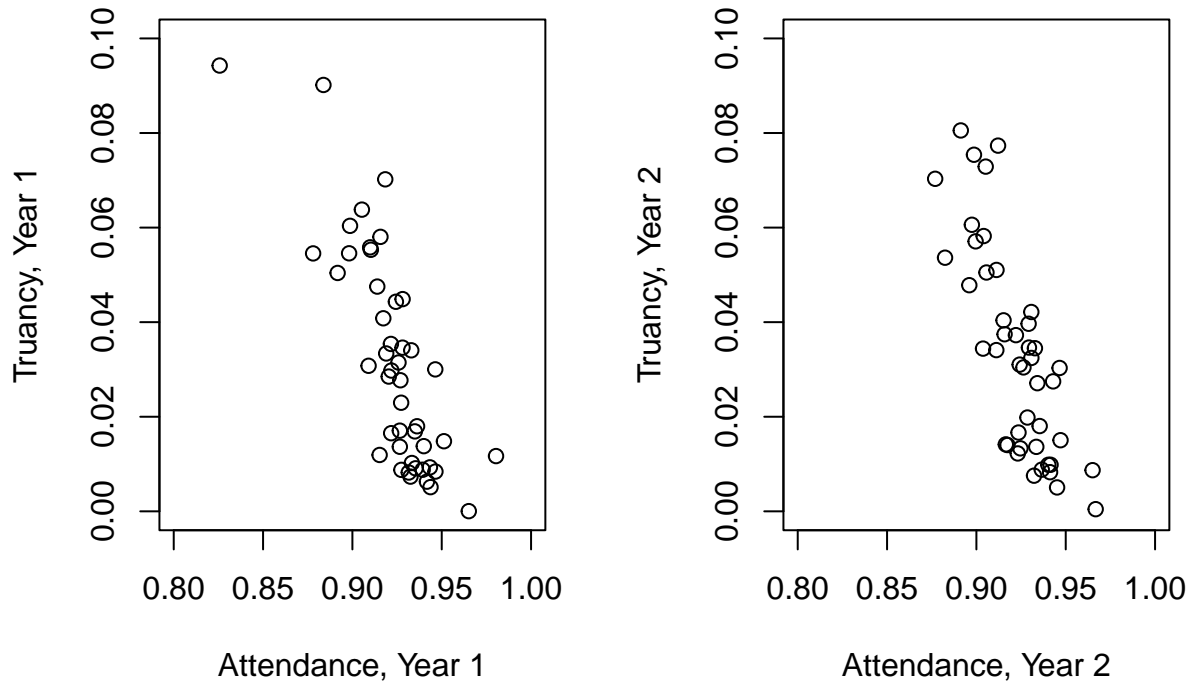
Individual-Level: Achievement

```
achievement <- load_achievement()
```

The researchers employed a matched-pairs experimental design with a waitlist control structure, divided into two cohorts. Cohort 1 data was measured during the 2016-17 and 2017-18 school years; Cohort 2 started one year later, with outcome data measured therefore during the 2017-18 and 2018-19 school years.

Observations were deleted "listwise" when outcome data were missing. No analysis of this attrition was performed in the paper.

```
# col_mapping <- c(
#   Z="TREATMENT",
#   GRADE="GRADE",
#   SCHOOL_YEAR="SCHOOL_YEAR",
#   paste0(
#     "Y",
#     rep(1:2, each=2),
#     c("BASELINE_R_Z", "BASELINE_M_Z")
#   )
# )
```

```r
knitr::kable(
  achievement %>% sample_n(5),
  digits=3,
  caption="School-Level Outcome Information"
  )
```

Table 3: School-Level Outcome Information

GRADE SCHOOL_YEAR FEMALE WHITE ZSCORE_R06 ZSCORE_R07 ZSCORE_R08 ...
07 2018 0 - - - - - - 1 2 pair13 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.948 1.699 2.872 4- 0 0 1.355 0.500 0.050 0.537 1.353 12
08 2018 1 0.782 0.476 0.870 NA 1 2 pair10 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.383 4.719 6.585 2.663 0 1
08 2019 1 NA 0.951 NA 0.069 2 2 pair20 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0.326 2.329 5.697 NA 0 1
08 2017 0 NA NA NA NA 1 1 pair05 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.496 8.523 7.086 NA 0 0.842 0.114
08 2017 0 NA NA NA NA 0 1 pair11 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.356 5.963 7.515 NA 0 1.868 2.057

```r
achievement %>% group_by(PID)
```

```
## # A tibble: 12,282 x 49
## # Groups:   PID [17]
##    GRADE SCHOOL_YEAR FEMALE WHITE ZSCORE_R06 ZSCORE_R07 ZSCORE_R08 ZSCORE_M06
##    <fct> <fct>        <dbl> <dbl>      <dbl>      <dbl>      <dbl>      <dbl>
##  1 07    2017             1     1      0.406     -0.237       1.15      0.613
##  2 07    2017             0     0     -0.156     -0.289      -0.35      0.226
##  3 08    2017             0     1     NA         NA           0.316    NA
##  4 07    2017             1     0     -0.312      0.132       0.225     0.742
##  5 08    2017             0     0     NA         NA          -1        NA
##  6 08    2017             0     0     NA         NA           0.237    NA
##  7 08    2017             1     0     NA         -0.865       0         NA
##  8 07    2017             1     1      0.969      1.37        0.925     0.839
##  9 07    2017             1     0     -1.19      -1.39       -0.575    -1.42
## 10 07    2017             1     1      0.531      0.263      -0.075     0.419
## # i 12,272 more rows
## # i 41 more variables: ZSCORE_M07 <dbl>, ZSCORE_M08 <dbl>, TREATMENT <dbl>,
## #   COHORT <dbl>, PID <fct>, PAIR1 <dbl>, PAIR2 <dbl>, PAIR3 <dbl>,
## #   PAIR4 <dbl>, PAIR5 <dbl>, PAIR6 <dbl>, PAIR7 <dbl>, PAIR8 <dbl>,
## #   PAIR9 <dbl>, PAIR10 <dbl>, PAIR11 <dbl>, PAIR12 <dbl>, PAIR13 <dbl>,
## #   PAIR14 <dbl>, PAIR15 <dbl>, PAIR16 <dbl>, PAIR17 <dbl>, PAIR18 <dbl>,
## #   PAIR19 <dbl>, PAIR20 <dbl>, PAIR21 <dbl>, PAIR22 <dbl>, PAIR23 <dbl>, ...
```

The researchers estimated two separate sets of models, one for reading scores and one for math scores.

Year 1 Models

```r
library(lme4)
```

```
## Loading required package: Matrix
```

```r
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':
##
##     lmer

## The following object is masked from 'package:stats':
##
##     step
```

```r
# In this regression, the authors do NOT consider grade range of the school.
# It is unclear why not.
indreg_school_covariates <- c(
  "FRL", "PUPTCH", "PCT_BL", "PCT_HS", "PCT_WH",
  "SMALLSCHOOL", "LARGESCHOOL"
)

ind_covariates <- c("FEMALE", "WHITE")

year_1_grade_7_df <- achievement %>%
  filter(!is.na(Y1BASELINE_R_Z), !is.na(Y1BASELINE_M_Z), GRADE=="07") %>%
  mutate(
    Y_R=ZSCORE_R07,
    Y_M=ZSCORE_M07
    )

formula <- as.formula(paste(
  "Y_R ~",
  "TREATMENT + Y1BASELINE_R_Z + ",
  paste(ind_covariates, collapse=" + "),
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g1 <- lmer(formula, year_1_grade_7_df)
```

```
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
```

```
## boundary (singular) fit: see help('isSingular')
```

```
summary(g1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_1_grade_7_df
##
## REML criterion at convergence: 5845.9
##
## Scaled residuals:
##     Min     1Q  Median     3Q     Max
## -4.4486 -0.6211  0.0428  0.6837  3.7554
##
## Random effects:
##  Groups      Name        Variance  Std.Dev.
##  PCT_WH      (Intercept) 1.301e-10 0.0000114
##  PUPTCH      (Intercept) 0.000e+00 0.0000000
##  FRL         (Intercept) 0.000e+00 0.0000000
##  PCT_HS      (Intercept) 0.000e+00 0.0000000
##  PCT_BL      (Intercept) 4.349e-02 0.2085399
##  LARGESCHOOL (Intercept) 0.000e+00 0.0000000
##  SMALLSCHOOL (Intercept) 2.166e-02 0.1471896
##  Residual                2.676e-01 0.5173107
## Number of obs: 3766, groups:
## PCT_WH, 34; PUPTCH, 34; FRL, 34; PCT_HS, 33; PCT_BL, 26; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                 Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)    1.222e-02  1.895e-01 5.718e+00   0.065  0.95075
## TREATMENT      1.764e-01  5.599e-02 5.065e+01   3.150  0.00274 **
## Y1BASELINE_R_Z 8.120e-01  9.638e-03 3.739e+03  84.255  < 2e-16 ***
## FEMALE         1.353e-01  1.749e-02 3.733e+03   7.736 1.31e-14 ***
## WHITE          2.248e-02  2.086e-02 3.739e+03   1.078  0.28124
## PAIR2         -1.606e-01  2.188e-01 1.233e+01  -0.734  0.47668
## PAIR3         -1.001e-01  1.982e-01 1.739e+01  -0.505  0.61972
## PAIR5         -5.243e-02  1.882e-01 1.524e+01  -0.279  0.78430
## PAIR6         -5.095e-02  2.135e-01 1.185e+01  -0.239  0.81549
## PAIR7         -2.653e-01  2.193e-01 1.242e+01  -1.210  0.24890
## PAIR8         -4.545e-02  1.995e-01 1.698e+01  -0.228  0.82252
## PAIR10        -1.260e-01  2.124e-01 1.160e+01  -0.593  0.56448
## PAIR11        -1.517e-01  2.131e-01 1.176e+01  -0.712  0.49043
## PAIR13        -2.253e-01  1.893e-01 1.562e+01  -1.190  0.25178
## PAIR15        -1.534e-01  2.242e-01 1.346e+01  -0.684  0.50554
## PAIR19        -1.588e-01  1.800e-01 1.445e+01  -0.883  0.39192
## PAIR20         1.500e-01  1.827e-01 1.463e+01   0.821  0.42498
## PAIR21        -7.282e-02  2.137e-01 1.187e+01  -0.341  0.73919
## PAIR22         1.462e-01  1.790e-01 1.374e+01   0.817  0.42808
## PAIR23        -1.003e-01  2.219e-01 1.304e+01  -0.452  0.65851
## PAIR24        -5.511e-01  2.014e-01 2.031e+01  -2.736  0.01261 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
```

```
## Correlation matrix not shown by default, as p = 21 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
formula <- as.formula(paste(
  "Y_M ~",
  "TREATMENT + Y1BASELINE_M_Z + ",
  paste(ind_covariates, collapse=" + "),
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g1_m_7 <- lmer(formula, year_1_grade_7_df)
```

```
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
summary(g1_m_7)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_1_grade_7_df
##
## REML criterion at convergence: 5201.4
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -5.3167 -0.5777  0.0406  0.6388  4.3191
##
## Random effects:
##  Groups      Name        Variance  Std.Dev.
##  PCT_WH      (Intercept) 6.199e-03 0.078732
##  PUPTCH      (Intercept) 1.079e-05 0.003285
##  FRL         (Intercept) 4.958e-05 0.007041
##  PCT_HS      (Intercept) 0.000e+00 0.000000
##  PCT_BL      (Intercept) 0.000e+00 0.000000
##  LARGESCHOOL (Intercept) 0.000e+00 0.000000
##  SMALLSCHOOL (Intercept) 1.746e-04 0.013213
##  Residual                2.301e-01 0.479719
## Number of obs: 3723, groups:
## PCT_WH, 34; PUPTCH, 34; FRL, 34; PCT_HS, 33; PCT_BL, 26; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
```

```
##                   Estimate Std. Error          df t value Pr(>|t|)
## (Intercept)     -4.167e-02  6.822e-02  9.003e+00  -0.611 0.556398
## TREATMENT        3.192e-02  3.432e-02  1.189e+01   0.930 0.370875
## Y1BASELINE_M_Z   8.516e-01  9.223e-03  3.700e+03  92.331  < 2e-16 ***
## FEMALE          -2.487e-03  1.581e-02  3.695e+03  -0.157 0.875027
## WHITE            6.425e-02  1.940e-02  3.698e+03   3.313 0.000933 ***
## PAIR2           -1.281e-02  9.133e-02  8.885e+00  -0.140 0.891567
## PAIR3            5.528e-02  9.977e-02  1.111e+01   0.554 0.590537
## PAIR5            1.456e-01  9.260e-02  9.296e+00   1.572 0.149230
## PAIR6            6.536e-02  8.994e-02  7.600e+00   0.727 0.489181
## PAIR7            1.086e-01  9.195e-02  8.606e+00   1.181 0.269097
## PAIR8            3.617e-02  1.011e-01  1.226e+01   0.358 0.726555
## PAIR10          -4.812e-02  8.750e-02  6.832e+00  -0.550 0.599878
## PAIR11           1.112e-01  8.907e-02  7.325e+00   1.249 0.250110
## PAIR13          -5.910e-02  9.565e-02  1.080e+01  -0.618 0.549431
## PAIR15          -5.813e-02  1.006e-01  1.142e+01  -0.578 0.574467
## PAIR19           1.630e-01  9.301e-02  8.723e+00   1.753 0.114561
## PAIR20           8.200e-02  9.502e-02  9.381e+00   0.863 0.409668
## PAIR21          -1.606e-01  8.980e-02  7.462e+00  -1.789 0.114191
## PAIR22           1.147e-01  9.533e-02  9.763e+00   1.203 0.257287
## PAIR23           9.557e-02  9.688e-02  1.118e+01   0.987 0.344753
## PAIR24          -2.553e-01  1.178e-01  2.015e+01  -2.167 0.042405 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##
## Correlation matrix not shown by default, as p = 21 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)         if you need it

## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
year_1_grade_8_df <- achievement %>%
  filter(!is.na(Y1BASELINE_R_Z), !is.na(Y1BASELINE_M_Z), GRADE=="08") %>%
  mutate(
    Y_R=ZSCORE_R08,
    Y_M=ZSCORE_M08
    )

formula <- as.formula(paste(
  "Y_R ~",
  "TREATMENT + Y1BASELINE_R_Z + ",
  paste(ind_covariates, collapse=" + "),
  " + ",
  # postpend _Y1 since each covariate is measured in both years
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g1_r_8 <- lmer(formula, year_1_grade_8_df)
```

15

```
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients

## boundary (singular) fit: see help('isSingular')
```

```
summary(g1_r_8)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_1_grade_8_df
##
## REML criterion at convergence: 5648.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.6738 -0.6059  0.0323  0.6530  3.6093
##
## Random effects:
##  Groups      Name        Variance  Std.Dev.
##  PCT_WH      (Intercept) 0.000e+00 0.000e+00
##  PUPTCH      (Intercept) 0.000e+00 0.000e+00
##  FRL         (Intercept) 1.255e-07 3.543e-04
##  PCT_HS      (Intercept) 3.216e-02 1.793e-01
##  PCT_BL      (Intercept) 8.370e-10 2.893e-05
##  LARGESCHOOL (Intercept) 7.769e-08 2.787e-04
##  SMALLSCHOOL (Intercept) 4.360e-02 2.088e-01
##  Residual                2.840e-01 5.329e-01
## Number of obs: 3503, groups:
## PCT_WH, 34; PUPTCH, 34; FRL, 34; PCT_HS, 33; PCT_BL, 26; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)   -2.904e-02  2.054e-01  2.844e+00  -0.141   0.8970
## TREATMENT      1.418e-01  7.232e-02  1.427e+01   1.961   0.0697 .
## Y1BASELINE_R_Z 7.980e-01  1.070e-02  3.481e+03  74.596   <2e-16 ***
## FEMALE         1.825e-01  1.885e-02  3.472e+03   9.681   <2e-16 ***
## WHITE          1.503e-02  2.223e-02  3.478e+03   0.676   0.4990
## PAIR2         -5.257e-02  1.947e-01  1.301e+01  -0.270   0.7914
## PAIR3         -3.127e-01  1.972e-01  1.696e+01  -1.586   0.1313
## PAIR5         -3.940e-02  1.945e-01  1.301e+01  -0.203   0.8426
## PAIR6          2.963e-02  1.866e-01  1.232e+01   0.159   0.8764
## PAIR7         -3.730e-02  1.938e-01  1.277e+01  -0.193   0.8504
## PAIR8          9.920e-02  1.982e-01  1.402e+01   0.501   0.6244
## PAIR10         1.629e-03  1.857e-01  1.211e+01   0.009   0.9931
## PAIR11        -4.935e-02  1.866e-01  1.234e+01  -0.264   0.7958
## PAIR13        -1.114e-01  1.961e-01  1.345e+01  -0.568   0.5794
## PAIR15        -2.880e-01  2.017e-01  1.456e+01  -1.428   0.1745
## PAIR19        -2.656e-01  1.900e-01  1.327e+01  -1.398   0.1851
## PAIR20         1.887e-01  1.898e-01  1.319e+01   0.994   0.3381
## PAIR21         5.736e-02  1.874e-01  1.253e+01   0.306   0.7646
## PAIR22        -2.257e-01  1.872e-01  1.507e+01  -1.206   0.2465
## PAIR23        -3.600e-01  1.962e-01  1.348e+01  -1.835   0.0887 .
## PAIR24        -2.454e-01  2.267e-01  1.673e+01  -1.083   0.2943
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation matrix not shown by default, as p = 21 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it
```

```
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```r
formula <- as.formula(paste(
  "Y_M ~",
  "TREATMENT + Y1BASELINE_M_Z + ",
  # postpend _Y1 since each covariate is measured in both years
  paste(ind_covariates, collapse=" + "),
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g1_m_8 <- lmer(formula, year_1_grade_8_df)
```

```
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
```

```
## boundary (singular) fit: see help('isSingular')
```

```
## Warning: Model failed to converge with 4 negative eigenvalues: -2.5e-02
## -4.0e-02 -5.8e+01 -4.4e+02
```

```r
summary(g1_m_8)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_1_grade_8_df
##
## REML criterion at convergence: 5191.3
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9515 -0.6276  0.0491  0.6737  3.2064
##
## Random effects:
##  Groups       Name        Variance  Std.Dev.
##  PCT_WH       (Intercept) 2.822e-06 1.680e-03
##  PUPTCH       (Intercept) 1.432e-02 1.197e-01
##  FRL          (Intercept) 3.213e-05 5.668e-03
```

```
## PCT_HS      (Intercept) 6.796e-10 2.607e-05
## PCT_BL      (Intercept) 0.000e+00 0.000e+00
## LARGESCHOOL (Intercept) 1.157e-06 1.076e-03
## SMALLSCHOOL (Intercept) 5.944e-10 2.438e-05
## Residual                3.260e-01 5.709e-01
## Number of obs: 2968, groups:
## PCT_WH, 34; PUPTCH, 34; FRL, 34; PCT_HS, 33; PCT_BL, 26; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                  Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)     2.208e-01  1.020e-01 1.507e+01   2.164  0.04691 *
## TREATMENT      -7.094e-03  4.992e-02 1.673e+01  -0.142  0.88869
## Y1BASELINE_M_Z  7.957e-01  1.256e-02 2.941e+03  63.380  < 2e-16 ***
## FEMALE          6.775e-02  2.113e-02 2.942e+03   3.207  0.00136 **
## WHITE           3.242e-02  2.580e-02 2.946e+03   1.257  0.20893
## PAIR2          -2.987e-02  1.373e-01 1.403e+01  -0.217  0.83097
## PAIR3          -4.259e-01  1.541e-01 1.843e+01  -2.763  0.01261 *
## PAIR5          -1.758e-01  1.401e-01 1.525e+01  -1.255  0.22846
## PAIR6          -5.267e-02  1.347e-01 1.331e+01  -0.391  0.70198
## PAIR7          -1.415e-01  1.360e-01 1.351e+01  -1.040  0.31658
## PAIR8          -1.161e-01  1.436e-01 1.697e+01  -0.808  0.43017
## PAIR10         -1.903e-01  1.333e-01 1.256e+01  -1.427  0.17794
## PAIR11         -3.884e-01  1.351e-01 1.355e+01  -2.874  0.01258 *
## PAIR13         -1.660e-01  1.397e-01 1.525e+01  -1.189  0.25263
## PAIR15         -3.508e-01  1.479e-01 1.690e+01  -2.372  0.02984 *
## PAIR19         -1.476e-01  1.435e-01 1.694e+01  -1.029  0.31798
## PAIR20          2.293e-02  1.446e-01 1.722e+01   0.159  0.87588
## PAIR21         -6.369e-02  1.348e-01 1.327e+01  -0.472  0.64428
## PAIR22         -6.523e-02  1.394e-01 1.505e+01  -0.468  0.64662
## PAIR23         -2.373e-01  1.403e-01 1.555e+01  -1.691  0.11083
## PAIR24         -2.152e-01  1.588e-01 2.057e+01  -1.356  0.18992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##
## Correlation matrix not shown by default, as p = 21 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)         if you need it


## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

Year 2 Models

```
schoolsize <- achievement %>%
  filter(!is.na(LARGESCHOOL)) %>%
  group_by(PID, TREATMENT) %>%
  select(PID, TREATMENT, LARGESCHOOL) %>%
  distinct()
```

```r
year_2_df <- achievement %>%
  filter(!is.na(Y2BASELINE_R_Z), !is.na(Y2BASELINE_M_Z)) %>%
  # only 8th graders in the sample
  mutate(
    Y_R=ZSCORE_R08,
    Y_M=ZSCORE_M08
    ) %>%
  left_join(
    schoolsize, by=c("PID", "TREATMENT")
  ) %>%
  select(-LARGESCHOOL.x) %>%   # remove the old NA column
  rename(LARGESCHOOL = LARGESCHOOL.y)   # rename the new column

formula <- as.formula(paste(
  "Y_R ~",
  "TREATMENT + Y2BASELINE_R_Z + ",
  paste(ind_covariates, collapse=" + "),
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g2_r_8 <- lmer(formula, year_2_df)
```

```
## fixed-effect model matrix is rank deficient so dropping 13 columns / coefficients
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00601348 (tol = 0.002, component 1)
```

```r
summary(g2_r_8)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_2_df
##
## REML criterion at convergence: 4573.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.2658 -0.6387  0.0608  0.6794  3.3959
##
## Random effects:
##  Groups       Name        Variance  Std.Dev.
##  PCT_WH       (Intercept) 8.526e-03 0.0923363
##  PCT_HS       (Intercept) 7.025e-03 0.0838176
##  PUPTCH       (Intercept) 4.856e-03 0.0696878
##  FRL          (Intercept) 7.198e-05 0.0084844
##  PCT_BL       (Intercept) 2.631e-08 0.0001622
##  LARGESCHOOL  (Intercept) 1.342e-08 0.0001159
##  SMALLSCHOOL  (Intercept) 4.553e-02 0.2133662
```

```
## Residual                     3.329e-01 0.5770075
## Number of obs: 2586, groups:
## PCT_WH, 22; PCT_HS, 22; PUPTCH, 22; FRL, 22; PCT_BL, 21; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                   Estimate Std. Error         df t value Pr(>|t|)
## (Intercept)     -6.396e-02  1.954e-01  2.008e+00  -0.327   0.7744
## TREATMENT        9.216e-02  7.033e-02  8.827e+00   1.310   0.2231
## Y2BASELINE_R_Z   7.826e-01  1.307e-02  2.569e+03  59.868   <2e-16 ***
## FEMALE           1.952e-01  2.348e-02  2.564e+03   8.316   <2e-16 ***
## WHITE            5.237e-02  2.773e-02  2.570e+03   1.889   0.0591 .
## PAIR5            9.869e-02  1.637e-01  8.186e+00   0.603   0.5630
## PAIR6           -1.935e-02  1.535e-01  7.742e+00  -0.126   0.9029
## PAIR7           -1.006e-01  1.629e-01  7.901e+00  -0.618   0.5543
## PAIR8           -2.828e-01  1.978e-01  1.055e+01  -1.429   0.1818
## PAIR10          -2.361e-03  1.508e-01  7.225e+00  -0.016   0.9879
## PAIR11          -6.383e-03  1.529e-01  7.619e+00  -0.042   0.9678
## PAIR20           3.785e-01  1.556e-01  8.186e+00   2.433   0.0404 *
## PAIR21           1.762e-01  1.536e-01  7.729e+00   1.147   0.2855
## PAIR22          -6.561e-02  1.652e-01  8.400e+00  -0.397   0.7012
## PAIR23          -1.854e-01  1.675e-01  8.982e+00  -1.107   0.2969
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##
## Correlation matrix not shown by default, as p = 15 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it


## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 13 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.00601348 (tol = 0.002, component 1)
```

```r
formula <- as.formula(paste(
  "Y_M ~",
  "TREATMENT + Y2BASELINE_M_Z + ",
  paste(ind_covariates, collapse=" + "),
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g2_m_8 <- lmer(formula, year_2_df)
```

```
## fixed-effect model matrix is rank deficient so dropping 13 columns / coefficients
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
summary(g2_m_8)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_2_df
##
## REML criterion at convergence: 3887.2
##
## Scaled residuals:
##     Min     1Q  Median     3Q     Max
## -4.3467 -0.6114  0.0633  0.6628  3.7693
##
## Random effects:
##  Groups      Name        Variance  Std.Dev.
##  PCT_WH      (Intercept) 0.000e+00 0.0000000
##  PCT_HS      (Intercept) 0.000e+00 0.0000000
##  PUPTCH      (Intercept) 3.277e-07 0.0005724
##  FRL         (Intercept) 5.499e-08 0.0002345
##  PCT_BL      (Intercept) 1.557e-02 0.1247783
##  LARGESCHOOL (Intercept) 1.702e-08 0.0001305
##  SMALLSCHOOL (Intercept) 3.448e-04 0.0185676
##  Residual                2.760e-01 0.5253332
## Number of obs: 2457, groups:
## PCT_WH, 22; PCT_HS, 22; PUPTCH, 22; FRL, 22; PCT_BL, 21; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)   -2.611e-03  1.046e-01 4.499e+00  -0.025  0.98114
## TREATMENT      1.292e-03  5.727e-02 1.167e+01   0.023  0.98238
## Y2BASELINE_M_Z 8.463e-01  1.274e-02 2.439e+03  66.412  < 2e-16 ***
## FEMALE         8.902e-02  2.127e-02 2.437e+03   4.185 2.95e-05 ***
## WHITE          6.724e-02  2.564e-02 2.441e+03   2.622  0.00879 **
## PAIR5          4.841e-02  1.353e-01 1.128e+01   0.358  0.72715
## PAIR6          8.858e-02  1.371e-01 9.304e+00   0.646  0.53371
## PAIR7         -4.169e-02  1.389e-01 1.025e+01  -0.300  0.77010
## PAIR8         -3.186e-01  1.440e-01 6.064e+00  -2.213  0.06844 .
## PAIR10        -1.259e-01  1.340e-01 8.536e+00  -0.939  0.37338
## PAIR11        -5.447e-02  1.358e-01 8.999e+00  -0.401  0.69777
## PAIR20         2.477e-01  1.384e-01 9.701e+00   1.790  0.10470
## PAIR21        -1.285e-01  1.365e-01 9.125e+00  -0.941  0.37080
## PAIR22        -1.082e-01  1.421e-01 1.113e+01  -0.761  0.46222
## PAIR23        -1.041e-01  1.426e-01 1.156e+01  -0.730  0.47978
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##
## Correlation matrix not shown by default, as p = 15 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it


## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 13 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

Robustness to Model Specification

In this section, we probe the analysis's robustness to model specification by grouping the grade analyses. We also hope to find alternative ways of analyzing the data to improve parameter standard errors.

We'll try to estimate 1-year grade-specific effects together in one model, rather than two separate models.

```r
year_1_df <- achievement %>%
  filter(!is.na(Y1BASELINE_R_Z), !is.na(Y1BASELINE_M_Z)) %>%
  mutate(
    Y_R=if_else(
      GRADE=="07",
      ZSCORE_R07,
      ZSCORE_R08
    ),
    Y_M=if_else(
      GRADE=="07",
      ZSCORE_M07,
      ZSCORE_M08
    ),
    GRADE_8 = (GRADE == "08")
  )

formula <- as.formula(paste(
  "Y_R ~",
  "TREATMENT:GRADE_8 + Y1BASELINE_R_Z + ",
  paste(ind_covariates, collapse=" + "),
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g1 <- lmer(formula, year_1_df)
```

```
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
```

```
## boundary (singular) fit: see help('isSingular')
```

```r
summary(g1)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_1_df
##
## REML criterion at convergence: 11647
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -4.5492 -0.6253  0.0374  0.6635  3.9265
##
## Random effects:
```

```
##  Groups       Name        Variance  Std.Dev.
##  PCT_WH       (Intercept) 0.000e+00 0.000e+00
##  PUPTCH       (Intercept) 0.000e+00 0.000e+00
##  FRL          (Intercept) 5.112e-11 7.150e-06
##  PCT_HS       (Intercept) 0.000e+00 0.000e+00
##  PCT_BL       (Intercept) 3.340e-02 1.828e-01
##  LARGESCHOOL (Intercept) 4.689e-11 6.848e-06
##  SMALLSCHOOL (Intercept) 3.461e-02 1.860e-01
##  Residual                 2.849e-01 5.337e-01
## Number of obs: 7269, groups:
## PCT_WH, 34; PUPTCH, 34; FRL, 34; PCT_HS, 33; PCT_BL, 26; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                        Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)          -9.793e-03  1.907e-01 3.341e+00  -0.051  0.96197
## Y1BASELINE_R_Z        8.030e-01  7.216e-03 7.246e+03 111.281  < 2e-16 ***
## FEMALE                1.587e-01  1.302e-02 7.237e+03  12.186  < 2e-16 ***
## WHITE                 2.180e-02  1.545e-02 7.244e+03   1.411  0.15824
## PAIR2                -1.285e-01  1.910e-01 1.267e+01  -0.672  0.51341
## PAIR3                -1.741e-01  1.717e-01 1.742e+01  -1.014  0.32430
## PAIR5                 1.751e-02  1.621e-01 1.489e+01   0.108  0.91546
## PAIR6                -1.635e-02  1.860e-01 1.226e+01  -0.088  0.93139
## PAIR7                -1.522e-01  1.911e-01 1.267e+01  -0.796  0.44045
## PAIR8                -7.185e-02  1.699e-01 1.623e+01  -0.423  0.67795
## PAIR10               -6.966e-02  1.854e-01 1.211e+01  -0.376  0.71368
## PAIR11               -1.062e-01  1.859e-01 1.223e+01  -0.571  0.57829
## PAIR13               -2.001e-01  1.630e-01 1.519e+01  -1.228  0.23826
## PAIR15               -2.103e-01  1.948e-01 1.358e+01  -1.080  0.29901
## PAIR19               -1.854e-01  1.555e-01 1.432e+01  -1.192  0.25275
## PAIR20                2.424e-01  1.573e-01 1.453e+01   1.541  0.14482
## PAIR21               -1.678e-02  1.863e-01 1.233e+01  -0.090  0.92969
## PAIR22               -3.079e-02  1.548e-01 1.376e+01  -0.199  0.84519
## PAIR23               -2.526e-01  1.924e-01 1.304e+01  -1.313  0.21191
## PAIR24               -4.633e-01  1.667e-01 1.694e+01  -2.779  0.01289 *
## TREATMENT:GRADE_8FALSE  2.145e-01  4.646e-02 6.317e+01   4.617 1.96e-05 ***
## TREATMENT:GRADE_8TRUE   1.501e-01  4.660e-02 6.366e+01   3.221  0.00201 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##
## Correlation matrix not shown by default, as p = 22 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it


## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```

```
formula <- as.formula(paste(
  "Y_M ~",
  "TREATMENT:GRADE_8 + Y1BASELINE_M_Z + ",
  paste(ind_covariates, collapse=" + "),
```

```
  " + ",
  paste(paste0('PAIR', 2:24), collapse=" + "),
  "+",
  paste(paste0("(1|", indreg_school_covariates, ")"), collapse=' + ')
))

g1_m <- lmer(formula, year_1_df)
```

fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients

boundary (singular) fit: see help('isSingular')

```
summary(g1_m)
```

```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: formula
##    Data: year_1_df
##
## REML criterion at convergence: 10680.5
##
## Scaled residuals:
##     Min     1Q  Median     3Q     Max
## -4.9028 -0.5931  0.0477  0.6451  3.5155
##
## Random effects:
##  Groups      Name        Variance  Std.Dev.
##  PCT_WH      (Intercept) 0.000e+00 0.000e+00
##  PUPTCH      (Intercept) 5.614e-09 7.493e-05
##  FRL         (Intercept) 1.076e-09 3.280e-05
##  PCT_HS      (Intercept) 2.595e-03 5.094e-02
##  PCT_BL      (Intercept) 0.000e+00 0.000e+00
##  LARGESCHOOL (Intercept) 0.000e+00 0.000e+00
##  SMALLSCHOOL (Intercept) 0.000e+00 0.000e+00
##  Residual                2.840e-01 5.329e-01
## Number of obs: 6691, groups:
## PCT_WH, 34; PUPTCH, 34; FRL, 34; PCT_HS, 33; PCT_BL, 26; LARGESCHOOL, 2; SMALLSCHOOL, 2
##
## Fixed effects:
##                   Estimate Std. Error        df t value Pr(>|t|)
## (Intercept)      5.747e-02  4.784e-02 1.047e+01   1.201 0.256088
## Y1BASELINE_M_Z   8.299e-01  7.630e-03 6.665e+03 108.762  < 2e-16 ***
## FEMALE           2.874e-02  1.310e-02 6.669e+03   2.194 0.028234 *
## WHITE            5.466e-02  1.600e-02 6.532e+03   3.416 0.000639 ***
## PAIR2           -7.181e-03  6.372e-02 8.848e+00  -0.113 0.912797
## PAIR3           -1.318e-01  7.248e-02 1.300e+01  -1.819 0.092025 .
## PAIR5            3.178e-02  6.541e-02 9.824e+00   0.486 0.637668
## PAIR6            3.525e-02  6.251e-02 8.605e+00   0.564 0.587202
## PAIR7            8.174e-03  6.341e-02 8.339e+00   0.129 0.900499
## PAIR8           -1.507e-02  7.109e-02 1.362e+01  -0.212 0.835229
## PAIR10          -1.005e-01  6.069e-02 7.727e+00  -1.655 0.137798
## PAIR11          -7.152e-02  6.221e-02 8.519e+00  -1.150 0.281549
```

```
## PAIR13                   -8.780e-02  6.705e-02  1.115e+01  -1.309 0.216714
## PAIR15                   -1.980e-01  7.136e-02  1.187e+01  -2.775 0.016952 *
## PAIR19                    5.202e-02  6.696e-02  1.139e+01   0.777 0.453078
## PAIR20                    6.500e-02  6.850e-02  1.208e+01   0.949 0.361257
## PAIR21                   -9.417e-02  6.237e-02  8.334e+00  -1.510 0.168004
## PAIR22                    4.515e-02  6.619e-02  1.075e+01   0.682 0.509538
## PAIR23                   -4.995e-02  6.793e-02  1.175e+01  -0.735 0.476533
## PAIR24                   -1.812e-01  8.199e-02  2.117e+01  -2.210 0.038217 *
## TREATMENT:GRADE_8FALSE -1.226e-02  2.530e-02  1.369e+01  -0.485 0.635609
## TREATMENT:GRADE_8TRUE   6.501e-02  2.610e-02  1.559e+01   2.491 0.024419 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


##
## Correlation matrix not shown by default, as p = 22 > 12.
## Use print(x, correlation=TRUE)  or
##     vcov(x)        if you need it


## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 7 columns / coefficients
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see help('isSingular')
```