Supplement to 'Do Most Students Need In-Person Lectures? A Study of a Large Statistics Class'

Data

The F17 and S18 semesters of STAT 200 had a final exam while the the S19 and F19 semesters of STAT 200 had an average exam score as the final exam. For these reasons we will analyze these data groups separately. A joint analysis will also be considered.

In this supplement we estimate the average treatment effect (ATE) for the causal effect of online learning for the STAT 200 course. To begin our analysis, we first clear R's global environment, load in the required software packages, load in the STAT 200 data, do a bit of data cleaning, and set the random seed so that everyone can reproduce the findings. We consider data summaries before jumping into causal analyses.

```
rm(list=ls())
set.seed(13)
library(tidyverse)
library(foreach)
library(doParallel)
library(betareg)
dat <- read.csv("~/research/online/manuscript/200_1105.csv", header = TRUE)[1:1105, ] %>%
  rename(ObjExam = ObjectiveExam)
dat$ACT <- as.numeric(as.character(dat$ACT))
dat$ACTMajor <- as.numeric(as.character(dat$ACTMajor))
dat$ACTMath <- as.numeric(as.character(dat$ACTMath))
dat$ACTVerbal <- as.numeric(as.character(dat$ACTVerbal))
head(dat)
```

```
##   ObjExam Online Gender International F17 S18 S19 Fa19 FR SO JR ACTMajor
## 1   86.75      1      0             0   1   0   0    0  0  0  1     31.5
## 2   91.57      1      0             0   1   0   0    0  0  1  0     31.5
## 3   86.75      1      0             1   1   0   0    0  0  0  0     33.5
## 4   96.39      0      0             0   1   0   0    0  0  1  0     33.5
## 5   85.54      1      0             0   1   0   0    0  0  0  1     29.5
## 6   78.31      0      0             1   1   0   0    0  1  0  0     31.5
##   Homework  ExamAve Final Testing  ACT ACTMath ACTVerbal HSGPA  X X.1
## 1 94.32643 96.33333 86.75       1 30.0    31.0      59.0  3.06 NA  NA
## 2 99.18116 95.33333 91.57       1 33.2    35.6      60.9  3.40 NA  NA
## 3 95.81584 94.33333 86.75       1 33.9    35.6      64.2    NA NA  NA
## 4 99.34553 99.00000 96.39       1 35.0    36.0      70.0  4.00 NA  NA
## 5 96.46742 92.66666 85.54       1 32.0    35.0      64.0  3.42 NA  NA
## 6 99.06542 87.33333 78.31       1 30.0    35.0      49.0    NA NA  NA
```

Summary statistics

We first reproduce Table 1 in the manuscript which contains summary information for exam scores across semesters and class format.

```r
table1 <- function(xx){
  n <- nrow(xx)
  ExamOL <- xx[xx$Online == 1, ]$ObjExam
  ExamIP <- xx[xx$Online == 0, ]$ObjExam
  nOL <- length(which(xx$Online == 1))
  nIP <- n - nOL
  round(c(n, mean(xx$ObjExam), sd(xx$ObjExam), nOL, nIP,
    mean(ExamOL), mean(ExamIP), mean(ExamOL) - mean(ExamIP),
    sqrt(var(ExamOL)/nOL + var(ExamIP)/nIP)), 2)
}

table_1 <- rbind(
  table1(dat %>% filter(F17 == 1)),
  table1(dat %>% filter(S18 == 1)),
  table1(dat %>% filter(S19 == 1)),
  table1(dat %>% filter(Fa19 == 1))
)

colnames(table_1) <- c("n", "Mean score", "SD", "nOL", "nIP", "MeanOL",
                       "MeanIP", "MeanOL - MeanIP", "SE(diff)")
table_1
```

```
##        n Mean score    SD nOL nIP MeanOL MeanIP MeanOL - MeanIP SE(diff)
## [1,] 271      87.27  8.82 135 136  86.52  88.01           -1.49     1.07
## [2,] 274      85.36 10.31 158 116  86.88  83.29            3.59     1.27
## [3,] 267      83.30 10.84 201  66  84.64  79.19            5.45     1.69
## [4,] 293      86.65 10.16 153 140  87.44  85.78            1.66     1.18
```

We now reproduce the information in Table 2 in the manuscript. This table contains summary information on the covariates used in the analysis. We first compute and reproduce the top half of Table 2.

```r
# permutation test function
one.test <- function(x,y){
  xstar <- sample(x)
  mean(y[xstar==1]) - mean(y[xstar==0])
}

# ACT summary
ACT <- dat[!is.na(dat$ACT), ]
nACT_OL <- nrow(ACT[ACT$Online == 1, ])
nACT_IP <- nrow(ACT[ACT$Online == 0, ])
ACT_OL <- ACT[ACT$Online == 1, ]$ACT
ACT_IP <- ACT[ACT$Online == 0, ]$ACT
ACT_pval <- length(which((mean(ACT_OL) - mean(ACT_IP)) -
  replicate(10000, one.test(x = ACT$Online, y = ACT$ACT)) < 0)) / 10000 * 2

# ACT math summary
ACTMath <- dat[!is.na(dat$ACTMath), ]
nACTMath_OL <- nrow(ACTMath[ACTMath$Online == 1, ])
nACTMath_IP <- nrow(ACTMath[ACTMath$Online == 0, ])
ACTMath_OL <- ACTMath[ACTMath$Online == 1, ]$ACTMath
ACTMath_IP <- ACTMath[ACTMath$Online == 0, ]$ACTMath
ACTMath_pval <- length(which((mean(ACTMath_OL) - mean(ACTMath_IP)) -
  replicate(10000, one.test(x = ACTMath$Online, y = ACTMath$ACTMath)) < 0)) / 10000 * 2
```

```r
# ACT verbal summary
ACTV <- dat[!is.na(dat$ACTV), ]
nACTV_OL <- nrow(ACTV[ACTV$Online == 1, ])
nACTV_IP <- nrow(ACTV[ACTV$Online == 0, ])
ACTV_OL <- ACTV[ACTV$Online == 1, ]$ACTV
ACTV_IP <- ACTV[ACTV$Online == 0, ]$ACTV
ACTV_pval <- length(which((mean(ACTV_OL) - mean(ACTV_IP)) -
  replicate(10000, one.test(x = ACTV$Online, y = ACTV$ACTV)) < 0)) / 10000 * 2

# ACT major summary
ACTMajor <- dat[!is.na(dat$ACTMajor), ]
nACTMajor_OL <- nrow(ACTMajor[ACTMajor$Online == 1, ])
nACTMajor_IP <- nrow(ACTMajor[ACTMajor$Online == 0, ])
ACTMajor_OL <- ACTMajor[ACTMajor$Online == 1, ]$ACTMajor
ACTMajor_IP <- ACTMajor[ACTMajor$Online == 0, ]$ACTMajor
ACTMajor_pval <- length(which((mean(ACTMajor_OL) - mean(ACTMajor_IP)) -
  replicate(10000, one.test(x = ACTMajor$Online, y = ACTMajor$ACTMajor)) < 0)) / 10000 * 2

# HSGPA summary
HSGPA <- dat[!is.na(dat$HSGPA), ]
nHSGPA_OL <- nrow(HSGPA[HSGPA$Online == 1, ])
nHSGPA_IP <- nrow(HSGPA[HSGPA$Online == 0, ])
HSGPA_OL <- HSGPA[HSGPA$Online == 1, ]$HSGPA
HSGPA_IP <- HSGPA[HSGPA$Online == 0, ]$HSGPA
HSGPA_pval <- length(which((mean(HSGPA_OL) - mean(HSGPA_IP)) -
  replicate(10000, one.test(x = HSGPA$Online, y = HSGPA$HSGPA)) < 0)) / 10000 * 2
```

Here is the top half of Table 2.

```r
table2_top <- round(rbind(
  c(nrow(ACT), nACT_OL, nACT_IP, mean(ACT$ACT), sd(ACT$ACT),
    mean(ACT_OL) - mean(ACT_IP), sqrt(var(ACT_OL)/nACT_OL + var(ACT_IP)/nACT_IP),
    ACT_pval),
  c(nrow(ACTMath), nACTMath_OL, nACTMath_IP, mean(ACTMath$ACTMath), sd(ACTMath$ACTMath),
    mean(ACTMath_OL) - mean(ACTMath_IP),
    sqrt(var(ACTMath_OL)/nACTMath_OL + var(ACTMath_IP)/nACTMath_IP),
    ACTMath_pval),
  c(nrow(ACTV), nACTV_OL, nACTV_IP, mean(ACTV$ACTV), sd(ACTV$ACTV),
    mean(ACTV_OL) - mean(ACTV_IP), sqrt(var(ACTV_OL)/nACTV_OL + var(ACTV_IP)/nACTV_IP),
    ACTV_pval),
  c(nrow(ACTMajor), nACTMajor_OL, nACTMajor_IP, mean(ACTMajor$ACTMajor), sd(ACTMajor$ACTMajor),
    mean(ACTMajor_OL) - mean(ACTMajor_IP),
    sqrt(var(ACTMajor_OL)/nACTMajor_OL + var(ACTMajor_IP)/nACTMajor_IP),
    ACTMajor_pval),
  c(nrow(HSGPA), nHSGPA_OL, nHSGPA_IP, mean(HSGPA$HSGPA), sd(HSGPA$HSGPA),
    mean(HSGPA_OL) - mean(HSGPA_IP),
    sqrt(var(HSGPA_OL)/nHSGPA_OL + var(HSGPA_IP)/nHSGPA_IP),
    HSGPA_pval)), 3)

colnames(table2_top) <- c("n", "nOL", "nIP", "Mean", "SD", "MeanOL - MeanIP",
  "SE(diff)", "Random null p-value")
table2_top
```

```
##          n nOL nIP   Mean    SD MeanOL - MeanIP SE(diff) Random null p-value
```

```
## [1,] 1105 647 458 30.507 3.465          0.382    0.212              0.065
## [2,] 1105 647 458 32.162 4.006          0.448    0.247              0.064
## [3,] 1105 647 458 58.772 8.001          0.603    0.488              0.216
## [4,] 1105 647 458 30.180 2.621          0.475    0.157              0.003
## [5,]  675 377 298  3.516 0.349          0.036    0.027              0.183
```

We now compute and reproduce the top half of Table 2.

```
# International summary
prop_Int_OL <- sum(dat[dat$Online == 1, ]$International)/sum(dat$Online)
prop_Int_IP <- sum(dat[dat$Online == 0, ]$International)/sum(1-dat$Online)
International <- c(sum(dat$International), sum(dat[dat$Online == 1, ]$International),
  sum(dat[dat$Online == 0, ]$International),
  prop_Int_OL, prop_Int_IP,
  length(which(prop_Int_OL - prop_Int_IP -
    replicate(10000, one.test(x = dat$Online, y = dat$International)) < 0)) / 10000 * 2)

# Gender summary
foo <- dat %>% mutate(Gender = as.numeric(as.character(Gender)))
prop_Gen_OL <- sum(foo[foo$Online == 1, ]$Gender)/sum(foo$Online)
prop_Gen_IP <- sum(foo[foo$Online == 0, ]$Gender)/sum(1-foo$Online)
Gender <- c(sum(foo$Gender), sum(foo[foo$Online == 1, ]$Gender),
  sum(foo[foo$Online == 0, ]$Gender), prop_Gen_OL, prop_Gen_IP,
  length(which(prop_Gen_OL - prop_Gen_IP -
    replicate(10000, one.test(x = foo$Online, y = foo$Gender)) < 0)) / 10000 * 2)

# Freshmen summary
prop_FR_OL <- sum(dat[dat$Online == 1, ]$FR)/sum(dat$Online)
prop_FR_IP <- sum(dat[dat$Online == 0, ]$FR)/sum(1-dat$Online)
FR <- c(sum(dat$FR), sum(dat[dat$Online == 1, ]$FR),
  sum(dat[dat$Online == 0, ]$FR),
  prop_FR_OL, prop_FR_IP,
  length(which(prop_FR_OL - prop_FR_IP -
    replicate(10000, one.test(x = dat$Online, y = dat$FR)) > 0)) / 10000 * 2)

# Sophomore summary
prop_SO_OL <- sum(dat[dat$Online == 1, ]$SO)/sum(dat$Online)
prop_SO_IP <- sum(dat[dat$Online == 0, ]$SO)/sum(1-dat$Online)
SO <- c(sum(dat$SO), sum(dat[dat$Online == 1, ]$SO),
  sum(dat[dat$Online == 0, ]$SO),
  prop_SO_OL, prop_SO_IP,
  length(which(prop_SO_OL - prop_SO_IP -
    replicate(10000, one.test(x = dat$Online, y = dat$SO)) > 0)) / 10000 * 2)

# Junior summary
prop_JR_OL <- sum(dat[dat$Online == 1, ]$JR)/sum(dat$Online)
prop_JR_IP <- sum(dat[dat$Online == 0, ]$JR)/sum(1-dat$Online)
JR <- c(sum(dat$JR), sum(dat[dat$Online == 1, ]$JR),
  sum(dat[dat$Online == 0, ]$JR),
  prop_JR_OL, prop_JR_IP,
  length(which(prop_JR_OL - prop_JR_IP -
    replicate(10000, one.test(x = dat$Online, y = dat$JR)) < 0)) / 10000 * 2)

# Senior summary
foo <- dat %>% mutate(SR = ifelse(FR + SO + JR == 0,1,0))
```

```r
prop_SR_OL <- sum(foo[foo$Online == 1, ]$SR)/sum(foo$Online)
prop_SR_IP <- sum(foo[foo$Online == 0, ]$SR)/sum(1-foo$Online)
SR <- c(sum(foo$SR), sum(foo[foo$Online == 1, ]$SR),
  sum(foo[foo$Online == 0, ]$SR),
  prop_SR_OL, prop_SR_IP,
  length(which(prop_SR_OL - prop_SR_IP -
    replicate(10000, one.test(x = foo$Online, y = foo$SR)) < 0)) / 10000 * 2)


table2_bottom <- round(rbind(International, Gender, FR, SO, JR, SR), 3)
colnames(table2_bottom) <- c("n", "nOL", "nIP", "% OL", "% IP", "Random null p-value")
table2_bottom
```

```
##                 n nOL nIP  % OL  % IP Random null p-value
## International 435 270 165 0.417 0.360               0.048
## Gender        474 290 184 0.448 0.402               0.119
## FR            155  40 115 0.062 0.251               0.000
## SO            407 212 195 0.328 0.426               0.001
## JR            312 226  86 0.349 0.188               0.000
## SR            231 169  62 0.261 0.135               0.000
```

Causal Analysis

Setting

In this analysis we estimate the causal effect of online learning. The response variable Y is the comprehensive 3-hour Final that was Scantron graded at the end of each semester, and the treatment variable A is the presence or absence of the online course. The online course is considered the treatment and the regular in-class course is the control. Most of the causal inference methodology follows from [1].

We estimate the causal effect of online learning by estimating the average treatment effect (ATE) using inverse propensity score weighting methods. The form of this estimator of the ATE is

$$\widehat{\text{ATE}} = \frac{1}{n}\sum_{i=1}^{n}\left(w_i A_i Y_i - w_i(1-A_i)Y_i\right),$$

where the weights w_i are a functions of the propensity scores $\hat{p}_i = \widehat{P}(A_i = 1|X_i)$, $i = 1,\dots,n$. The propensity scores are each subjects conditional probability of belonging to the treatment group given their covariate information. We will estimate the propensity scores using a logistic regression model.

The most basic inverse propensity score weighted estimator has weights w_i of the form

$$w_i = \frac{A_i}{\hat{p}_i} + \frac{1-A_i}{1-\hat{p}_i}.$$

This estimator is unstable, it does not create balanced pseudo populations, and is inappropriate for continuous predictors in our analysis. We consider an alternative stable estimator of the ATE (referred to as ATE_{IPW} in the manuscript) mentioned in [1,3,4]. The alternative stable estimator of the ATE [1,3,4] has conventional inverse propensity score weights with changes being made to the aggregation. This estimator of the ATE is

$$\widehat{\text{ATE}}_{\text{alt}} = \left(\sum_{i=1}^{n}\frac{A_i}{\hat{p}_i}\right)^{-1}\sum_{i=1}^{n} w_i A_i Y_i - \left(\sum_{i=1}^{n}\frac{1-A_i}{1-\hat{p}_i}\right)^{-1}\sum_{i=1}^{n} w_i(1-A_i)Y_i.$$

The estimators $\widehat{\text{ATE}}$ and $\widehat{\text{ATE}}_{\text{alt}}$ are consistent if the propensity score model is correctly specified, so should be approximately unbiased in finite samples [3]. We will also consider a double robust (DR) estimator of the

ATE compared in [3] and developed in [5]. This estimator is

$$\widehat{\text{ATE}}_{\text{DR}} = \frac{1}{n}\sum_{i=1}^{n}\frac{A_iY_i - (A_i - \hat{p}_i)m_1(X_i,\hat{\alpha}_1)}{\hat{p}_i} - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-A_i)Y_i + (A_i - \hat{p}_i)m_0(X_i,\hat{\alpha}_0)}{1-\hat{p}_i},$$

where $m_a(X, \alpha_a) = E(Y|A = a, X)$ is the regression of the response on X in each level of $A \in \{0, 1\}$, depending on parameters α_a, and $\hat{\alpha}_a$ is the estimator for α_a based on the data from subjects with $A = a$ [3].

We will also consider a linear regression model to estimate the effect of online learning. The term in this regression coefficient vector corresponding to the effect of online learning will be denoted as β_{online}.

We now display summary information for the inverse propensity score weights for each of the considered data sets.

```
# basic wrangling
dat_small <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                            International, F17, S18, S19, Fa19, FR, SO, JR)
dat_first <- dat_small %>% filter(F17 + S18 == 1)
dat_second <- dat_small %>% filter(S19 + Fa19 == 1)

# propensity weights for first two semesters
m_first <- glm(Online ~. -Fa19 -S19 -F17, data = dat_first, family = "binomial")
trt_first <- dat_first$Online
p_first <- predict(m_first, type = "response")
w_first <- trt_first / p_first  + (1 - trt_first) / (1 - p_first)
summary(w_first)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.097   1.358   1.635   1.994   2.386   8.911
```

```
# propensity weights for last two semesters
m_second <- glm(Online ~. -S18 -Fa19 -F17, data = dat_second, family = "binomial")
trt_second <- dat_second$Online
p_second <- predict(m_second, type = "response")
w_second <- trt_second / p_second  + (1 - trt_second) / (1 - p_second)
summary(w_second)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.104   1.329   1.561   2.016   2.180  10.005
```

```
# propensity weights for all semesters
m_full <- glm(Online ~. -Fa19, data = dat_small, family = "binomial")
trt_full <- dat_small$Online
p_full <- predict(m_full, type = "response")
w_full <- trt_full / p_full  + (1 - trt_full) / (1 - p_full)
summary(w_full)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.108   1.342   1.620   2.010   2.274  10.684
```

Estimate the effect of online learning

We now estimate the ATE for the causal effect of online learning using the two estimators of the ATE that are presented in the Setting. We also estimate the effect of online learning using a linear regression model and beta regression model.

```
effects <- function(dat){
```

```r
# make numeric variables numeric
dat$ACT <- as.numeric(as.character(dat$ACT))
dat$ACTMajor <- as.numeric(as.character(dat$ACTMajor))
dat$ACTMath <- as.numeric(as.character(dat$ACTMath))

# basic data wrangling (will induce non problematic rank deficiency)
dat_small <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR) %>%
filter(ACTMajor > 0 & !is.na(ACTMajor) &
       ACTMath > 0 & !is.na(ACTMath))

## estimate ATE
m <- glm(Online ~., data = dat_small, family = "binomial")
trt <- dat_small$Online
preds <- predict(m, type = "response")
#w <- trt / preds - (1 - trt) / (1 - preds)
#ATE <- mean(w * dat$ObjExam)

## estimate alternative stabilized ATE
scale_alt_trt <- 1 / sum(trt / preds)
scale_alt_notrt <- 1 / sum((1 - trt)/(1 - preds))
ATE_alt <- scale_alt_trt * sum((1/preds) * trt * dat$ObjExam) -
  scale_alt_notrt * sum((1-trt)/(1-preds)* dat$ObjExam)

## estimate DR version ATE (will induce non problematic rank deficiency)
m_trt <- lm(ObjExam ~ ACTMath + ACTMajor + ACT + International + Gender +
              FR + SO + JR + F17 + S18 + S19,
            data = dat[trt == 1, ])
Y_trt <- predict(m_trt, newdata = dat)

m_notrt <- lm(ObjExam ~ ACTMath + ACTMajor + ACT + International + Gender +
              FR + SO + JR + F17 + S18 + S19,
            data = dat[trt == 0, ])
Y_notrt <- predict(m_notrt, newdata = dat)
ATE_DR <- mean( (dat$ObjExam * trt - (trt - preds) * Y_trt) / preds -
  (dat$ObjExam * (1 - trt) + (trt - preds)*Y_notrt) / (1 - preds))

## regression model (will induce non problematic rank deficiency)
m <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
          FR + SO + JR + F17 + S18 + S19,
        data = dat)

# beta regression
#dat_Y <- dat %>% mutate(Y = ObjExam/100)
#dat_Y$Y[which(dat$Y == 1)] <- 0.9999
#dat_Y$Y[which(dat$Y == 0)] <- 0.0001
#m_beta <- betareg(Y ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
#          FR + SO + JR + F17 + S18 + S19 | ACTMath + S18 + S19 + Fa19 + FR + SO + JR,
#          data = dat, link = "log", link.phi = "sqrt"),

#c("ATE" = ATE, "ATE_alt" = ATE_alt, "ATE_DR" = ATE_DR,
#  "Online" = as.numeric(m$coef[2]))
#c("Online" = as.numeric(m$coef[2]), "Online_Beta" = summary(m_beta)$coe$me[2,1],
```

```
  #  "ATE_alt" = ATE_alt, "ATE_DR" = ATE_DR)
  c("Online" = as.numeric(m$coef[2]),  "ATE_alt" = ATE_alt,
    "ATE_DR" = ATE_DR)
}
```

Here are the results

```
## F17 semesters
effects(dat = dat %>% filter(F17 == 1))
```

```
##    Online   ATE_alt    ATE_DR
## -1.743206 -1.140049 -1.550791
```

```
## S18 semesters
effects(dat = dat %>% filter(S18 == 1))
```

```
##   Online  ATE_alt   ATE_DR
## 1.748684 1.820145 1.710989
```

```
## S19 semesters
effects(dat = dat %>% filter(S19 == 1))
```

```
##   Online  ATE_alt   ATE_DR
## 2.066649 2.165275 2.223289
```

```
## Fa19 semesters
effects(dat = dat %>% filter(Fa19 == 1))
```

```
##   Online  ATE_alt   ATE_DR
## 1.185532 1.393317 1.046269
```

```
## First two semesters
effects(dat = dat %>% filter(F17 + S18 == 1))
```

```
##      Online     ATE_alt      ATE_DR
## -0.01399159  0.15698401  0.04686280
```

```
## Second two semesters
effects(dat = dat %>% filter(S19 + Fa19 == 1))
```

```
##   Online  ATE_alt   ATE_DR
## 1.535600 1.661728 1.356780
```

```
## All semesters - S19
effects(dat = dat %>% filter(F17 + S18 + Fa19 == 1))
```

```
##    Online   ATE_alt    ATE_DR
## 0.3491160 0.4673822 0.3913971
```

```
## All semesters - F17
effects(dat = dat %>% filter(S18 + Fa19 + S19 == 1))
```

```
##   Online  ATE_alt   ATE_DR
## 1.462210 1.505607 1.421703
```

```
## All semesters
effects(dat = dat)
```

```
##    Online   ATE_alt    ATE_DR
## 0.6253318 0.7467678 0.5839065
```

Balance of methods

We now check the balance of the pseudo-samples created by the weighting regimes used to construct the classical ATE and $\widehat{\text{ATE}}_{\text{alt}}$. The alternate weighting regime used to construct $\widehat{\text{ATE}}_{\text{alt}}$ corrects for slight inbalances that are observed in the classical weighting approach.

```
w_ATE <- w_full
scale_alt_trt <- 1 / sum(trt_full / p_full)
scale_alt_notrt <- 1 / sum((1 - trt_full)/(1 - p_full))
w_ATEalt <- scale_alt_trt * (1/p_full) * trt_full +
    scale_alt_notrt * (1-trt_full)/(1-p_full)

dat_weights <- cbind(dat_small,w_ATE,w_ATEalt)
lapply(split(dat_weights, as.factor(dat_weights$Online)),
       function(xx){
         c('classical' = sum(xx$w_ATE), 'alternative' =sum(xx$w_ATEalt))
})
```

```
## $`0`
##   classical alternative
##    1128.493       1.000
##
## $`1`
##   classical alternative
##    1092.335       1.000
```

We now take a more detailed look at balance of the stabilized weights across both categorical and discretized continuous covariates. The balance across treatment groups is solid overall (as judged by 100 * standardized difference being less than 10 [1,7]).

```
dat_ATEalt <- data.frame(dat, weights = w_ATEalt)
balance <- function(w, x){
  mean_w <- sum(x*w)/sum(w)
  sd_w <- sum(w) / (sum(w)^2 - sum(w^2) ) * sum(w*(x - mean_w)^2)
  c(mean_w, sd_w)
}

## balance in Gender
as_tibble(dat_ATEalt %>% mutate(Gender = as.numeric(Gender)) %>% group_by(Online) %>%
  summarise(balance = balance(weights, Gender))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] 1.45964
```

```
## balance in International
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, International))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] 2.822587
```

```
## balance in Freshmen
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, FR))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] 6.746332
```

```r
## balance in Sophomores
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, SO))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -8.613175
```

```r
## balance in Juniors
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, JR))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -2.21941
```

```r
## balance in Fall 17
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, F17))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] 2.265054
```

```r
## balance in Spring 18
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, S18))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -5.413468
```

```r
## balance in Spring 19
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, S19))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] 9.838854
```

```r
## balance in Fall 19
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, Fa19))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -6.679236
```

```r
## balance in ACT
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, ACT))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -0.4801436
```

```r
## balance in ACTMajor
as_tibble(dat_ATEalt %>% group_by(Online) %>%
```

```r
  summarise(balance = balance(weights, ACTMajor))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -0.4940572
```

```r
## balance in ACT Math
as_tibble(dat_ATEalt %>% group_by(Online) %>%
  summarise(balance = balance(weights, ACTMath))) %>%
  summarise(d = 100 * (balance[1] - balance[3])/(sqrt( (balance[2]^2 + balance[4]^2)/2 ))) %>%
  pull(d)
```

```
## [1] -0.3387809
```

Bootstrapping estimates of the ATE

We now construct a nonparametric bootstrap procedure to estimate 95% percentile intervals for each estimator for the online effect of learning.

```r
ncores <- detectCores() - 1
boot <- function(data, B, ncores){
  registerDoParallel(cores=ncores)
  out <- foreach(j = 1:B, .combine = 'rbind') %dopar% {
    set.seed(j)
    ind <- sample(1:nrow(data), replace = TRUE)
    effects(dat = data[ind, ])
  }
  out
}
```

Here are 95% bootstrap percentile intervals for each estimator of the effect of online learning in each of our analyses.

```r
system.time(boot_F17 <- boot(data = dat %>% filter(F17 == 1),
                  B = 1e4, ncores = ncores))
```

F17 Semester

```
##    user  system elapsed
## 387.543   5.788  27.850
```

```r
apply(boot_F17, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##            Online    ATE_alt     ATE_DR
## 2.5%  -3.48379169 -2.7301497 -3.2417210
## 97.5%  0.05851159  0.5889001  0.3857383
```

```r
system.time(boot_S18 <- boot(data = dat %>% filter(S18 == 1),
                  B = 1e4, ncores = ncores))
```

S18 Semester

```
##    user  system elapsed
## 360.253   5.567  25.910
```

```r
apply(boot_S18, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##           Online    ATE_alt     ATE_DR
## 2.5%  -0.4233774 -0.3707566 -0.5273323
## 97.5%  4.0401997  4.1138883  4.0802369
```

```r
system.time(boot_S19 <- boot(data = dat %>% filter(S19 == 1),
                  B = 1e4, ncores = ncores))
```

S19 Semester

```
##    user  system elapsed
## 382.319   6.294  28.754
```

```r
apply(boot_S19, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##           Online    ATE_alt     ATE_DR
## 2.5%  -0.7768883 -0.6036961 -0.5504172
## 97.5%  4.9637091  5.1325361  5.1235144
```

```r
system.time(boot_Fa19 <- boot(data = dat %>% filter(Fa19 == 1),
                  B = 1e4, ncores = ncores))
```

Fa19 Semester

```
##    user  system elapsed
## 608.270   8.659  45.211
```

```r
apply(boot_Fa19, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##           Online    ATE_alt    ATE_DR
## 2.5%  -0.9101499 -0.8305599 -1.087487
## 97.5%  3.2641489  3.6587516  3.137763
```

```r
system.time(boot_full <- boot(data = dat %>% filter(F17 + S18 == 1),
                  B = 1e4, ncores = ncores))
```

First two semesters (F17 and S18)

```
##    user  system elapsed
## 547.382   9.090  40.042
```

```r
apply(boot_full, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##         Online   ATE_alt    ATE_DR
## 2.5%  -1.42421 -1.209527 -1.349317
## 97.5%  1.43115  1.582026  1.494199
```

```r
system.time(boot_full <- boot(data = dat %>% filter(S19 + Fa19 == 1),
                  B = 1e4, ncores = ncores))
```

Second two semesters (S19 and F19)

```
##    user  system elapsed
## 536.176   8.168  39.077
```

```r
apply(boot_full, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##            Online    ATE_alt     ATE_DR
## 2.5%  -0.09267256 -0.0702575 -0.2756142
## 97.5%  3.21169980  3.4491651  3.0111125
```

```r
system.time(boot_full <- boot(data = dat %>% filter(S18 + S19 + Fa19 == 1),
                  B = 1e4, ncores = ncores))
```

All semesters minus F17

```
##     user   system  elapsed
## 2390.312   66.848  193.993
```

```r
apply(boot_full, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##          Online   ATE_alt    ATE_DR
## 2.5%  0.1491896 0.1622605 0.1273671
## 97.5% 2.7846294 2.8200060 2.6956848
```

```r
system.time(boot_full <- boot(data = dat, B = 1e4, ncores = ncores))
```

Joint analysis

```
##    user  system elapsed
## 620.647   9.458  44.809
```

```r
apply(boot_full, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##           Online    ATE_alt     ATE_DR
## 2.5%  -0.4390801 -0.3428699 -0.4400942
## 97.5%  1.7067833  1.8635747  1.6272127
```

Checking HS GPA

We now restrict attention to individuals who had a high school GPA recorded. Only about half of the individuals in the original study have a recorded high school GPA value.

```r
dat_HS <- dat %>% filter(!is.na(HSGPA))
nrow(dat_HS)
```

```
## [1] 675
```

The R function below computes the differences between the online learning estimates of models that include high school GPA and models that ignore high school GPA.

```r
effects_HS <- function(dat){

  # make numeric variables numeric
  dat$ACT <- as.numeric(as.character(dat$ACT))
  dat$ACTMajor <- as.numeric(as.character(dat$ACTMajor))
  dat$ACTMath <- as.numeric(as.character(dat$ACTMath))

  # basic data wrangling
  dat_small <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR) %>%
```

```r
filter(ACTMajor > 0 & !is.na(ACTMajor) &
       ACTMath > 0 & !is.na(ACTMath))

dat_HS_small <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR, HSGPA) %>%
filter(ACTMajor > 0 & !is.na(ACTMajor) &
       ACTMath > 0 & !is.na(ACTMath))

## estimate ATE for both data sets
m <- glm(Online ~., data = dat_small, family = "binomial")
trt <- dat_small$Online
preds <- predict(m, type = "response")
w <- trt / preds - (1 - trt) / (1 - preds)
#ATE <- mean(w * dat$ObjExam)
m_HS <- glm(Online ~., data = dat_HS_small, family = "binomial")
trt_HS <- dat_HS_small$Online
preds_HS <- predict(m_HS, type = "response")
w_HS <- trt_HS / preds_HS - (1 - trt_HS) / (1 - preds_HS)
#ATE_HS <- mean(w_HS * dat_HS$ObjExam)
#ATE_diff <- ATE - ATE_HS

## estimate alternative stabilized ATE for both data sets
scale_alt_trt <- 1 / sum(trt / preds)
scale_alt_notrt <- 1 / sum((1 - trt)/(1 - preds))
ATE_alt <- scale_alt_trt * sum((1/preds) * trt * dat$ObjExam) -
  scale_alt_notrt * sum((1-trt)/(1-preds)* dat$ObjExam)

scale_alt_trt_HS <- 1 / sum(trt_HS / preds_HS)
scale_alt_notrt_HS <- 1 / sum((1 - trt_HS)/(1 - preds_HS))
ATE_alt_HS <- scale_alt_trt_HS * sum((1/preds_HS) * trt_HS * dat$ObjExam) -
  scale_alt_notrt_HS * sum((1-trt_HS)/(1-preds_HS)* dat$ObjExam)
ATE_alt_diff <- ATE_alt_HS - ATE_alt

## estimate DR version ATE for both data sets
m_trt <- lm(ObjExam ~ ACTMath + ACTMajor + ACT + International + Gender +
              FR + SO + JR + F17 + S18 + S19,
            data = dat[trt == 1, ])
Y_trt <- predict(m_trt, newdata = dat)

m_notrt <- lm(ObjExam ~ ACTMath + ACTMajor + ACT + International + Gender +
              FR + SO + JR + F17 + S18 + S19,
            data = dat[trt == 0, ])
Y_notrt <- predict(m_notrt, newdata = dat)
ATE_DR <- mean( (dat$ObjExam * trt - (trt - preds) * Y_trt) / preds -
  (dat$ObjExam * (1 - trt) + (trt - preds)*Y_notrt) / (1 - preds))

m_trt_HS <- lm(ObjExam ~ ACTMath + ACTMajor + ACT + International + Gender +
              FR + SO + JR + F17 + S18 + S19 + HSGPA,
            data = dat[trt == 1, ])
Y_trt_HS <- predict(m_trt_HS, newdata = dat)

m_notrt_HS <- lm(ObjExam ~ ACTMath + ACTMajor + ACT + International + Gender +
              FR + SO + JR + F17 + S18 + S19 + HSGPA,
```

```r
              data = dat[trt == 0, ])
  Y_notrt_HS <- predict(m_notrt_HS, newdata = dat)
  ATE_DR_HS <- mean( (dat$ObjExam * trt_HS - (trt_HS - preds_HS) * Y_trt_HS) / preds_HS -
    (dat$ObjExam * (1 - trt_HS) + (trt_HS - preds_HS)*Y_notrt_HS) / (1 - preds_HS))
  ATE_DR_diff <- ATE_DR_HS - ATE_DR

  ## regression model
  m <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
            FR + SO + JR + F17 + S18 + S19,
          data = dat)
  m_HS <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
            FR + SO + JR + F17 + S18 + S19 + HSGPA,
          data = dat)
  diff_beta <- as.numeric(m_HS$coef[2]) - as.numeric(m$coef[2])

  #c("ATE_diff" = ATE_diff, "ATE_alt_diff" = ATE_alt_diff, "ATE_DR_diff" = ATE_DR_diff,
  #  "Online_diff" = diff_beta)
  c("Online_diff" = diff_beta, "ATE_alt_diff" = ATE_alt_diff,
    "ATE_DR_diff" = ATE_DR_diff)
}
```

We can see that the effect of High School GPA does not greatly effect the estimators of the online learning effect.

```r
effects_HS(dat_HS)
```

```
##  Online_diff ATE_alt_diff  ATE_DR_diff
##   0.10894243   0.08516013   0.14746157
```

Nonparametric bootstrapping reveals that the distribution of the discrepancy between estimators including or ignoring the high school GPA variable is not practically significant for any of the estimators.

```r
ncores <- detectCores() - 1
boot_HS <- function(data, B, ncores){
  registerDoParallel(cores=ncores)
  n <- nrow(data)
  out <- foreach(j = 1:B, .combine = 'rbind') %dopar% {
    set.seed(j)
    effects_HS(dat = data[sample(1:n, replace = TRUE), ])
  }
  out
}
```

```r
system.time(boot_full_HS <- boot_HS(data = dat_HS, B = 1e4, ncores = ncores))
```

```
##     user   system  elapsed
## 1098.002   11.853   78.819
```

```r
apply(boot_full_HS, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))
```

```
##       Online_diff ATE_alt_diff ATE_DR_diff
## 2.5%    -0.3102021   -0.2260648  -0.2090609
## 97.5%    0.5471583    0.4209767   0.5532218
```

Multiple linear regression inference

We expand on the multiple linear regression analysis of the full data set. The following is the summary output for the full model considered. The effect for online learning can be seen below (effect size of 0.625, p-value of 0.254).

```
m_MLR <- lm(ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
            FR + SO + JR + ACTMajor + ACT + ACTMath, data = dat)
summary(m_MLR)
```

```
##
## Call:
## lm(formula = ObjExam ~ Online + Gender + International + F17 +
##     S18 + S19 + FR + SO + JR + ACTMajor + ACT + ACTMath, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -41.424  -3.922   1.445   5.480  21.984
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   29.47928    3.46130   8.517  < 2e-16 ***
## Online         0.62533    0.54751   1.142 0.253651
## Gender1        0.40658    0.52320   0.777 0.437263
## International  0.86411    0.62302   1.387 0.165735
## F17            0.05405    0.70312   0.077 0.938739
## S18           -1.09210    0.69766  -1.565 0.117787
## S19           -2.60146    0.71287  -3.649 0.000275 ***
## FR            -1.89505    0.92222  -2.055 0.040129 *
## SO            -0.26733    0.71540  -0.374 0.708715
## JR            -0.49974    0.72236  -0.692 0.489198
## ACTMajor       0.51778    0.11152   4.643 3.85e-06 ***
## ACT            0.22348    0.12665   1.765 0.077918 .
## ACTMath        1.06518    0.12319   8.647  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.267 on 1092 degrees of freedom
## Multiple R-squared:  0.3451, Adjusted R-squared:  0.3379
## F-statistic: 47.95 on 12 and 1092 DF,  p-value: < 2.2e-16
```

Note that classical inference for the treatment effect may be misleading since the normal regression assumptions are not perfectly satisfied. In particular, there is non-constant variance and large negative residuals are more extreme than what is expected.

Moreover, we see that large negative and large positive residuals are, respectively, associated with primarily low and high exam scores. We also note that the response variable is upper bounded by the value of 100. This constraint partially motivates the observed departures from residual modeling assumptions.

```
summary(dat[as.numeric(scale(m_MLR$residuals)) < -1, ]$ObjExam)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   36.90   63.87   71.39   69.50   77.38   84.43
```

```
summary(dat[as.numeric(scale(m_MLR$residuals)) < -2, ]$ObjExam)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   36.90   54.76   59.52   58.81   65.48   73.05
```

```r
summary(dat[as.numeric(scale(m_MLR$residuals)) > 1, ]$ObjExam)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   78.31   89.29   93.44   92.79   96.39  100.00
```

```r
summary(dat[as.numeric(scale(m_MLR$residuals)) > 2, ]$ObjExam)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   89.45   90.29   93.94   93.58   95.31  100.00
```

Earlier we conducted inference via a nonparametric bootstrap. This procedure is appropriate considering these model assumption violations. However, we note that inference from the multiple linear regression model does not deviate too far from the nonparametric bootstrap procedure implemented above.

```r
inference <- cbind(apply(boot_full, 2, function(xx) quantile(xx, probs = c(0.025, 0.975)))[, 1],
coef(summary(m_MLR))[2,1] + qnorm(c(0.025,0.975)) * coef(summary(m_MLR))[2,2])
colnames(inference) <- c("bootstrap", "lm")
inference
```

```
##        bootstrap         lm
## 2.5%  -0.4390801 -0.4477762
## 97.5%  1.7067833  1.6984397
```

We will consider a residual bootstrap and parametric bootstrap procedure. The code below performances these techniques. The conclusions from these bootstrap procedures are the same as the original nonparametric bootstrap procedure considered previously (included again for completeness). Note that the table produced below includes bootstrap intervals for all regression coefficients.

```r
p <- length(m_MLR$coefficients)
bootMLR <- function(B, dat){
  m <- lm(ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
            FR + SO + JR + ACTMajor + ACT + ACTMath,
          data = dat, x = TRUE, y = TRUE)
  y <- m$y
  modmat <- m$x
  beta <- m$coef
  sd <- summary(m)$sigma
  n <- nrow(dat)
  resids <- m$residuals

  registerDoParallel(cores=ncores)
  out <- foreach(j = 1:B, .combine = 'rbind') %dopar% {
    set.seed(j)

    ## nonparametric bootstrap
    index <- sample(1:n, replace = TRUE)
    betastar_nonparam <- as.numeric(solve(crossprod(modmat[index, ])) %*%
      t(modmat[index, ]) %*% y[index])

    ## residual bootstrap
    ystar_resid <- modmat %*% beta + resids[index]
    betastar_resid <- as.numeric(solve(crossprod(modmat)) %*% t(modmat) %*% ystar_resid)

    ## parametric bootstrap
    ystar_param <- modmat %*% beta + rnorm(n, sd = sd)
    betastar_param <- as.numeric(solve(crossprod(modmat)) %*% t(modmat) %*% ystar_param)
```

```
    c("beta_nonparam" = betastar_nonparam,
      "beta_resid" = betastar_resid,
      "beta_param" = betastar_param)
  }
  out
}
```

All bootstrap procedures are ran below.

```
system.time(foo <- bootMLR(B = 1000, dat = dat))
```

```
##    user  system elapsed
##   4.237   2.288   0.701
```

The complete table of bootstrap percentile estimates is made below.

```
bar <- as.numeric(apply(foo, 2, function(xx) quantile(xx, probs = c(0.025, 0.975))))
table <- cbind(matrix(bar[1:(2*p)], nrow = p, ncol = 2, byrow = TRUE),
  matrix(bar[(2*p+1):(4*p)], nrow = p, ncol = 2, byrow = TRUE),
  matrix(bar[(4*p+1):(6*p)], nrow = p, ncol = 2, byrow = TRUE))
rownames(table) <- names(m_MLR$coef)
colnames(table) <- c("lwr_nonparam", "upr_nonparam", "lwr_resid",
                     "upr_resid", "lwr_param", "upr_param")
round(table, 3)
```

```
##                 lwr_nonparam upr_nonparam lwr_resid upr_resid lwr_param upr_param
## (Intercept)           22.524       36.053    22.920    36.495    22.816    36.456
## Online                -0.364        1.683    -0.463     1.669    -0.426     1.765
## Gender1               -0.631        1.429    -0.622     1.428    -0.591     1.399
## International         -0.258        2.005    -0.363     2.069    -0.316     2.093
## F17                   -1.279        1.329    -1.326     1.416    -1.271     1.401
## S18                   -2.471        0.218    -2.550     0.344    -2.449     0.295
## S19                   -4.018       -1.165    -4.136    -1.068    -4.011    -1.191
## FR                    -3.663       -0.102    -3.549    -0.167    -3.775    -0.048
## SO                    -1.451        1.207    -1.590     1.111    -1.659     1.140
## JR                    -1.785        0.811    -1.844     0.915    -1.851     0.918
## ACTMajor               0.309        0.741     0.303     0.720     0.283     0.736
## ACT                   -0.039        0.492    -0.031     0.462    -0.028     0.475
## ACTMath                0.801        1.317     0.825     1.303     0.839     1.314
```

For completeness we will also consider a beta regression [8,9] approach to the transformed response $Y =$ ObjExam/100 which we show does not have problematic residuals resulting from the natural constraints in the response variable that we saw with linear regression. We consider three beta regression models: one contains all the main effects for both the mean and precision parameter (formula `response   mean | precision`), one contains all main effects for the mean parameter and just the ACTMath term for the precision parameter, and the last model is the previously mentioned reduced model fit with different link functions.

```
library(betareg)
dat_beta <- dat %>% mutate(Y = ObjExam / 100) %>%
  dplyr::select(-ObjExam)
dat_beta$Y[which(dat_beta$Y == 1)] <- 0.9999
dat_beta$Y[which(dat_beta$Y == 0)] <- 0.0001

# standard model
m1 <- betareg(Y ~ Online + Gender + International + F17 + S18 + S19 +
            FR + SO + JR + ACTMajor + ACT + ACTMath|Online + Gender +
```

```
            International + F17 + S18 + S19 +
            FR + SO + JR + ACTMajor + ACT + ACTMath, data = dat_beta)

# smaller model
m2 <- betareg(Y ~ Online + Gender + International + F17 + S18 + S19 +
            FR + SO + JR + ACTMajor + ACT + ACTMath|ACTMath,
            data = dat_beta)

# smaller model (different link functions)
m3 <- betareg(Y ~ Online + Gender + International + F17 + S18 + S19 +
            FR + SO + JR + ACTMajor + ACT + ACTMath|ACTMath,
            data = dat_beta, link = "log", link.phi = "sqrt")
```

AIC suggests that the reduced model with different link functions fits the data the best.

```
AIC(m1)
```

```
## [1] -2710.258
```

```
AIC(m2)
```

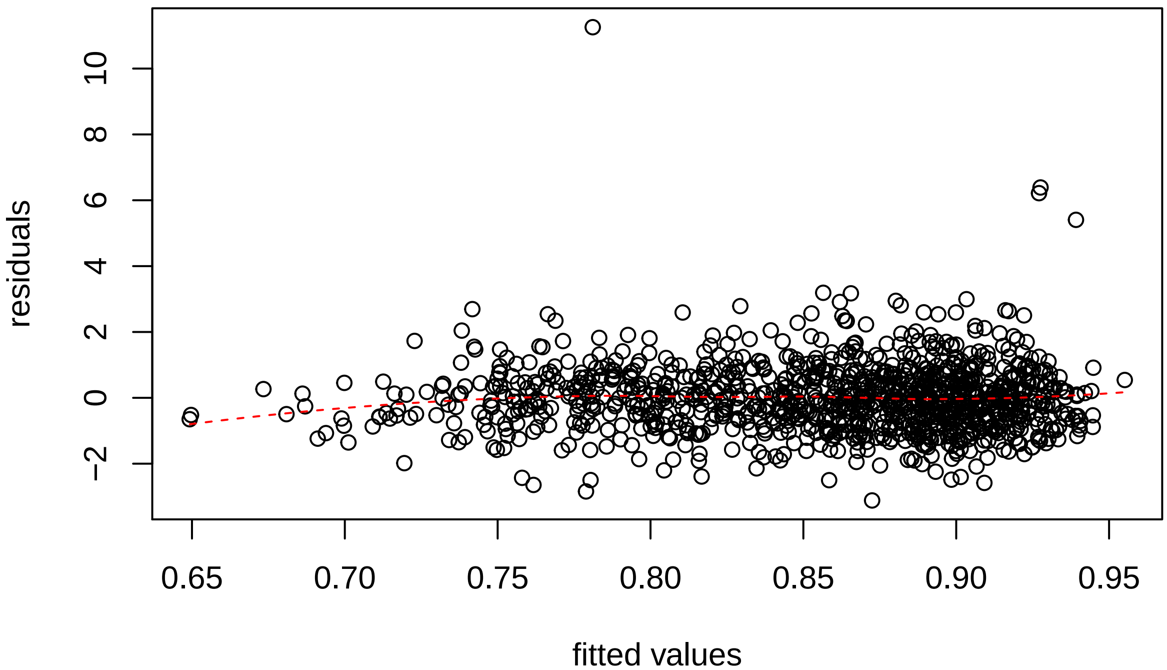
```
## [1] -2720.405
```

```
AIC(m3)
```

```
## [1] -2729.24
```

A residuals vs. fitted values plot with residuals defined in [10] indicates that the beta regression approach has constant variance across the design space.

```
plot(m3$fitted.values, residuals(m3), xlab = "fitted values",
     ylab = "residuals")
fit_resid <- data.frame(resid = residuals(m3),
                        fitted = m3$fitted.value)
lo <- loess(resid ~ fitted, data = fit_resid)
lines(seq(0.3,1,length=500),
      predict(lo, data.frame(fitted = seq(from = 0.3, to = 1, length = 500))),
      lty = 2, col = "red")
```



None of these models suggest that the online learning effect is insignificant at reasonable testing levels.

```
summary(m1)$coef$me[2, ]
```

```
##   Estimate Std. Error    z value   Pr(>|z|)
## 0.05603693 0.04397726 1.27422499 0.20258369
```

```
summary(m2)$coef$me[2, ]
```

```
##   Estimate Std. Error    z value   Pr(>|z|)
## 0.06080280 0.04186502 1.45235333 0.14640338
```

```
summary(m3)$coef$me[2, ]
```

```
##    Estimate  Std. Error     z value    Pr(>|z|)
## 0.006560642 0.005261822 1.246838414 0.212456752
```

We will continue using the multiple linear regression model since beta regression does not conflict with multiple linear regression modeling and bootstrap inference does not depart too far from inferences obtained via the multiple linear regression model.

Conclusions of causal analysis

It appears that the causal effect for online learning is not statistically significant at the $\alpha = 0.05$ significance level as judged by bootstrap percentile interval. The ATE estimator $\widehat{\text{ATE}}_{\text{alt}}$ and the double robust estimator $\widehat{\text{ATE}}_{\text{DR}}$ appear to be stable. The linear regression model estimator for the effect of online learning is consistent with the causal effect estimators.

The alternative stable estimator $\widehat{\text{ATE}}_{\text{alt}}$ of the ATE achieves pseudo-sample balance across important variables.

Additional Analyses

We verify that ACTVerbal is well predicted from both ACT and ACTMath and whose inclusion would have a tiny effect ($< +0.02$) on the OLS estimator of the online learning effect. We will perform these analyses using multiple linear regression models. We note that linear regression modeling assumptions are slightly violated.

We previously showed that inferences from the slightly misspecified multiple linear regression model did not deviate too far from those obtained from the nonparametric bootstrap. We also note that a nonparametric outcome highly adaptive lasso (OHAL) procedure below yielded similar estimates and inferences of the ATE as the procedures considered so far.

```
## ACT verbal predicted from ACT and ACT math
m_V <- lm(ACTVerbal ~ ACT + ACTMath, data = dat)
summary(m_V)
```

```
##
## Call:
## lm(formula = ACTVerbal ~ ACT + ACTMath, data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.8057 -1.8968 -0.1908  1.6751 11.1571
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.03172    0.78632   7.671 3.74e-14 ***
## ACT          3.25672    0.04245  76.711  < 2e-16 ***
## ACTMath     -1.44928    0.03672 -39.465  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.902 on 1102 degrees of freedom
## Multiple R-squared:  0.8687, Adjusted R-squared:  0.8684
## F-statistic:  3645 on 2 and 1102 DF,  p-value: < 2.2e-16
```

```
## add ACT verbal
m_V_full <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + ACTVerbal +
  International + Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat)

## change in effect resulting from including ACTVerbal
coefficients(m_V_full)[2] - effects(dat)[1]
```

```
##     Online
## 0.01300506
```

We see that the OLS estimator of the online learning effect was not very sensitive to omission of any of these covariates. Removal of all the ACT scores increases the OLS estimator by about 0.5. Removal of the college class decreases the OLS estimator by about 0.1.

```
## remove ACT variables
m_noACT <- lm(ObjExam ~ Online + International + Gender + FR + SO +
  JR + F17 + S18 + S19, data = dat)
coefficients(m_noACT)[2] - effects(dat)[1]

## remove class variables
m_noClass <- lm(ObjExam ~ Online + International + Gender + ACTMath +
  ACTMajor + ACT + ACTVerbal, data = dat)
coefficients(m_noClass)[2] - effects(dat)[1]

## remove all variables
m_Online <- lm(ObjExam ~ Online, data = dat)
coefficients(m_Online)[2] - effects(dat)[1]
```

Inclusion of HSGPA in the model on the subsample of 675 students for which it was available increased R^2 from 0.32 to 0.37. It changed the OLS estimator in this subsample by 0.11 (also computed above), from 0.85 to 0.96.

```r
## HSGPA in model
m_HS <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
  FR + SO + JR + F17 + S18 + S19 + HSGPA,
  data = dat_HS)
summary(m_HS)
```

```
##
## Call:
## lm(formula = ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
##     Gender + FR + SO + JR + F17 + S18 + S19 + HSGPA, data = dat_HS)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -39.281  -4.516   1.386   5.406  21.935
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    11.7110     4.8110   2.434  0.01519 *
## Online          0.9564     0.7277   1.314  0.18920
## ACTMath         0.8636     0.1534   5.629 2.69e-08 ***
## ACTMajor        0.3338     0.1552   2.151  0.03185 *
## ACT             0.2697     0.1589   1.698  0.09003 .
## International  -1.9913     1.1991  -1.661  0.09726 .
## Gender1        -1.0849     0.7273  -1.492  0.13625
## FR             -1.4514     1.2886  -1.126  0.26042
## SO              0.3949     0.9446   0.418  0.67601
## JR             -0.2241     0.9534  -0.235  0.81422
## F17             1.2276     0.9613   1.277  0.20202
## S18            -0.1433     0.9390  -0.153  0.87875
## S19            -2.5499     0.9497  -2.685  0.00744 **
## HSGPA           7.8441     1.0347   7.581 1.16e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.582 on 661 degrees of freedom
## Multiple R-squared:  0.3728, Adjusted R-squared:  0.3605
## F-statistic: 30.22 on 13 and 661 DF,  p-value: < 2.2e-16
```

```r
## HSGPA not in model
m_HS_noGPA <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
  FR + SO + JR + F17 + S18 + S19,
  data = dat_HS)
summary(m_HS_noGPA)
```

```
##
## Call:
## lm(formula = ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
##     Gender + FR + SO + JR + F17 + S18 + S19, data = dat_HS)
##
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -39.847  -4.466   1.695   5.961  20.896
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    27.5835     4.5125   6.113 1.67e-09 ***
## Online          0.8475     0.7580   1.118  0.26393
## ACTMath         0.9742     0.1591   6.123 1.57e-09 ***
## ACTMajor        0.6221     0.1567   3.969 8.00e-05 ***
## ACT             0.2679     0.1655   1.619  0.10603
## International  -1.7564     1.2488  -1.406  0.16007
## Gender1        -0.1088     0.7457  -0.146  0.88404
## FR             -2.4624     1.3352  -1.844  0.06559 .
## SO             -0.1584     0.9811  -0.161  0.87178
## JR             -0.7328     0.9908  -0.740  0.45981
## F17             0.6797     0.9986   0.681  0.49636
## S18            -0.2404     0.9781  -0.246  0.80591
## S19            -2.6127     0.9893  -2.641  0.00846 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.94 on 662 degrees of freedom
## Multiple R-squared:  0.3183, Adjusted R-squared:  0.3059
## F-statistic: 25.75 on 12 and 662 DF,  p-value: < 2.2e-16
```

```
## difference in effects
coef(m_HS)[2] - coef(m_HS_noGPA)[2]
```

```
##    Online
## 0.1089424
```

We think the best estimate of ATE should then be increased by roughly +0.1 in the overall sample, assuming that the traits measured by HSGPA have roughly similar effects and similar group differences in the 39% of the sample for which HSGPA was not available. The resulting ATE estimates would then be 0.73, 0.83, and 0.73 for the OLS, IPW, and DR methods respectively, each with 95% CIs of roughly ± 1.2 around the point estimate.

```
effects(dat) + effects_HS(dat_HS)
```

```
##    Online   ATE_alt    ATE_DR
## 0.7342742 0.8319280 0.7313681
```

We explored interaction effects of the treatment with the variables suspected of being relevant to the effectiveness of the online treatment (gender, citizenship, freshman status and ACT math) to see if there were any indications of special effects for those groups. Adding these effects, either individually or all simultaneously, gave no interaction terms significant at the 95% confidence level. Stratifying by the same variables gave insignificantly positive OLS estimators of the online learning effect for both US and non-US citizens, both freshmen and non-freshman, and both the upper and lower halves of the ACT math distribution.

```
## added effect of Online*Gender
m_Online_Gen <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
  Gender + Gender*Online + FR + SO + JR + F17 + S18 + S19,
  data = dat)
anova(m_MLR, m_Online_Gen)
```

```
## Analysis of Variance Table
##
```

```
## Model 1: ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
##     FR + SO + JR + ACTMajor + ACT + ACTMath
## Model 2: ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
##     Gender + Gender * Online + FR + SO + JR + F17 + S18 + S19
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1   1092 74635
## 2   1091 74483  1    152.03 2.2269 0.1359
```

```
coef(m_Online_Gen)[2] + c(1,0) * coef(m_Online_Gen)[14]
```

```
## [1] -0.2582874  1.2712004
```

```
## added effect of Online*International
m_Online_Int <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
  I(International*Online) + Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat)
anova(m_MLR, m_Online_Int)
```

```
## Analysis of Variance Table
##
## Model 1: ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
##     FR + SO + JR + ACTMajor + ACT + ACTMath
## Model 2: ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
##     I(International * Online) + Gender + FR + SO + JR + F17 +
##     S18 + S19
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1   1092 74635
## 2   1091 74584  1    50.608 0.7403 0.3898
```

```
coef(m_Online_Int)[2] + c(1,0) * coef(m_Online_Int)[7]
```

```
## [1] 0.06736862 0.97003234
```

```
## added effect of Online*FR
m_Online_FR <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
  Gender + FR + I(Online*FR) + SO + JR + F17 + S18 + S19,
  data = dat)
anova(m_MLR, m_Online_FR)
```

```
## Analysis of Variance Table
##
## Model 1: ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
##     FR + SO + JR + ACTMajor + ACT + ACTMath
## Model 2: ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
##     Gender + FR + I(Online * FR) + SO + JR + F17 + S18 + S19
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1   1092 74635
## 2   1091 74629  1    5.9889 0.0876 0.7674
```

```
coef(m_Online_FR)[2] + c(0,1) * coef(m_Online_FR)[9]
```

```
## [1] 0.5676914 1.0536288
```

```
## added effect of Online*ACTMath
m_Online_ACTMath <- lm(ObjExam ~ Online + ACTMath + I(Online*ACTMath) + ACTMajor +
  ACT + International + Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat)
anova(m_MLR, m_Online_ACTMath)
```

```
## Analysis of Variance Table
##
## Model 1: ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
##     FR + SO + JR + ACTMajor + ACT + ACTMath
## Model 2: ObjExam ~ Online + ACTMath + I(Online * ACTMath) + ACTMajor +
##     ACT + International + Gender + FR + SO + JR + F17 + S18 +
##     S19
##   Res.Df   RSS Df Sum of Sq     F Pr(>F)
## 1   1092 74635
## 2   1091 74605  1    29.678 0.434 0.5102
```

```r
coef(m_Online_ACTMath)[2] + range(dat$ACTMath) * coef(m_Online_ACTMath)[4]
```

```
## [1] -0.7050538  0.9630548
```

```r
## added effect of all interactions
m_Online_all <- lm(ObjExam ~ Online + ACTMath + I(Online*ACTMath) + ACTMajor +
  ACT + International + I(International*Online) + Gender + Gender*Online + FR +
  I(Online*FR) + SO + JR + F17 + S18 + S19,
  data = dat)
anova(m_MLR, m_Online_all)
```

```
## Analysis of Variance Table
##
## Model 1: ObjExam ~ Online + Gender + International + F17 + S18 + S19 +
##     FR + SO + JR + ACTMajor + ACT + ACTMath
## Model 2: ObjExam ~ Online + ACTMath + I(Online * ACTMath) + ACTMajor +
##     ACT + International + I(International * Online) + Gender +
##     Gender * Online + FR + I(Online * FR) + SO + JR + F17 + S18 +
##     S19
##   Res.Df   RSS Df Sum of Sq      F Pr(>F)
## 1   1092 74635
## 2   1088 74372  4    262.69 0.9607 0.4281
```

The estimated ATE given above are appropriate if the variation in homework is not determined by any properties that differ systematically between the treatment groups except the treatment and the measured covariates. Students' particular commitment to the course is perhaps one of the most obvious variables likely to affect learning outcomes, to differ between groups choosing different versions, and not to show up in more general covariates. That commitment should affect how much effort students put into doing the homework. We verify that the homework scores were less predictable than exam scores from the covariates to which we had access. We see that $R^2 = 0.11$ for homework in the group for which HSGPA could be used, in contrast to 0.37 for objective exam scores (verified above).

```r
m_HW_HS <- lm(Homework ~ Online + ACTMath + ACTMajor + ACT + International +
  Gender + FR + SO + JR + F17 + S18 + S19 + HSGPA,
  data = dat_HS)
summary(m_HW_HS)
```

```
##
## Call:
## lm(formula = Homework ~ Online + ACTMath + ACTMajor + ACT + International +
##     Gender + FR + SO + JR + F17 + S18 + S19 + HSGPA, data = dat_HS)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -88.285  -1.575   2.955   6.738  18.679
##
```

```
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    37.88584    7.22586   5.243 2.13e-07 ***
## Online          1.97409    1.09296   1.806   0.0713 .
## ACTMath         0.22585    0.23044   0.980   0.3274
## ACTMajor        0.21061    0.23307   0.904   0.3665
## ACT             0.03383    0.23860   0.142   0.8873
## International   0.92481    1.80104   0.513   0.6078
## Gender1         2.25645    1.09234   2.066   0.0392 *
## FR              3.69119    1.93536   1.907   0.0569 .
## SO              0.40541    1.41870   0.286   0.7752
## JR              0.01985    1.43202   0.014   0.9889
## F17            -0.83117    1.44378  -0.576   0.5650
## S18            -0.76492    1.41027  -0.542   0.5877
## S19             0.98770    1.42639   0.692   0.4889
## HSGPA          10.04534    1.55407   6.464 1.98e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.89 on 661 degrees of freedom
## Multiple R-squared:  0.1119, Adjusted R-squared:  0.09443
## F-statistic: 6.406 on 13 and 661 DF,  p-value: 1.642e-11
```

Furthermore, homework was much better predicted by HSGPA than by ACT math, while objective exam scores were slightly better predicted by ACT math, in a model where the other ACT scores were dropped, consistent with our intuition that HSGPA and homework might show relatively large effort-dependent contributions compared to ACT math and objective exam scores, respectively.

```
## HSGPA effect higher than ACT math when predicting HW
m_HW_HS_noACT <- lm(Homework ~ Online + ACTMath + International +
  Gender + FR + SO + JR + F17 + S18 + S19 + HSGPA,
  data = dat_HS)
summary(m_HW_HS_noACT)
```

```
##
## Call:
## lm(formula = Homework ~ Online + ACTMath + International + Gender +
##     FR + SO + JR + F17 + S18 + S19 + HSGPA, data = dat_HS)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -89.432  -1.575   3.009   6.749  19.322
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    41.61042    6.03861   6.891 1.29e-11 ***
## Online          1.95951    1.09089   1.796   0.0729 .
## ACTMath         0.30590    0.13589   2.251   0.0247 *
## International   0.74590    1.76365   0.423   0.6725
## Gender1         2.19067    1.08230   2.024   0.0434 *
## FR              3.54411    1.92349   1.843   0.0658 .
## SO              0.19678    1.39884   0.141   0.8882
## JR              0.04595    1.43055   0.032   0.9744
## F17            -0.72160    1.43278  -0.504   0.6147
## S18            -0.77751    1.40856  -0.552   0.5811
```

```
## S19            0.99601    1.42503   0.699   0.4848
## HSGPA         10.40252    1.50461   6.914 1.11e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.88 on 663 degrees of freedom
## Multiple R-squared:  0.1107, Adjusted R-squared:  0.09595
## F-statistic: 7.503 on 11 and 663 DF,  p-value: 3.134e-12
```

```
## model with HSGP prefered when predicting HW
AIC(lm(Homework ~ Online + ACTMath + International +
  Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat_HS)) -
AIC(lm(Homework ~ Online + HSGPA + International +
  Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat_HS))
```

```
## [1] 41.85122
```

```
## ACT math effect slightly higher than HSGPA when predicting Exam
m_exam_HS_noACT <- lm(ObjExam ~ Online + ACTMath + International +
  Gender + FR + SO + JR + F17 + S18 + S19 + HSGPA,
  data = dat_HS)
summary(m_exam_HS_noACT)
```

```
##
## Call:
## lm(formula = ObjExam ~ Online + ACTMath + International + Gender +
##     FR + SO + JR + F17 + S18 + S19 + HSGPA, data = dat_HS)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -39.150  -4.644   1.358   5.506  22.232
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   18.81874    4.04403   4.653 3.94e-06 ***
## Online         0.88747    0.73057   1.215   0.2249
## ACTMath        1.15744    0.09100  12.719  < 2e-16 ***
## International -2.58969    1.18111  -2.193   0.0287 *
## Gender1       -1.09978    0.72481  -1.517   0.1297
## FR            -1.83270    1.28815  -1.423   0.1553
## SO             0.05987    0.93679   0.064   0.9491
## JR            -0.17516    0.95803  -0.183   0.8550
## F17            1.32511    0.95953   1.381   0.1677
## S18           -0.19836    0.94330  -0.210   0.8335
## S19           -2.55117    0.95434  -2.673   0.0077 **
## HSGPA          8.45506    1.00763   8.391 2.91e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.625 on 663 degrees of freedom
## Multiple R-squared:  0.3646, Adjusted R-squared:  0.3541
## F-statistic: 34.59 on 11 and 663 DF,  p-value: < 2.2e-16
```

```r
## model with ACT math prefered when predicting Exam
AIC(lm(ObjExam ~ Online + ACTMath + International +
  Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat_HS)) -
AIC(lm(ObjExam ~ Online + HSGPA + International +
  Gender + FR + SO + JR + F17 + S18 + S19,
  data = dat_HS))
```

```
## [1] -79.23859
```

We now verify that including homeowrk reduced the OLS estimator of the online learning effect from 0.63 to 0.33 in the overall sample for which HSGPA was not used. In the subset for which HSGPA was available and used in the model, including homework reduced the OLS estimator from 0.96 to 0.43.

```r
## full model with homework added
m_HW <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
  FR + SO + JR + F17 + S18 + S19 + Homework,
  data = dat)
coef(m_HW)[2]
```

```
##    Online
## 0.3281311
```

```r
## HSGPA model with HW excluded
m_HS_noHW <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
  FR + SO + JR + F17 + S18 + S19 + HSGPA,
  data = dat_HS)

## HSGPA model with HW included
m_HS_HW <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
  FR + SO + JR + F17 + S18 + S19 + HSGPA + Homework,
  data = dat_HS)
coef(m_HS_noHW)[2]
```

```
##    Online
## 0.9564226
```

```r
coef(m_HS_HW)[2]
```

```
##    Online
## 0.4264326
```

Bootstrapping interaction terms

Earlier, we explored interaction effects of the treatment with the variables suspected of being relevant to the effectiveness of the online treatment (gender, citizenship, freshman status and ACT math) to see if there were any indications of special effects for those groups. In the analysis above we considered p-values for the interaction terms and anova tests. We now consider a 95% percentile interval obtained from a nonparametric bootstrap. All of the 95% percentile intervals for the estimated treatment effects and interaction terms contain zero.

```r
set.seed(13)
B <- 1e4
output <- matrix(0, nrow = B, ncol = 13)
j <- 1

while(j <= B){
```

```
  ind <- sample(1:nrow(dat), replace = TRUE)

  ## added effect of Online*Gender
  m_Online_Gen <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
    Gender + FR + SO + JR + F17 + S18 + S19 + Gender*Online,
    data = dat[ind, ])
  output[j, 1:2] <-  coef(m_Online_Gen)[c(2,14)]

  ## added effect of Online*International
  m_Online_Int <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
    Gender + FR + SO + JR + F17 + S18 + S19 + I(International*Online),
    data = dat[ind, ])
  output[j, 3:4] <-  coef(m_Online_Int)[c(2,14)]

  ## added effect of Online*FR
  m_Online_FR <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
    Gender + FR + SO + JR + F17 + S18 + S19 + I(Online*FR),
    data = dat[ind, ])
  output[j, 5:6] <-  coef(m_Online_FR)[c(2,14)]

  ## added effect of Online*ACTMath
  m_Online_ACTMath <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT +
    International + Gender + FR + SO + JR + F17 + S18 + S19 + I(Online*ACTMath),
    data = dat[ind, ])
  output[j, 7:8] <-  coef(m_Online_ACTMath)[c(2,14)]

  ## added effect of all interactions
  m_Online_all <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International +
    Gender + FR + SO + JR + F17 + S18 + S19 + I(Online*FR) + I(Online*ACTMath) +
      I(International*Online) + Gender*Online, data = dat[ind, ])
  output[j, 9:13] <-  coef(m_Online_all)[c(2,14:17)]
  j <- j + 1
}
```

```
apply(output, 2, function(xx) quantile(xx, probs = c(0.025,0.975)))
```

```
##              [,1]      [,2]       [,3]      [,4]       [,5]      [,6]
## 2.5%  -0.05959127 -3.553128 -0.4752994 -2.848765 -0.5689434 -3.047683
## 97.5%  2.63311043  0.527920  2.4383218  1.062957  1.7493762  3.777721
##             [,7]       [,8]       [,9]     [,10]      [,11]      [,12]
## 2.5%  -11.895460 -0.2154712 -13.718826 -2.806869 -0.1991492 -3.6287280
## 97.5%   7.902685  0.3796041   8.266593  3.967240  0.4756259  0.7401861
##            [,13]
## 2.5%  -3.3004099
## 97.5%  0.8273998
```

Attenuation bias and the F17 semester

```
effects_atten <- function(dat){

  # make numeric variables numeric
  dat$ACT <- as.numeric(as.character(dat$ACT))
  dat$ACTMajor <- as.numeric(as.character(dat$ACTMajor))
  dat$ACTMath <- as.numeric(as.character(dat$ACTMath))
```

```r
## basic data wrangling (will induce non problematic rank deficiency)
# no ACT or HW
dat_small <- dat %>% dplyr::select(Online, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR)
# ACT
dat_ACT <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR) %>%
filter(ACTMajor > 0 & !is.na(ACTMajor) &
       ACTMath > 0 & !is.na(ACTMath))
# HW
dat_HW <- dat %>% dplyr::select(Online, Homework, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR)
# ACT and HW
dat_big <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR,
                          Homework) %>%
filter(ACTMajor > 0 & !is.na(ACTMajor) &
       ACTMath > 0 & !is.na(ACTMath))

## estimate ATE
# no ACT or HW
m_small <- glm(Online ~., data = dat_small, family = "binomial")
trt_small <- dat_small$Online
preds_small <- predict(m_small, type = "response")
# ACT
m_ACT <- glm(Online ~., data = dat_ACT, family = "binomial")
trt_ACT <- dat_ACT$Online
preds_ACT <- predict(m_ACT, type = "response")
# HW
m_HW <- glm(Online ~., data = dat_HW, family = "binomial")
trt_HW <- dat_HW$Online
preds_HW <- predict(m_HW, type = "response")
# ACT and HW
m_big <- glm(Online ~., data = dat_big, family = "binomial")
trt_big <- dat_big$Online
preds_big <- predict(m_big, type = "response")



## estimate alternative stabilized ATE
# no ACT or HW
scale_alt_trt_small <- 1 / sum(trt_small / preds_small)
scale_alt_notrt_small <- 1 / sum((1 - trt_small)/(1 - preds_small))
ATE_alt_small <- scale_alt_trt_small * sum((1/preds_small) * trt_small * dat$ObjExam) -
  scale_alt_notrt_small * sum((1-trt_small)/(1-preds_small)* dat$ObjExam)
# ACT
scale_alt_trt_ACT <- 1 / sum(trt_ACT / preds_ACT)
scale_alt_notrt_ACT <- 1 / sum((1 - trt_ACT)/(1 - preds_ACT))
ATE_alt_ACT <- scale_alt_trt_ACT * sum((1/preds_ACT) * trt_ACT * dat$ObjExam) -
  scale_alt_notrt_ACT * sum((1-trt_ACT)/(1-preds_ACT)* dat$ObjExam)
# HW
scale_alt_trt_HW <- 1 / sum(trt_HW / preds_HW)
scale_alt_notrt_HW <- 1 / sum((1 - trt_HW)/(1 - preds_HW))
```

```r
ATE_alt_HW <- scale_alt_trt_HW * sum((1/preds_HW) * trt_HW * dat$ObjExam) -
  scale_alt_notrt_HW * sum((1-trt_HW)/(1-preds_HW)* dat$ObjExam)
# ACT and HW
scale_alt_trt_big <- 1 / sum(trt_big / preds_big)
scale_alt_notrt_big <- 1 / sum((1 - trt_big)/(1 - preds_big))
ATE_alt_big <- scale_alt_trt_big * sum((1/preds_big) * trt_big * dat$ObjExam) -
  scale_alt_notrt_big * sum((1-trt_big)/(1-preds_big)* dat$ObjExam)


## estimate DR version ATE (will induce non problematic rank deficiency)
# no ACT or HW
m_trt_small <- lm(ObjExam ~ International + Gender +
              FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_small == 1, ])
Y_trt_small <- predict(m_trt_small, newdata = dat)
m_notrt_small <- lm(ObjExam ~ International + Gender +
              FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_small == 0, ])
Y_notrt_small <- predict(m_notrt_small, newdata = dat)
ATE_DR_small <- mean( (dat$ObjExam * trt_small - (trt_small - preds_small) * Y_trt_small) / preds_smal
  (dat$ObjExam * (1 - trt_small) + (trt_small - preds_small)*Y_notrt_small) / (1 - preds_small))
# ACT
m_trt_ACT <- lm(ObjExam ~ International + Gender + ACTMath +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_ACT == 1, ])
Y_trt_ACT <- predict(m_trt_ACT, newdata = dat)
m_notrt_ACT <- lm(ObjExam ~ International + Gender + ACTMath +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_ACT == 0, ])
Y_notrt_ACT <- predict(m_notrt_ACT, newdata = dat)
ATE_DR_ACT <- mean( (dat$ObjExam * trt_ACT - (trt_ACT - preds_ACT) * Y_trt_ACT) / preds_ACT -
  (dat$ObjExam * (1 - trt_ACT) + (trt_ACT - preds_ACT)*Y_notrt_ACT) / (1 - preds_ACT))
# HW
m_trt_HW <- lm(ObjExam ~ International + Gender + Homework +
                 FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_HW == 1, ])
Y_trt_HW <- predict(m_trt_HW, newdata = dat)
m_notrt_HW <- lm(ObjExam ~ International + Gender + Homework +
                   FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_HW == 0, ])
Y_notrt_HW <- predict(m_notrt_HW, newdata = dat)
ATE_DR_HW <- mean( (dat$ObjExam * trt_HW - (trt_HW - preds_HW) * Y_trt_HW) / preds_HW -
  (dat$ObjExam * (1 - trt_HW) + (trt_HW - preds_HW)*Y_notrt_HW) / (1 - preds_HW))
# ACT and HW
m_trt_big <- lm(ObjExam ~ International + Gender + ACTMath + Homework +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_big == 1, ])
Y_trt_big <- predict(m_trt_big, newdata = dat)
m_notrt_big <- lm(ObjExam ~ International + Gender + ACTMath + Homework +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_big == 0, ])
Y_notrt_big <- predict(m_notrt_big, newdata = dat)
ATE_DR_big <- mean( (dat$ObjExam * trt_big - (trt_big - preds_big) * Y_trt_big) / preds_big -
```

```
    (dat$ObjExam * (1 - trt_big) + (trt_big - preds_big)*Y_notrt_big) / (1 - preds_big))

  ## regression model (will induce non problematic rank deficiency)
  # no ACT or HW
  m_small <- lm(ObjExam ~ Online + International + Gender +
            FR + SO + JR + F17 + S18 + S19,
          data = dat)
  # ACT
  m_ACT <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
            FR + SO + JR + F17 + S18 + S19,
          data = dat)
  # HW
  m_HW <- lm(ObjExam ~ Online + Homework + International + Gender +
            FR + SO + JR + F17 + S18 + S19,
          data = dat)
  # ACT and HW
  m_big <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
            Homework + FR + SO + JR + F17 + S18 + S19,
          data = dat)

  # beta regression
  #dat_Y <- dat %>% mutate(Y = ObjExam/100)
  #dat_Y$Y[which(dat$Y == 1)] <- 0.9999
  #dat_Y$Y[which(dat$Y == 0)] <- 0.0001
  #m_beta <- betareg(Y ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
  #           FR + SO + JR + F17 + S18 + S19 | ACTMath + S18 + S19 + Fa19 + FR + SO + JR,
  #           data = dat, link = "log", link.phi = "sqrt"),

  #c("ATE" = ATE, "ATE_alt" = ATE_alt, "ATE_DR" = ATE_DR,
  #  "Online" = as.numeric(m$coef[2]))
  #c("Online" = as.numeric(m$coef[2]), "Online_Beta" = summary(m_beta)$coe$me[2,1],
  #  "ATE_alt" = ATE_alt, "ATE_DR" = ATE_DR)
  out <- list(out_small = c("Online" = as.numeric(m_small$coef[2]),
                            "ATE_alt" = ATE_alt_small,
                            "ATE_DR" = ATE_DR_small),
              out_ACT = c("Online" = as.numeric(m_ACT$coef[2]),
                            "ATE_alt" = ATE_alt_ACT,
                            "ATE_DR" = ATE_DR_ACT),
              out_HW = c("Online" = as.numeric(m_HW$coef[2]),
                            "ATE_alt" = ATE_alt_HW,
                            "ATE_DR" = ATE_DR_HW),
              out_big = c("Online" = as.numeric(m_big$coef[2]),
                            "ATE_alt" = ATE_alt_big,
                            "ATE_DR" = ATE_DR_big)
  )
  out

}
```

We now investigate attenuation bias for the whole data set, the F17 semester only, and the data set with the F17 semester removed.

```
## F17 semester
effects_atten(dat %>% filter(F17 == 1))
```

```
## $out_small
##    Online   ATE_alt    ATE_DR
## -2.410252 -1.800868 -1.761319
##
## $out_ACT
##    Online   ATE_alt    ATE_DR
## -1.743206 -1.140049 -1.550791
##
## $out_HW
##    Online   ATE_alt    ATE_DR
## -1.641146 -1.148427 -1.182634
##
## $out_big
##     Online    ATE_alt     ATE_DR
## -1.2059902 -0.7525027 -1.3313546
```

```
## F17 semester removed
effects_atten(dat %>% filter(F17 != 1))
```

```
## $out_small
##   Online  ATE_alt   ATE_DR
## 2.452081 2.568785 2.438197
##
## $out_ACT
##   Online  ATE_alt   ATE_DR
## 1.462210 1.505607 1.421703
##
## $out_HW
##   Online  ATE_alt   ATE_DR
## 1.237880 1.386361 1.136449
##
## $out_big
##    Online   ATE_alt    ATE_DR
## 0.5102844 0.6022103 0.3571616
```

```
## all semesters combined
effects_atten(dat)
```

```
## $out_small
##   Online  ATE_alt   ATE_DR
## 1.170456 1.360552 1.154516
##
## $out_ACT
##    Online   ATE_alt    ATE_DR
## 0.6253318 0.7467678 0.5839065
##
## $out_HW
##    Online   ATE_alt    ATE_DR
## 0.7428589 0.9396642 0.7112390
##
## $out_big
##    Online   ATE_alt    ATE_DR
## 0.3281311 0.4633829 0.2513331
```

We now investigate attenuation bias for the subset of observations for which HSGPA was recorded.

```
effects_atten_HS <- function(dat){

  # make numeric variables numeric
  dat$ACT <- as.numeric(as.character(dat$ACT))
  dat$ACTMajor <- as.numeric(as.character(dat$ACTMajor))
  dat$ACTMath <- as.numeric(as.character(dat$ACTMath))

  ## basic data wrangling (will induce non problematic rank deficiency)
  # no ACT or HW
  dat_small <- dat %>% dplyr::select(Online, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR, HSGPA)
  # ACT
  dat_ACT <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR, HSGPA) %>%
  filter(ACTMajor > 0 & !is.na(ACTMajor) &
         ACTMath > 0 & !is.na(ACTMath))
  # HW
  dat_HW <- dat %>% dplyr::select(Online, Homework, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR, HSGPA)
  # ACT and HW
  dat_big <- dat %>% dplyr::select(Online, ACTMath, ACTMajor, ACT, Gender,
                          International, F17, S18, S19, Fa19, FR, SO, JR,
                          Homework, HSGPA) %>%
  filter(ACTMajor > 0 & !is.na(ACTMajor) &
         ACTMath > 0 & !is.na(ACTMath))

  ## estimate ATE
  # no ACT or HW
  m_small <- glm(Online ~., data = dat_small, family = "binomial")
  trt_small <- dat_small$Online
  preds_small <- predict(m_small, type = "response")
  # ACT
  m_ACT <- glm(Online ~., data = dat_ACT, family = "binomial")
  trt_ACT <- dat_ACT$Online
  preds_ACT <- predict(m_ACT, type = "response")
  # HW
  m_HW <- glm(Online ~., data = dat_HW, family = "binomial")
  trt_HW <- dat_HW$Online
  preds_HW <- predict(m_HW, type = "response")
  # ACT and HW
  m_big <- glm(Online ~., data = dat_big, family = "binomial")
  trt_big <- dat_big$Online
  preds_big <- predict(m_big, type = "response")



  ## estimate alternative stabilized ATE
  # no ACT or HW
  scale_alt_trt_small <- 1 / sum(trt_small / preds_small)
  scale_alt_notrt_small <- 1 / sum((1 - trt_small)/(1 - preds_small))
  ATE_alt_small <- scale_alt_trt_small * sum((1/preds_small) * trt_small * dat$ObjExam) -
    scale_alt_notrt_small * sum((1-trt_small)/(1-preds_small)* dat$ObjExam)
  # ACT
```

```
scale_alt_trt_ACT <- 1 / sum(trt_ACT / preds_ACT)
scale_alt_notrt_ACT <- 1 / sum((1 - trt_ACT)/(1 - preds_ACT))
ATE_alt_ACT <- scale_alt_trt_ACT * sum((1/preds_ACT) * trt_ACT * dat$ObjExam) -
  scale_alt_notrt_ACT * sum((1-trt_ACT)/(1-preds_ACT)* dat$ObjExam)
# HW
scale_alt_trt_HW <- 1 / sum(trt_HW / preds_HW)
scale_alt_notrt_HW <- 1 / sum((1 - trt_HW)/(1 - preds_HW))
ATE_alt_HW <- scale_alt_trt_HW * sum((1/preds_HW) * trt_HW * dat$ObjExam) -
  scale_alt_notrt_HW * sum((1-trt_HW)/(1-preds_HW)* dat$ObjExam)
# ACT and HW
scale_alt_trt_big <- 1 / sum(trt_big / preds_big)
scale_alt_notrt_big <- 1 / sum((1 - trt_big)/(1 - preds_big))
ATE_alt_big <- scale_alt_trt_big * sum((1/preds_big) * trt_big * dat$ObjExam) -
  scale_alt_notrt_big * sum((1-trt_big)/(1-preds_big)* dat$ObjExam)


## estimate DR version ATE (will induce non problematic rank deficiency)
# no ACT or HW
m_trt_small <- lm(ObjExam ~ International + Gender +
              FR + SO + JR + F17 + S18 + S19 + HSGPA,
            data = dat[trt_small == 1, ])
Y_trt_small <- predict(m_trt_small, newdata = dat)
m_notrt_small <- lm(ObjExam ~ International + Gender +
              FR + SO + JR + F17 + S18 + S19 + HSGPA,
            data = dat[trt_small == 0, ])
Y_notrt_small <- predict(m_notrt_small, newdata = dat)
ATE_DR_small <- mean( (dat$ObjExam * trt_small - (trt_small - preds_small) * Y_trt_small) / preds_smal
  (dat$ObjExam * (1 - trt_small) + (trt_small - preds_small)*Y_notrt_small) / (1 - preds_small))
# ACT
m_trt_ACT <- lm(ObjExam ~ International + Gender + ACTMath + HSGPA +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_ACT == 1, ])
Y_trt_ACT <- predict(m_trt_ACT, newdata = dat)
m_notrt_ACT <- lm(ObjExam ~ International + Gender + ACTMath + HSGPA +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19,
            data = dat[trt_ACT == 0, ])
Y_notrt_ACT <- predict(m_notrt_ACT, newdata = dat)
ATE_DR_ACT <- mean( (dat$ObjExam * trt_ACT - (trt_ACT - preds_ACT) * Y_trt_ACT) / preds_ACT -
  (dat$ObjExam * (1 - trt_ACT) + (trt_ACT - preds_ACT)*Y_notrt_ACT) / (1 - preds_ACT))
# HW
m_trt_HW <- lm(ObjExam ~ International + Gender + Homework +
                 FR + SO + JR + F17 + S18 + S19 + HSGPA,
            data = dat[trt_HW == 1, ])
Y_trt_HW <- predict(m_trt_HW, newdata = dat)
m_notrt_HW <- lm(ObjExam ~ International + Gender + Homework +
                  FR + SO + JR + F17 + S18 + S19 + HSGPA,
            data = dat[trt_HW == 0, ])
Y_notrt_HW <- predict(m_notrt_HW, newdata = dat)
ATE_DR_HW <- mean( (dat$ObjExam * trt_HW - (trt_HW - preds_HW) * Y_trt_HW) / preds_HW -
  (dat$ObjExam * (1 - trt_HW) + (trt_HW - preds_HW)*Y_notrt_HW) / (1 - preds_HW))
# ACT and HW
m_trt_big <- lm(ObjExam ~ International + Gender + ACTMath + Homework +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19 + HSGPA,
```

```
            data = dat[trt_big == 1, ])
Y_trt_big <- predict(m_trt_big, newdata = dat)
m_notrt_big <- lm(ObjExam ~ International + Gender + ACTMath + Homework +
                  ACTMajor + ACT + FR + SO + JR + F17 + S18 + S19 + HSGPA,
            data = dat[trt_big == 0, ])
Y_notrt_big <- predict(m_notrt_big, newdata = dat)
ATE_DR_big <- mean( (dat$ObjExam * trt_big - (trt_big - preds_big) * Y_trt_big) / preds_big -
  (dat$ObjExam * (1 - trt_big) + (trt_big - preds_big)*Y_notrt_big) / (1 - preds_big))

## regression model (will induce non problematic rank deficiency)
# no ACT or HW
m_small <- lm(ObjExam ~ Online + International + Gender +
          FR + SO + JR + F17 + S18 + S19 + HSGPA,
        data = dat)
# ACT
m_ACT <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
          FR + SO + JR + F17 + S18 + S19 + HSGPA,
        data = dat)
# HW
m_HW <- lm(ObjExam ~ Online + Homework + International + Gender +
          FR + SO + JR + F17 + S18 + S19 + HSGPA,
        data = dat)
# ACT and HW
m_big <- lm(ObjExam ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
          Homework + FR + SO + JR + F17 + S18 + S19 + HSGPA,
        data = dat)

# beta regression
#dat_Y <- dat %>% mutate(Y = ObjExam/100)
#dat_Y$Y[which(dat$Y == 1)] <- 0.9999
#dat_Y$Y[which(dat$Y == 0)] <- 0.0001
#m_beta <- betareg(Y ~ Online + ACTMath + ACTMajor + ACT + International + Gender +
#          FR + SO + JR + F17 + S18 + S19 | ACTMath + S18 + S19 + Fa19 + FR + SO + JR,
#          data = dat, link = "log", link.phi = "sqrt"),

#c("ATE" = ATE, "ATE_alt" = ATE_alt, "ATE_DR" = ATE_DR,
#  "Online" = as.numeric(m$coef[2]))
#c("Online" = as.numeric(m$coef[2]), "Online_Beta" = summary(m_beta)$coe$me[2,1],
#  "ATE_alt" = ATE_alt, "ATE_DR" = ATE_DR)
out <- list(out_small = c("Online" = as.numeric(m_small$coef[2]),
                          "ATE_alt" = ATE_alt_small,
                          "ATE_DR" = ATE_DR_small),
            out_ACT = c("Online" = as.numeric(m_ACT$coef[2]),
                          "ATE_alt" = ATE_alt_ACT,
                          "ATE_DR" = ATE_DR_ACT),
            out_HW = c("Online" = as.numeric(m_HW$coef[2]),
                          "ATE_alt" = ATE_alt_HW,
                          "ATE_DR" = ATE_DR_HW),
            out_big = c("Online" = as.numeric(m_big$coef[2]),
                          "ATE_alt" = ATE_alt_big,
                          "ATE_DR" = ATE_DR_big)
)
out
```

```
}
```

We now investigate attenuation bias as we did before.

```
## F17 semester
effects_atten_HS(dat_HS %>% filter(F17 == 1))
```

```
## $out_small
##    Online   ATE_alt    ATE_DR
## -1.921619 -1.156743 -1.082125
##
## $out_ACT
##     Online    ATE_alt     ATE_DR
## -1.2088539 -0.3553720 -0.6401177
##
## $out_HW
##     Online    ATE_alt     ATE_DR
## -1.4968814 -0.8271209 -0.9525560
##
## $out_big
##     Online    ATE_alt     ATE_DR
## -0.9010391 -0.1384423 -0.7451444
```

```
## F17 semester removed
effects_atten_HS(dat_HS %>% filter(F17 != 1))
```

```
## $out_small
##   Online  ATE_alt   ATE_DR
## 2.684789 2.920453 2.769526
##
## $out_ACT
##   Online  ATE_alt   ATE_DR
## 1.634437 1.619028 1.544258
##
## $out_HW
##   Online  ATE_alt   ATE_DR
## 1.278256 1.479142 1.184117
##
## $out_big
##    Online   ATE_alt    ATE_DR
## 0.4489123 0.3976356 0.2859509
```

```
## all semesters combined
effects_atten_HS(dat_HS)
```

```
## $out_small
##   Online  ATE_alt   ATE_DR
## 1.550229 1.738722 1.617449
##
## $out_ACT
##    Online   ATE_alt    ATE_DR
## 0.9564226 0.9274249 0.8693833
##
## $out_HW
##    Online   ATE_alt    ATE_DR
## 0.9156962 1.0508543 0.9062267
```

```
##
## $out_big
##    Online   ATE_alt    ATE_DR
## 0.4264326 0.3495311 0.2799441
```

The differences between the attenuation biases for the regular data set and the smaller data set with HSGPA are computed below.

```
## F17 semester
a <- effects_atten_HS(dat_HS %>% filter(F17 == 1))
b <- effects_atten(dat %>% filter(F17 == 1))
for(j in 1:length(a)){
  print(a[[j]] - b[[j]])
}
```

```
##    Online   ATE_alt    ATE_DR
## 0.4886326 0.6441244 0.6791943
##    Online   ATE_alt    ATE_DR
## 0.5343521 0.7846768 0.9106734
##    Online   ATE_alt    ATE_DR
## 0.1442644 0.3213060 0.2300780
##    Online   ATE_alt    ATE_DR
## 0.3049511 0.6140604 0.5862101
```

```
## F17 semester removed
a <- effects_atten_HS(dat_HS %>% filter(F17 != 1))
b <- effects_atten(dat %>% filter(F17 != 1))
for(j in 1:length(a)){
  print(a[[j]] - b[[j]])
}
```

```
##    Online   ATE_alt    ATE_DR
## 0.2327089 0.3516685 0.3313290
##    Online   ATE_alt    ATE_DR
## 0.1722268 0.1134208 0.1225549
##     Online     ATE_alt      ATE_DR
## 0.04037631 0.09278096 0.04766855
##       Online      ATE_alt       ATE_DR
## -0.06137205 -0.20457466 -0.07121070
```

```
## all semesters combined
a <- effects_atten_HS(dat_HS)
b <- effects_atten(dat)
for(j in 1:length(a)){
  print(a[[j]] - b[[j]])
}
```

```
##    Online   ATE_alt    ATE_DR
## 0.3797733 0.3781702 0.4629331
##    Online   ATE_alt    ATE_DR
## 0.3310908 0.1806571 0.2854768
##    Online   ATE_alt    ATE_DR
## 0.1728374 0.1111900 0.1949877
##       Online      ATE_alt       ATE_DR
##  0.09830147 -0.11385182  0.02861103
```

Outcome highly adaptive lasso

We now perform the outcome highly adaptive lasso [6]. We run the code from the supplement to [6].

```
## steps in HAL paper
source("~/research/online/biom13121-sup-0001-code/ohal_functions.R")
library(SuperLearner)
library(glmnet)
library(gam)
library(tmle)
library(drtmle)
library(hal9001)
head(dat_small)
```

```
##   Online ACTMath ACTMajor  ACT Gender International F17 S18 S19 Fa19 FR SO JR
## 1      1    31.0     31.5 30.0      0             0   1   0   0    0  0  0  1
## 2      1    35.6     31.5 33.2      0             0   1   0   0    0  0  1  0
## 3      1    35.6     33.5 33.9      0             1   1   0   0    0  0  0  0
## 4      0    36.0     33.5 35.0      0             0   1   0   0    0  0  1  0
## 5      1    35.0     29.5 32.0      0             0   1   0   0    0  0  0  1
## 6      0    35.0     31.5 30.0      0             1   1   0   0    0  1  0  0
```

```
dat_small$Gender <- as.numeric(dat_small$Gender)
Y <- dat$ObjExam
mse <- function(preds, y) {
    mean((preds - y)^2)
}
```

We first call `ohal_nuisance` to get nuisance estimates.

```
# call ohal_nuisance to get nuisance estimates
system.time(fit <- ohal_nuisance(W = dat_small[, -1],
                     A = dat_small$Online,
                     Y = Y, V = 5,
  outcome_family = "gaussian",
  lambda_seq = exp(seq(-1, -13, length=100))))
```

```
##    user  system elapsed
## 403.252   3.802 407.427
```

```
system.time(drtmle_ohal_fit <- drtmle(W = dat_small[, -1],
  A = dat_small$Online,  Y = Y,
  a_0 = c(0,1), tolg = 1e-4,
  Qn = list(fit$Q0W, fit$Q1W),
  # Qsteps = 1, # uncomment for one-step targeting
  gn = list(1 - fit$G1W_OHAL, fit$G1W_OHAL),
  SL_gr = "SL.hal9001",
  guard = "g"))
```

```
##     user   system  elapsed
## 1303.192    5.316 1322.191
```

We then get report the drtmle (doubly robust targeted minimum loss-based estimators) for the ATE with the outcome highly adaptive lasso.

```
# get drtmle for OHAL ATE
drtmle_ohal_ci <- ci(drtmle_ohal_fit, contrast = c(-1, 1))
drtmle_ohal_ci
```

```
## $drtmle
##                     est    cil  ciu
## E[Y(1)]-E[Y(0)] 0.627 -0.376 1.63
```

We now report the drtmle with cross-validated outcome highly adaptive lasso.

```
# call drtmle with cross-validated OHAL
system.time(drtmle_ohal_fit_cv <- drtmle(W = dat_small[, -1],
  A = dat_small$Online,  Y = Y,
  a_0 = c(0, 1), tolg = 1e-4,
  Qn = list(fit$cv_Q0W, fit$cv_Q1W),
  # Qsteps = 1, # uncomment for one-step targeting
  gn = list(1 - fit$cv_G1W_OHAL, fit$cv_G1W_OHAL),
  cvFolds = fit$fold_vec,
  SL_gr = "SL.hal9001",
  guard = "g", returnModels = FALSE, verbose = TRUE))
```

```
## Mean of IC       = 0 0 0.01081507 0.000505884 0
## Mean of IC       = 5.52e-08 0 0.003490272 3.3212e-06 0
## Mean of IC       = 6e-10 3.3e-09 0.001110561 7.756e-07 0
```

```
##     user   system  elapsed
## 2542.549    7.962 2568.224
```

```
# get drtmle for OHAL ATE
tmp <- ci(drtmle_ohal_fit_cv, contrast = c(-1, 1), est = "aiptw_c")

# get a confidence interval about drtmle + ohal using cross-validated standard errors
cv_se <- (tmp$aiptw_c[3] - tmp$aiptw_c[2]) / 2 / 1.96
drtmle_ohal_ci_cv_se <- c(drtmle_ohal_fit$drtmle$est[2] - drtmle_ohal_fit$drtmle$est[1],
                          drtmle_ohal_fit$drtmle$est[2] - drtmle_ohal_fit$drtmle$est[1] +
                          c(-1.96, 1.96) * cv_se)
drtmle_ohal_ci_cv_se
```

```
## [1]  0.6270819 -0.4954712  1.7496349
```

References

[1] Austin, PC, Stuart, EA (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Statistics in medicine, 34(28), 3661–3679.

[2] Cole SR, Hernan, MA (2008). Constructing inverse probability weights for marginal structural models. American Journal of Epidemiology, 168(6), 656-–664.

[3] Lunceford JK, Davidian M (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. Statistics in Medicine, 23(19), 2937-–2960.

[4] Rosenbaum PR (1998). Propensity score. In Encyclopedia of Biostatistics, Volume 5, Armitage P, Colton T (eds). Wiley: New York, 3551–3555.

[5] Robins, JM, Rotnitzky A, Zhao LP (1994). Estimation of regression coefficients when some regressors are not always observed, Journal of the American Statistical Association, 89, 846–866.

[6] Ju, C, Benkeser, D, and van Der Laan, MJ (2020). Robust inference on the average treatment effect using the outcome highly adaptive lasso. Biometrics, 76(1), 109–118.

[7] Austin, PC (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res, 46(3): 399–424.

[8] Cribari-Neto, F, and Zeileis, A (2010). Beta Regression in R. Journal of Statistical Software, 34(2), 1–24. doi: 10.18637/jss.v034.i02

[9] Zeileis, A, Cribari-Neto, F, Gruen, B, Kosmidis, I, Simas, AB, Rocha, AV, & Zeileis, MA (2016). Package 'betareg'. R package, 3(2).

[10] Espinheira, PL, Ferrari, SLP, and Cribari-Neto, F (2008). On Beta Regression Residuals. Journal of Applied Statistics, 35(4), 407–419.