

Assignment 8.b for **STATGR6103**

submitted to Professor Andrew Gelman

Advait Rajagopal

31 October 2016

1 Question 1

A study was performed to measure the effect of a new educational television program, “The Electric Company”, on children’s reading abilities. An experiment was performed on children in grades 1 - 4 in two cities (Fresno and Youngstown). For each city and grade, the experimenters selected a small number of schools (10 - 20) and, within each school they selected the two poorest reading classes of that grade. For each pair, one of these classes was randomly assigned to continue with their regular reading course and the other was assigned to view the TV program. Once these treatments had been assigned, the teacher for each class assigned to the Electric Company treatment chooses to either replace or supplement the regular reading program with the Electric Company. (That is, the classes in the treatment group all get the Electric Company, but some get it instead of the regular reading program and others get it in addition.)

The students in every class are given a test (before treatment), with 2nd, 3rd, and 4th graders receiving the same exam, but 1st graders getting a slightly different version. After the treatments, the students take another test, equivalent to that which the 2nd, 3rd, and 4th graders received on the pre-test. The file <http://www.stat.columbia.edu/~gelman/bda.course/electric.txt> contains the average pre-test and post-test scores for the students in each of the treatment and control classes (cities are labeled 1=Fresno, 2=Youngstown).

1.1 Part A

Using Stan, estimate the effect of viewing the TV program. (For the purpose of this part of the assignment, ignore the distinction between replacement and supplement.) Give uncertainty intervals for all point estimates, and be sensitive to the possibility that effects may differ between cities and between grades.

I know that a number of grades (1,2,3,4) have been considered and the schools have been selected from two cities so at the outset we have some information about the hierarchical structure in the model and the need to capture this in any Stan model becomes immediately apparent. I start by plotting some histograms as an exploratory exercise to find out if there are variations in the treatment and control condition across grades. Figure 1 shows that there appears to be some differences in the post-test scores for the four grades considered and more importantly there is variation between grades in the post-test scores.

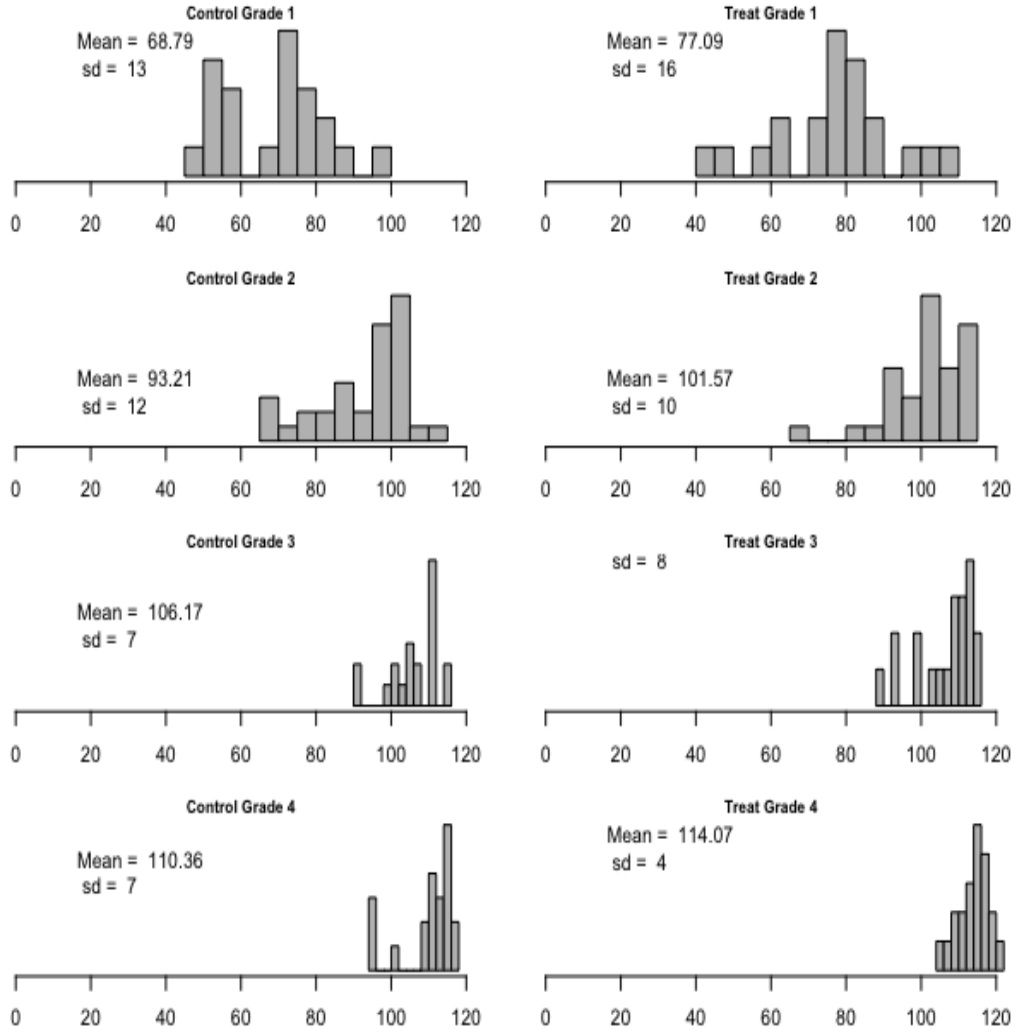


Figure 1: Histogram of post - test scores for control and treatment conditions varies across grades.

Next I have some reason to believe that pre-test score will be a good predictor of post-test scores. While the relationship between pre and post scores is not a causal one, but it is important to factor in pre post improvement regardless of the presence or absence of a treatment. Therefore I fit a linear regression on with post test scores as a function of pre test scores and the presence of the treatment with some error, however I control for pre test scores so that the treatment indicator now represents the average treatment effect controlled for pre test score.

$$y_i = \alpha + \theta T_i + \beta x_i + \epsilon_i$$

Figure 2 shows the relationship between the pre and post scores for all grades in the study. As we suspect the grades 2, 3 and 4 show a similar improving trend after the treatment as the pre and post test were the same. Since grade 1 was given a different test, we observe different pre test scores. However even grade 1 improves on the post test scoring. as is evidenced by the positive sloping regression lines.

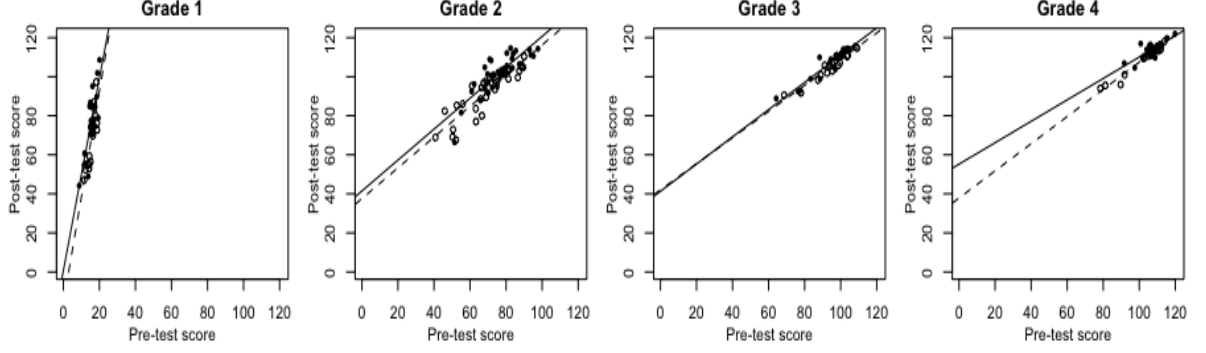


Figure 2: Pre and post test scores for each grade in the experiment in each panel. Treated and control classes are shown by dots and circles respectively. The solid and dotted lines in each panel show the treatment and control conditions respectively.

Some trends are immediately obvious. As mentioned earlier, there seems to be an effect of treatment (The Electric Company TV show) on students in all grades but the effect is clearly different across grades. For example in grades 1 and 2, the treatment line is above the control line showing a positive impact of treatment. In grade 4, for incremental pre test scores, the effect of the treatment becomes less as the two regression lines appear to converge at higher pre test scores.

Now I consider a full hierarchical regression accounting for state level and grade level differences in treatment effects. Owing to the difference in the type of test taken by grade 1 and the other grades I believe incorporating grade level and city level variation in pre test scores is also important. The full exposition of the model is given as follows;

$$y_i \sim \mathbf{N}(\alpha + \theta_{cg}T_i + \beta_{cg}x_i, \sigma^2)$$

$$\beta_{cg} \sim \mathbf{N}(\mu_\beta, \tau_\beta^2)$$

$$\theta_{cg} \sim \begin{cases} \mathbf{N}(\mu_{\theta F}, \tau_{\theta F}^2), & \text{if } c = F \\ \mathbf{N}(\mu_{\theta Y}, \tau_{\theta Y}^2), & \text{if } c = Y \end{cases}$$

Where y_i is the post test score. T_i and x_i are indicators for treatment and the value of the pre test score respectively. α is the intercept, β_{cg} is the coefficient that capture grade level effects of pre test score on the post test score and is given a normal prior distribution with mean μ_β and variance parameter τ_β^2 . The coefficient θ_{cg} is given a normal prior distribution that is city specific. Therefore treatments in city F (Fresno) come from a common distribution with a mean $\mu_{\theta F}$ and a variance $\tau_{\theta F}^2$. Similarly treatments in city Y (Youngstown) come from a distribution with mean $\mu_{\theta Y}$ and a variance $\tau_{\theta Y}^2$. I also use weakly informative prior distributions on the hyperparameters $\mu_{\theta F}$, $\tau_{\theta F}$, $\mu_{\theta Y}$ and $\tau_{\theta Y}$. σ is given a noninformative prior distribution.

Table 1 shows the posterior estimates of the hyperparameters and their 95% interval. The chains all converge satisfactorily with R-hats between 1 and 1.05.

Table 1: Posterior Means and Standard Deviations of Hyperparameters

	mean	sd	2.5%	25%	50%	75%	97.5%
$\mu_{\theta F}$	3.87	1.73	0.55	2.78	3.85	4.95	7.39
$\tau_{\theta F}$	2.26	1.44	0.22	1.16	2.07	3.11	5.55
$\mu_{\theta Y}$	4.10	2.25	-1.08	2.72	4.19	5.58	8.39
$\tau_{\theta Y}$	3.84	2.15	0.58	2.25	3.57	4.98	9.03
μ_{β}	1.38	0.44	0.51	1.16	1.38	1.68	2.29
τ_{β}	1.14	0.41	0.65	0.92	1.10	1.37	2.18
σ	6.03	0.32	5.45	5.82	6.01	6.23	6.69

Figure 3 shows the city wise variation in treatment effects over grades. There does not seem to be too substantial a difference for the same grade across cities. However there is inter city and inter grade variation which can be seen in the figure.

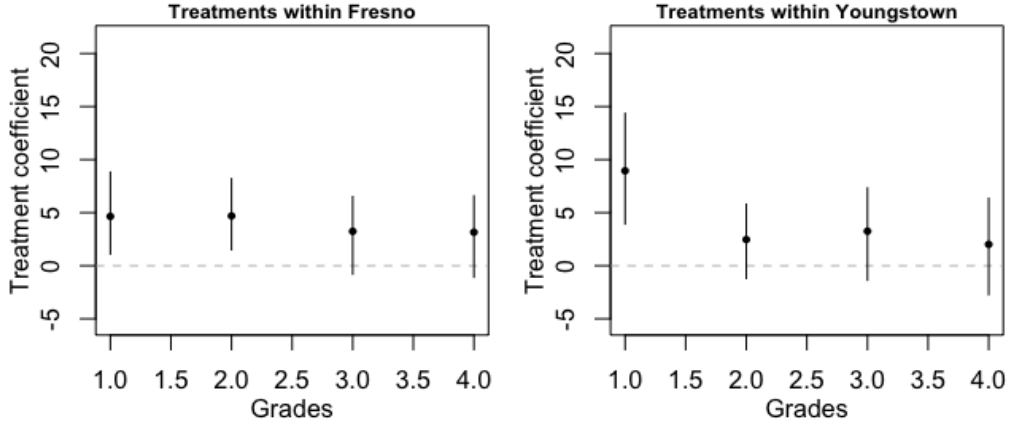


Figure 3: The points show the effect of treatment and the lines on either side represent the 95% interval. There is variation in treatment effects across grades in both cities and there is a positive treatment effect since all the points are above zero.

Figure 4 shows the pre and post score relationship divided according to cities. The regression lines drawn are using the city and grade level coefficients for each panel. So for example a grade 1 class in Fresno has it's own coefficient of treatment effect and pre test score and so on for all grades in Fresno and Youngstown.

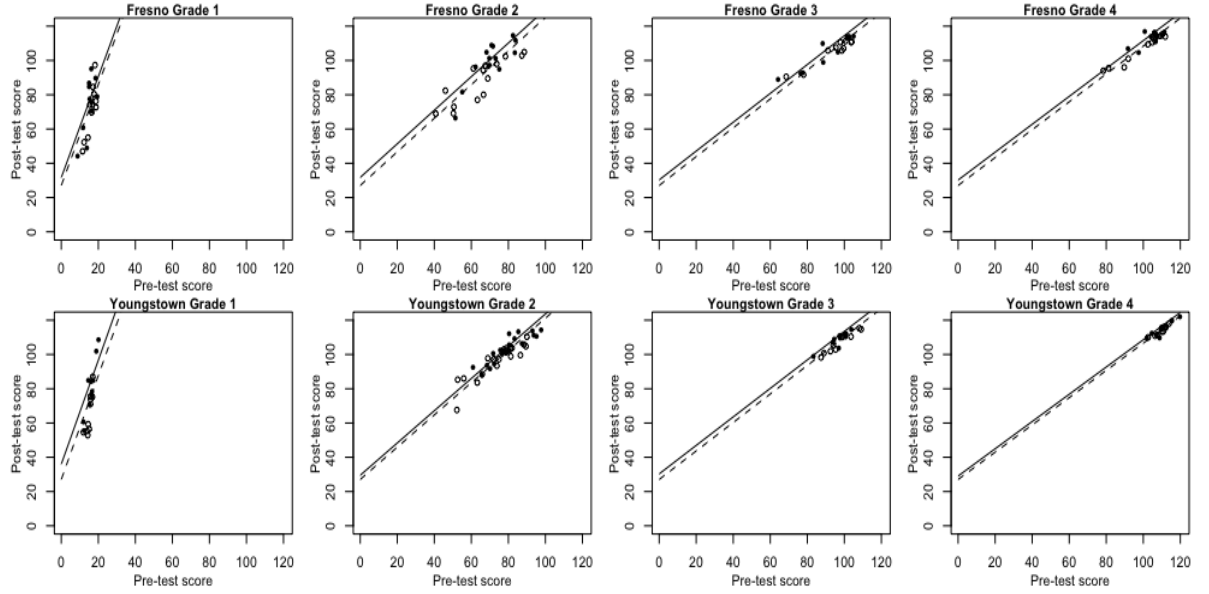


Figure 4: Pre and post test scores for each grade in each city. Treated and control classes are shown by dots and circles respectively. The solid and dotted lines in each panel show the treatment and control conditions respectively. The dotted lines are parallel displacements of the solid lines as they represent merely the absence of a treatment ($T_i = 0$).

1.2 Part B

Check that the data satisfy any assumptions made in your analysis.

I assume that treatments and controls were assigned randomly to classes. This is an intrinsic assumption of the experiment and I believe it is necessary to make this explicit. Now I assume that given the grade the pre test scores are treated as though recorded from the same test. This is a strong assumption and is not in general supported by the data. The reason for this is that grade 1 was given a different pre test but the same post test. The post test scores are therefore meaningful, but the baseline pre test scores are quite different for particularly grade 1. Figure 5 shows this explicitly.

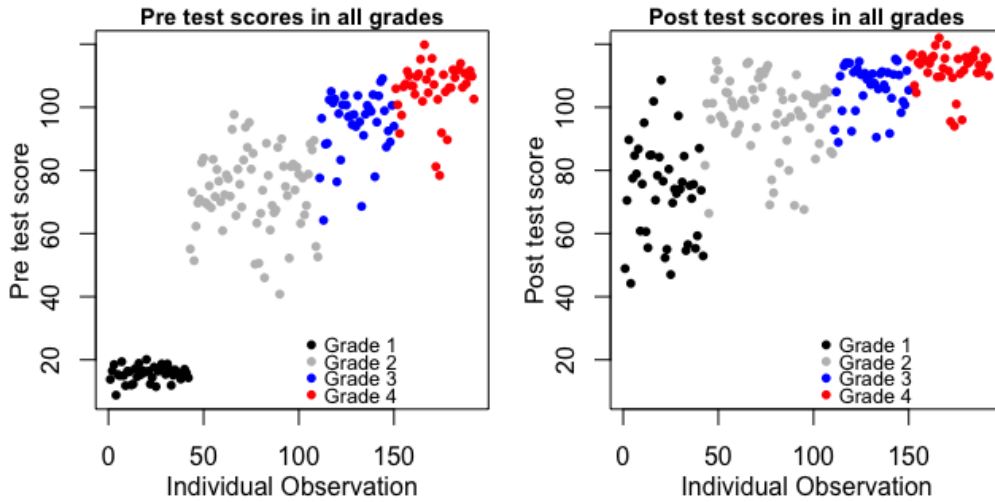


Figure 5: The pre scores for grade 1 are extremely low because of a different kind of test.

We observe in Figure 5 that the pre scores for grade 1 cluster at a very low value, naturally arising from the differential nature of testing. Moreover we observe that higher grades perform better upon testing. This makes sense intuitively. However, we see that many grade 1 students perform extremely well after watching the TV show, which means the treatment is positively impacting them as well but we would tend to overestimate the treatment effect for grade 1 students because in all my models I use the pre score as a predictor of the post test score, but the pre and post test not being the same for grade 1 is a fact that dents this kind of approach to analyzing the treatment effects. Further we assume that pre and post test scores differ across cities. This does not mean extreme differences, but there is definitely differences across grades between cities that has to be taken into account. This spread is shown by city in Figure 6.

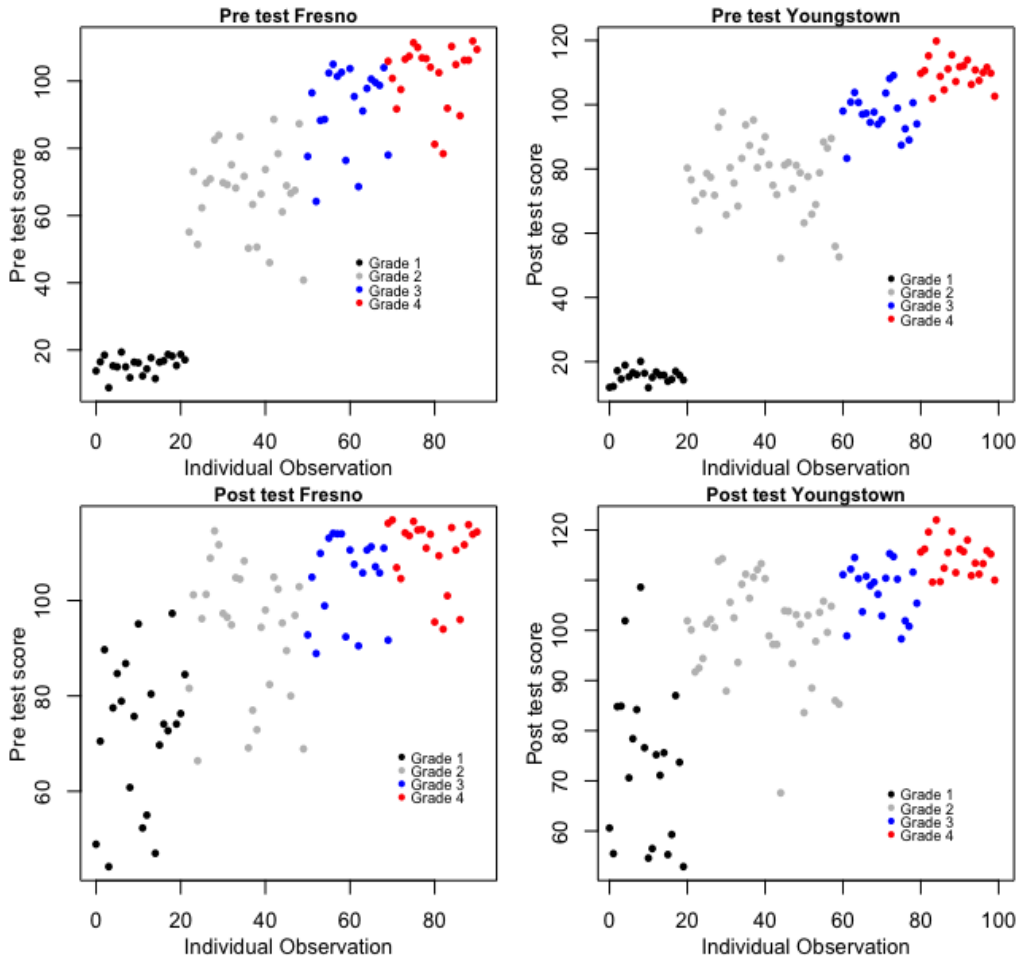


Figure 6: Pre and Post test scores for the two cities.

This last figure shows us that it is important to consider city wise and grade wise variation in any analysis and the data satisfies this assumption. Lastly there is the complete disregard of the replacement versus supplementary usage of the TV show in classes and this affects the assumption that classes are assigned the treatment and control randomly as well as does not accurately reflect the effect of just the TV show on test scores. However as instructed, we have ignored this aspect of the data for this particular problem.

2 code

2.1 R Code

Listing 1: R Code

```
##Question 1
rm(list = ls())
setwd("/Users/Advait/Desktop/New_School/Fall16/BDA/Class15")
library(rstan)
rstan_options(auto_write = TRUE)
options(mc.cores = parallel::detectCores())
##
elec <- read.table("http://www.stat.columbia.edu
/~gelman/bda.course/electric.txt", header = T, skip = 1)
str(elec)
dim(elec)
elec[1:5,]
##
par(mfcol = c(4,2), mar = c(3,1,1,1))
hist(elec$Posttest.1[elec$Grade==1], col = "gray",
      breaks = 10, xlim = c(0,120),
      yaxt = "n", ylab = NULL,
      xlab = NULL, main = "Control_Grade_1",
      cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest.1[elec$Grade==1]),2), "\n_sd_=",
round(sd(elec$Posttest.1[elec$Grade==1]))))

hist(elec$Posttest.1[elec$Grade==2], col = "gray",
      breaks = 10, xlim = c(0,120), yaxt = "n",
      ylab = NULL, xlab = NULL, main = "Control_Grade_2",
      cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest.1[elec$Grade==2]),2), "\n_sd_=",
round(sd(elec$Posttest.1[elec$Grade==2]))))

hist(elec$Posttest.1[elec$Grade==3], col = "gray", breaks = 10,
xlim = c(0,120), yaxt = "n", ylab = NULL,
xlab = NULL, main = "Control_Grade_3",
      cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest.1[elec$Grade==3]),2), "\n_sd_=",
round(sd(elec$Posttest.1[elec$Grade==3]))))
```

Listing 2: R code contd.

```
hist(elec$Posttest.1[elec$Grade==4], col = "gray",
breaks = 10, xlim = c(0,120), yaxt = "n", ylab = NULL, xlab = NULL,
main = "Control_Grade_4",
    cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest.1[elec$Grade==4]),2), "\n sd_=",
round(sd(elec$Posttest.1[elec$Grade==4]))))

#
hist(elec$Posttest[elec$Grade==1], col = "gray", breaks = 10,
xlim = c(0,120), yaxt = "n", ylab = NULL, xlab = NULL,
main = "Treat_Grade_1",
    cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest[elec$Grade==1]),2), "\n sd_=",
round(sd(elec$Posttest[elec$Grade==1]))))

hist(elec$Posttest[elec$Grade==2], col = "gray", breaks = 10,
xlim = c(0,120), yaxt = "n", ylab = NULL, xlab = NULL,
main = "Treat_Grade_2",
    cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest[elec$Grade==2]),2), "\n sd_=",
round(sd(elec$Posttest[elec$Grade==2]))))

hist(elec$Posttest[elec$Grade==3], col = "gray",
breaks = 10, xlim = c(0,120), yaxt = "n", ylab = NULL,
xlab = NULL, main = "Treat_Grade_3",
    cex.main = 0.8)
legend(10, 5, bty = "n", legend = paste("Mean_=",
round(mean(elec$Posttest[elec$Grade==3]),2), "\n
sd_=", round(sd(elec$Posttest[elec$Grade==3]))))

hist(elec$Posttest[elec$Grade==4], col = "gray",
breaks = 10, xlim = c(0,120), yaxt = "n",
ylab = NULL, xlab = NULL, main = "Treat_Grade_4",
    cex.main = 0.8)
legend(10, 5, bty = "n", legend =
paste("Mean_=", round(mean(elec$Posttest[elec$Grade==4]),2),
"\n sd_=", round(sd(elec$Posttest[elec$Grade==4]))))
```


Listing 3: R code contd.

```
# ####
# PLOT 2
# ####

g1treat <- elec$Posttest[elec$Grade==1]
g1treat.pre <- elec$Pretest[elec$Grade==1]
g1cont <- elec$Posttest.1[elec$Grade==1]
g1cont.pre <- elec$Pretest.1[elec$Grade==1]
g1post <- c(g1treat,g1cont)
g1pre <- c(g1treat.pre,g1cont.pre)
t1 <- c(rep(1,21),rep(0,21))
city1 <- c(rep("F",11),rep("Y",10),rep("F",11),rep("Y",10))
g1 <- data.frame(g1pre,g1post,t1,city1)
names(g1) <- c("Pre", "Post", "Treat", "City")

#
g2treat <- elec$Posttest[elec$Grade==2]
g2treat.pre <- elec$Pretest[elec$Grade==2]
g2cont <- elec$Posttest.1[elec$Grade==2]
g2cont.pre <- elec$Pretest.1[elec$Grade==2]
length(g2treat)
g2post <- c(g2treat,g2cont)
g2pre <- c(g2treat.pre,g2cont.pre)
t2 <- c(rep(1,34),rep(0,34))
city2 <- c(rep("F",14),rep("Y",20),rep("F",14),rep("Y",20))
g2 <- data.frame(g2pre,g2post,t2,city2)
names(g2) <- c("Pre", "Post", "Treat", "City")

#
g3treat <- elec$Posttest[elec$Grade==3]
g3treat.pre <- elec$Pretest[elec$Grade==3]
g3cont <- elec$Posttest.1[elec$Grade==3]
g3cont.pre <- elec$Pretest.1[elec$Grade==3]
length(g3treat)
g3post <- c(g3treat,g3cont)
g3pre <- c(g3treat.pre,g3cont.pre)
t3 <- c(rep(1,20),rep(0,20))
city3 <- c(rep("F",10),rep("Y",10),rep("F",10),rep("Y",10))
g3 <- data.frame(g3pre,g3post,t3,city3)
names(g3) <- c("Pre", "Post", "Treat", "City")
```

Listing 4:

```
#
g4treat <- elec$Posttest[elec$Grade==4]
g4treat.pre <- elec$Pretest[elec$Grade==4]
g4cont <- elec$Posttest.1[elec$Grade==4]
g4cont.pre <- elec$Pretest.1[elec$Grade==4]
length(g4treat)
g4post <- c(g4treat,g4cont)
g4pre <- c(g4treat.pre,g4cont.pre)
t4 <- c(rep(1,21),rep(0,21))
city4 <- c(rep("F",11),rep("Y",10),rep("F",11),rep("Y",10))
g4 <- data.frame(g4pre,g4post,t4,city4)
names(g4) <- c("Pre", "Post", "Treat", "City")

##
snoop <- rbind(g1,g2,g3,g4)
snoop <- data.frame(snoop)
str(snoop)
names(snoop) <- c("pre", "post", "treat","city")
snoop$grade <- c(rep(1,42),rep(2,68),rep(3,40), rep(4,42))
#####
par(mfcol = c(1,4),
    mar=c(3, 3, 2, 1),
    oma=c(.5, .5, .5, .5),
    mgp=c(2, 1, 0))
for(i in 1:4){
plot(snoop$pre[snoop$grade==i & snoop$treat==1],
     snoop$post[snoop$grade==i & snoop$treat==1],
     xlim = c(1,120),ylim = c(1,120),
     xlab = "Pre-test_score",
     ylab = "Post-test_score",
     main = paste("Grade",i), cex = 0.8, pch = 16)
points(snoop$pre[snoop$grade== i & snoop$treat==0],
       snoop$post[snoop$grade== i & snoop$treat==0],
       xlim = c(1,120),ylim = c(1,120),cex = 0.8)
abline(lm(snoop$post[snoop$grade==i &
snoop$treat==1]~
snoop$pre[snoop$grade== i & snoop$treat==1]))
abline(lm(snoop$post[snoop$grade==i &
snoop$treat==0]~
snoop$pre[snoop$grade==i & snoop$treat==0]), lty = 2)
}
```

Listing 5: R code contd.

```
#####
snoop[1:5,]
str(snoop)
snoop$treat <- ifelse(snoop$treat==1, 0.5,-0.5)
snoop$city <- ifelse(snoop$city=="F", 1,2)
pre <- snoop$pre
post <- snoop$post
treat <- snoop$treat
city <- snoop$city
grade <- snoop$grade
stanc("8b.stan")
fit1 <- stan("8b.stan",
             data = list("pre", "post", "treat", "city","grade"),
             iter = 1000,
             chains = 3)

print(fit1)
ext1 <- extract(fit1)
str(ext1)
thetas <- matrix(ext1$theta, nrow=1500, ncol=8)
lower <- NULL
for(i in 1:8){
  lower[i] <- quantile(thetas[,i], 0.025) }
upper <- NULL
for(i in 1:8){
  upper[i] <- quantile(thetas[,i], 0.975) }
lower
treatment_fx <- colMeans(ext1$theta)
#####Plot 3
par(mfcol = c(1,2))
plot(c(1,2,3,4), treatment_fx[,1], main = "Treatments_within_Fresno",
     cex.main = 0.8, xlab = "Grades", ylab = "Treatment_coefficient",
     pch = 16, cex = 0.7, ylim = c(2*min(lower), 1.5*max(upper)))
abline(h = 0, lty = 2,col = "gray")
arrows(c(1:4), treatment_fx[,1],c(1:4),
upper[1:4], col = "black",length = 0)
arrows(c(1:4), treatment_fx[,1],c(1:4),
lower[1:4], col = "black",length = 0)
plot(seq(1,4, by = 1), treatment_fx[,2],
main = "Treatments_within_Youngstown",
     cex.main = 0.8, xlab = "Grades",
     ylab = "Treatment_coefficient",
     pch = 16, cex = 0.7, ylim = c(2*min(lower), 1.5*max(upper)))
```

Listing 6: R Code contd.

```
abline(h = 0, lty = 2,col = "gray")
arrows(c(1:4), treatment_fx[,2],c(1:4),
upper[5:8], col = "black",length = 0)
arrows(c(1:4), treatment_fx[,2],c(1:4),
lower[5:8], col = "black",length = 0)
###
#Plot city wise treatment-control relationships
##
#Plot Relationship between pre and post for all grades and cities
par(mfrow = c(2,4),mar = c(3, 3, 1, 1),
oma = c(.5, .5, .5, .5), mgp=c(2,1,0))
str(snoop)
for(i in 1:4){
  plot(snoop$pre[snoop$grade == i &
snoop$treat==0.5 & snoop$city == 1],
      snoop$post[snoop$grade == i &
snoop$treat==0.5 & snoop$city == 1],
      xlab = "Pre-test_score",
      ylab = "Post-test_score",
      cex = 0.8, pch = 16, main = paste("Fresno_Grade",i),
      xlim = c(1,120),
      ylim = c(0,120),
      cex.main = 1)
  points(snoop$pre[snoop$grade == i &
snoop$treat== -0.5 & snoop$city == 1],
        snoop$post[snoop$grade == i &
snoop$treat== -0.5 & snoop$city == 1],
        cex = 0.8)
  lines(c(0,120),
        mean(ext1$b0)
        + colMeans(ext1$beta)[i,1]*c(0,120)
        + colMeans(ext1$theta)[i,1])
  lines(c(0,120),
        mean(ext1$b0)
        + colMeans(ext1$beta)[i,1]*c(0,120)
        ,lty = 2)
}
```

Listing 7: R code contd.

```
for(i in 1:4){
  plot(snoop$pre[snoop$grade == i & snoop$treat==0.5 &
    snoop$city == 2],
    snoop$post[snoop$grade == i & snoop$treat==0.5 &
    snoop$city == 2],
    xlab = "Pre-test_score",
    ylab = "Post-test_score",
    cex = 0.8, pch = 16, main = paste("Youngstown_Grade",i),
    xlim = c(1,120),
    ylim = c(0,120),
    cex.main = 1)
  points(snoop$pre[snoop$grade == i & snoop$treat== -0.5 &
    snoop$city == 2],
    snoop$post[snoop$grade == i & snoop$treat== -0.5 &
    snoop$city == 2],
    cex = 0.8)
  lines(c(0,120),
    mean(ext1$b0)
    + colMeans(ext1$beta)[i,2]*c(0,120)
    + colMeans(ext1$theta)[i,2])
  lines(c(0,120),
    mean(ext1$b0)
    + colMeans(ext1$beta)[i,2]*c(0,120)
    ,lty = 2)
}
###Part B
str(snoop)
par(mfrow = c(1,2),mar = c(3, 3, 1, 1),
oma = c(.5, .5, .5, .5), mgp=c(2,1,0))
plot(c(1:192), snoop$pre ,col = "white",
  main = "Pre_test_scores_in_all_grades",
  cex.main = 0.9,
  xlab = "Individual_Observation",
  ylab = "Pre_test_score")
points(c(1:42),snoop$pre[snoop$grade==1],
col = "black", cex = 0.8, pch = 16)
points(c(43:110),snoop$pre[snoop$grade==2],
col = "gray", cex = 0.8, pch = 16)
points(c(111:150),snoop$pre[snoop$grade==3],
col = "blue", cex = 0.8, pch = 16)
points(c(151:192), snoop$pre[snoop$grade==4],
col = "red", cex = 0.8, pch = 16)
```

Listing 8: R Code contd.

```
legend(100,30, bty = "n",
      legend = c("Grade_1", "Grade_2",
                  "Grade_3", "Grade_4"),
      col = c("black", "gray", "blue", "red"), pch = 16, cex = 0.8)
plot(c(1:192), snoop$pre ,col = "white",
     main = "Post_test_scores_in_all_grades",
     cex.main = 0.9,
     xlab = "Individual_Observation",
     ylab = "Post_test_score")
points(c(1:42), snoop$post[snoop$grade==1],
       col = "black", cex = 0.8, pch = 16)
points(c(43:110), snoop$post[snoop$grade==2],
       col = "gray", cex = 0.8, pch = 16)
points(c(111:150), snoop$post[snoop$grade==3],
       col = "blue", cex = 0.8, pch = 16)
points(c(151:192), snoop$post[snoop$grade==4],
       col = "red", cex = 0.8, pch = 16)
legend(100,30, bty = "n",
      legend = c("Grade_1", "Grade_2",
                  "Grade_3", "Grade_4"),
      col = c("black", "gray", "blue", "red"), pch = 16, cex = 0.8)
#####
par(mfrow = c(2,2),mar = c(3, 3, 1, 1),
    oma = c(.5, .5, .5, .5), mgp=c(2,1,0))
plot(c(0:91), snoop$pre[snoop$city==1] ,col = "white",
     main = "Pre_test_Fresno",
     cex.main = 0.9,
     xlab = "Individual_Observation",
     ylab = "Pre_test_score")
points(c(0:21), snoop$pre[snoop$grade==1 & snoop$city==1],
       col = "black", cex = 0.8, pch = 16)
points(c(22:49), snoop$pre[snoop$grade==2 & snoop$city==1],
       col = "gray", cex = 0.8, pch = 16)
points(c(50:69), snoop$pre[snoop$grade==3 & snoop$city==1],
       col = "blue", cex = 0.8, pch = 16)
points(c(69:90), snoop$pre[snoop$grade==4 & snoop$city==1],
       col = "red", cex = 0.8, pch = 16)
legend(60,50, bty = "n",
      legend = c("Grade_1", "Grade_2", "Grade_3", "Grade_4"),
      col = c("black", "gray", "blue", "red"), pch = 16, cex = 0.7)
##
```

Listing 9: R code contd.

```
plot(c(1:100), snoop$pre[snoop$city==2] ,col = "white",
     main = "Pre_test_Youngstown",
     cex.main = 0.9,
     xlab = "Individual_Observation",
     ylab = "Post_test_score")
points(c(0:19),snoop$pre[snoop$grade==1 &
  snoop$city==2],
col = "black", cex = 0.8, pch = 16)
points(c(20:59),snoop$pre[snoop$grade==2 &
  snoop$city==2],
col = "gray", cex = 0.8, pch = 16)
points(c(60:79),snoop$pre[snoop$grade==3 &
  snoop$city==2],
col = "blue", cex = 0.8, pch = 16)
points(c(80:99), snoop$pre[snoop$grade==4 &
  snoop$city==2],
col = "red", cex = 0.8, pch = 16)
legend(70,50, bty = "n",
      legend = c("Grade_1", "Grade_2", "Grade_3", "Grade_4"),
      col = c("black", "gray", "blue", "red"), pch = 16, cex = 0.7)
#####
plot(c(0:91), snoop$post[snoop$city==1] ,col = "white",
     main = "Post_test_Fresno",
     cex.main = 0.9,
     xlab = "Individual_Observation",
     ylab = "Pre_test_score")
points(c(0:21),snoop$post[snoop$grade==1 &
  snoop$city==1],
col = "black", cex = 0.8, pch = 16)
points(c(22:49),snoop$post[snoop$grade==2 &
  snoop$city==1],
col = "gray", cex = 0.8, pch = 16)
points(c(50:69),snoop$post[snoop$grade==3 &
  snoop$city==1],
col = "blue", cex = 0.8, pch = 16)
points(c(69:90), snoop$post[snoop$grade==4 &
  snoop$city==1],
col = "red", cex = 0.8, pch = 16)
legend(70,70, bty = "n",
      legend = c("Grade_1", "Grade_2", "Grade_3", "Grade_4"),
      col = c("black", "gray", "blue", "red"), pch = 16, cex = 0.7)
```

Listing 10: R code contd.

```
##
plot(c(1:100), snoop$post[snoop$city==2] ,col = "white",
     main = "Post_test_Youngstown",
     cex.main = 0.9,
     xlab = "Individual_Observation",
     ylab = "Post_test_score")
points(c(0:19), snoop$post[snoop$grade==1 & snoop$city==2],
       col = "black", cex = 0.8, pch = 16)
points(c(20:59), snoop$post[snoop$grade==2 & snoop$city==2],
       col = "gray", cex = 0.8, pch = 16)
points(c(60:79), snoop$post[snoop$grade==3 & snoop$city==2],
       col = "blue", cex = 0.8, pch = 16)
points(c(80:99), snoop$post[snoop$grade==4 & snoop$city==2],
       col = "red", cex = 0.8, pch = 16)
legend(70,70, bty = "n",
      legend = c("Grade_1", "Grade_2", "Grade_3", "Grade_4"),
      col = c("black", "gray", "blue", "red"), pch = 16, cex = 0.7)
```


2.2 Stan Code

Listing 11: Stan Code

```
data{
  vector[192]post;
  vector[192]pre;
  vector[192]treat;
  int grade[192];
  int city[192];
}
parameters{
  real b0;
  matrix [4,2] beta;
  matrix [4,2] theta;
  real <lower = 0> sigma;
  real mu_F;
  real<lower = 0> tau_F;
  real mu_Y;
  real<lower = 0> tau_Y;
  real mu_beta;
  real<lower = 0> tau_beta;
}
model{
  for(i in 1:192){
    post[i] ~ normal(b0 + beta[grade[i], city[i]]*pre[i] +
                     theta[grade[i], city[i]]*treat[i] ,
                     sigma);
  }
  theta[, 1] ~ normal(mu_F, tau_F);
  theta[, 2] ~ normal(mu_Y, tau_Y);
  mu_F ~ normal(5,5);
  mu_Y ~ normal(5,5);
  tau_F ~ normal(3,2);
  tau_Y ~ normal(3,2);
  for(i in 1:4){
    beta[i,] ~ normal(mu_beta, tau_beta);
  }
}
```