The 3-billion School Improvement Grants (SIG) program has no significant effects on student outcome

Sharon Allman, Diego Mamanche Castellano & Ke-Li Chiu

22/03/2020

Abstract

The objective of this analysis was to examine if there was a relationship between a school's participation in a School Improvement Grant program and an increase in that school's students' achievement. Using linear regression models, we failed to reject the null hypothesis that the SIG programs had no statistically significant effect on students' performance over time.

Introduction

In 2009, the US Government allocated $3 billion dollars under the American Recovery and Reinvestment Act of 2009 into School Improvement Grants (SIG) to turn around the lowest performing schools. As one of the Obama administration's signature programs, SIG is also one of the largest investments in education grants made by the US government. Did SIG drive changes in outcome of schools' performance and students' achievement? That is the 3-billion-dollar question.

The SIG awarded persistently low-performing schools sizeable funding to bring about radical school changes by implementing their intervention models that included specific practices such as increased learning time, integration of technology in classrooms, and the replacement of administration management. In this study, we evaluated the impact of the SIG program by observing the schools that have implemented the models in all four years to see the change of percentage of students meeting a level of proficiency in math assessments. As a result, despite the substantial amount of investment, we found no statistically significant effect of the program on the students' proficiency outcome.

Dataset Description and Data Cleaning

The dataset sy1011-1314 includes data for four school years from 2010-11 to 2013-14. The data contains the list of schools who got the funding to implement SIG programs in each year. Year 2010-11 is implementation year 1, 2011-12 is implementation year 2, 2012-13 is implementation year 3, and 2013-14 is implementation year 4. We kept only the rows of schools who had implemented SIG models in all four implementation years. It is possible to assess the program in each implementation year. However, we choose to observe the long-term effect of the program. Therefore, the data from 2009-10 (before implementation year 1) is compared with data from implementation year 4.

To assess the effect of the SIG program, we wanted to observe the percentage change of students in the schools that scored above proficiency in math assessment before implementation 1 (2009-10) and at implementation year 4 (2013-14). Therefore, we gathered math achievement data sets from school years 2009-10 and 2013-14 and extracted the column of math proficiency rate from each data set. The math proficiency rate has a range from 0 to 100. Moreover, the datasets contain schools that did not implement SIG programs in any of the implementation years, which allowed us to formulate a control group with random assignment method. In the end, we had a data frame observing 183 schools across 5 school years with 8 variables concerning the schools' name, region, model implemented, math proficiency rate, and school year.

Method

The method we used to estimate the effects of the SIG programs was difference-in-differences (diff-in-diff). The initial difference between the treatment and control group was taken into account. Because the SIG program only awarded the schools that ranked as low performing, the average proficiency rate in the treatment group was substantially lower than the ones in the control group— hence we did not compare the differences of proficiency rate between treatment and control group. What we wanted to compare were the changes in the treatment group over time to changes in the control group over time.

Summary Statistics

The distribution of the control group and treatment group before and after the program implementation is shown in Figure 1. The mean of the proficiency rate in both school years remained under 40%. The maximum proficiency rate was 93% in the school year 2009-2010. The lowest proficiency rate was 2% in the school year 2013-2014. In the control group, the mean of the proficiency rate across the five school years was above 60%. The maximum proficiency rate was 98% in both school years 2009-2010 and 2013-2014. The lowest proficiency rate was 5% in the school year 2013-2014.
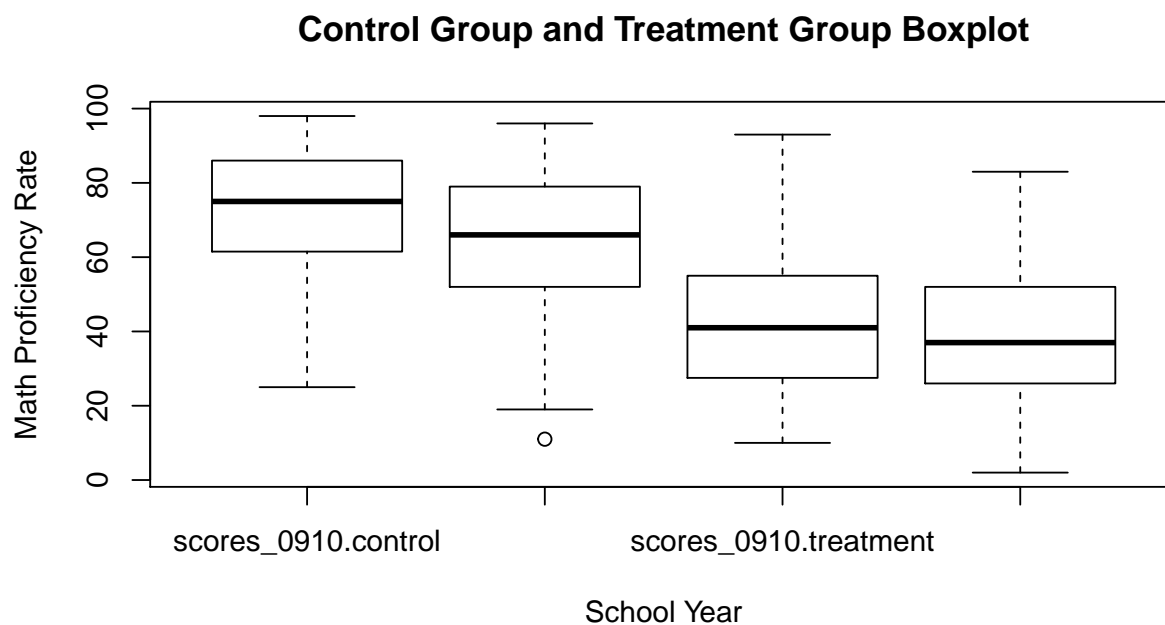

Figure 1. Summary statistics

Research Question and Hypotheses

Within this research, we aimed to examine the impact of SIG programs over time.

H0: The SIG programs had no significant effect on students' performance over time.

H1: The SIG programs had a significant effect on students' performance over time.

H2: The SIG programs had a significant, positive effect on students' performance over time.

Ethical Issues

This data was sourced from the United States Department of Education and is publicly available, so logically this data has been largely anonymized. This data does not contain specific demographic information that could directly identify any of the participants. However, it does contain very specific information about the schools involved, especially the names of the schools and the states they are located in. This may be problematic as this could be theoretically used to de-anonymize data specific to individuals if combined with other demographic data (about students/participants in relation to these states or the names of these schools). It is unlikely that individual respondents would be able to request to have data related to them removed from this data, as this is data collected by the government and does not have personally identifying information. Additionally, clearly identifying the worst-performing schools could be further stigmatizing to the institutions and individuals associated with them.

This data should be representative of the program, as this data is cumulative of most of the cases of this program within the specific time frames. As such, research conducted with this data should also have higher external validity. However, it may be ethically ambiguous for the government to be publishing positive findings about their program (The U.S. Department of Education, 2013), when independent research sources were unable to demonstrate that this program had a significant impact (Le Floch, 2018). Independent research firms have suggested that SIG programs should have been studied at the state level, as this would have provided more concise findings in terms of how the SIG schools were doing compared to their state counterparts, rather than comparing the best and worst cases in the country overall (Smith & Ginsburg, 2018). Even the U.S. Department of Education itself released a report after the fact outlining the ambiguity of the effectiveness of the program (Dragoset et al., 2017), contradicting what they had first reported in 2013 (The U.S. Department of Education, 2013).

In terms of the contradictory reports published by the government itself, this leads into another issue with this data and research that uses it. as multiple independent researchers have found, as well as the government itself, it is difficult to reproduce findings consistently. Depending on how models are constructed and evaluated, you can both reach the conclusion that these programs were effective and ineffective. This is a broader ethical issue in terms of both the reproducibility of the data within government studies as well as speaking to the quality and consistency of the data.

The U.S. Department of Education has broad ethical guidelines in terms of collection, dissemination, and analysis of data. However, ethical considerations specific to this program are not outlined in the data documentation or summary of findings. A multidisciplinary team was assembled in order to conceptualize and implement these programs (Kutash et al., 2010), however, it is unclear if this was also the case when it came to the data collection and analysis. This means that there may be some selection bias when it comes to the types of variables measured and reported on by the researchers. This could have been mitigated by ensuring they reported a transparent process of how the data was collected and analyzed, including the documentation and human resources used to reduce sampling bias.

Regarding the consent of the participants, especially as it was a government program, it is ambiguous if individual schools were able to opt out of data collection and subsequent government studies with this data while still being able to take part in SIG programs. This would make sense on one hand, as it is very useful to collect data like this about programs in order to evaluate them later. However, it does leave room for ethical issues regarding whether the schools or individuals themselves regarding informed consent (all students and faculty knew this data was being collected and what it was being collected for) and being able to provide consent not under any form of duress (not feeling like they needed to consent to data collection for fear of losing SIG funding).

Weaknesses & Limitations

As previously mentioned, research done using this data could theoretically have higher external validity, as this data is directly representative of the population impacted by the program. As this was a government-run program, analyzed by the government itself, the sampling included the entire population, ensuring that

the sample would be fully representative. Another consideration is that this data spans from 2009 to 2014. While it may be representative of that specific timeframe, it would not be representative of school settings today. Additionally, those organizing and analyzing this data have done so in a particular way, which may have included some sampling bias as they would have selected the measured variables according to what they had already planned to study.

As noted in a federal report by the U.S. Department of Education, they concluded that schools that had implemented SIG programs reported using more SIG program practices than other schools, but were unable to confidently assert that the SIG programs themselves caused those changes (Dragoset et al., 2017). This is relevant as an additional point to note when examining the final model; while this model is able to elaborate on the differences between the control and treatment groups, it may not have full power to assert that the SIG programs themselves caused the differences.

Analysis

Construction of Treatment Group and Control Group

To assure the comparability of the treatment group and the control group, we ensured that the two groups had the composition in the distribution of schools across states. The distribution of schools across states in treatment group is shown in Figure 2, and the control group distribution is shown in Figure 5.
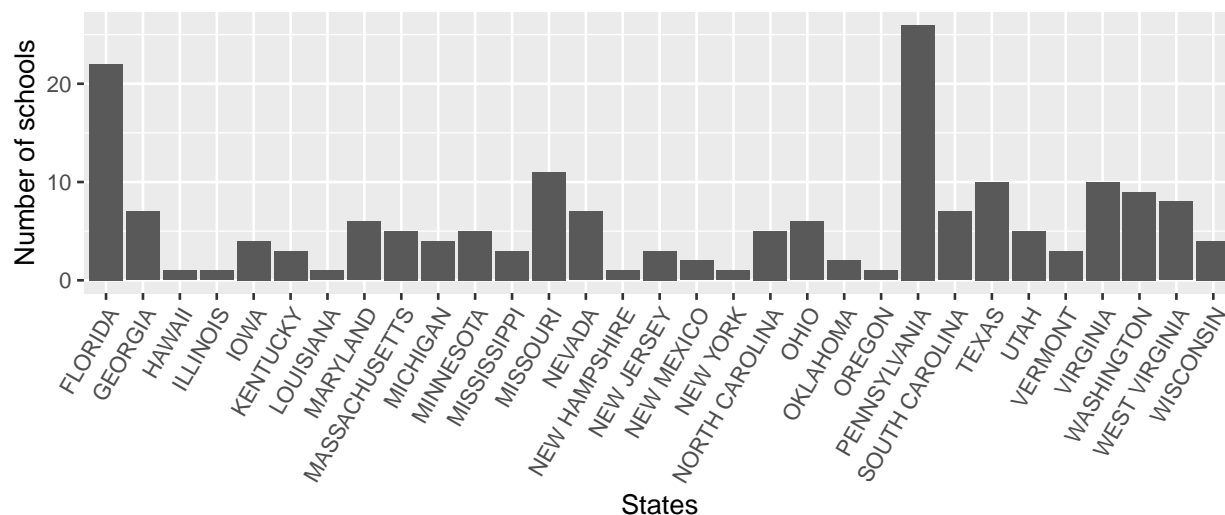

Figure 2. Distribution of schools which implemented the SIG models across states

According to the Parallel Trend Assumption, the difference between the 'treatment' and 'control' group is constant over time. To validate, we plot the data from 2009-10 (before the program implementation) to 2013-14 to observe the trends of the two groups. As shown in Figure 3, the two groups share a similar trend in the math proficiency rate.
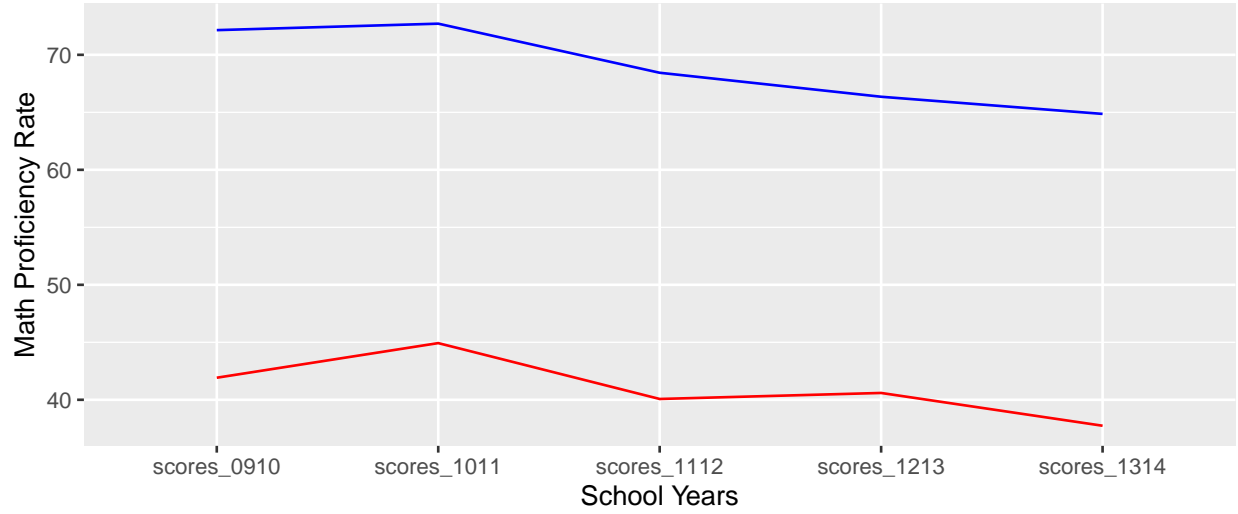
4

Figure 3. Trends of control group and treatment group

We plotted the proficiency rate change between 2009-10 and 2013-14 of each school in both treatment and control groups. As shown in Figure 4, there is no clear trend in either group.
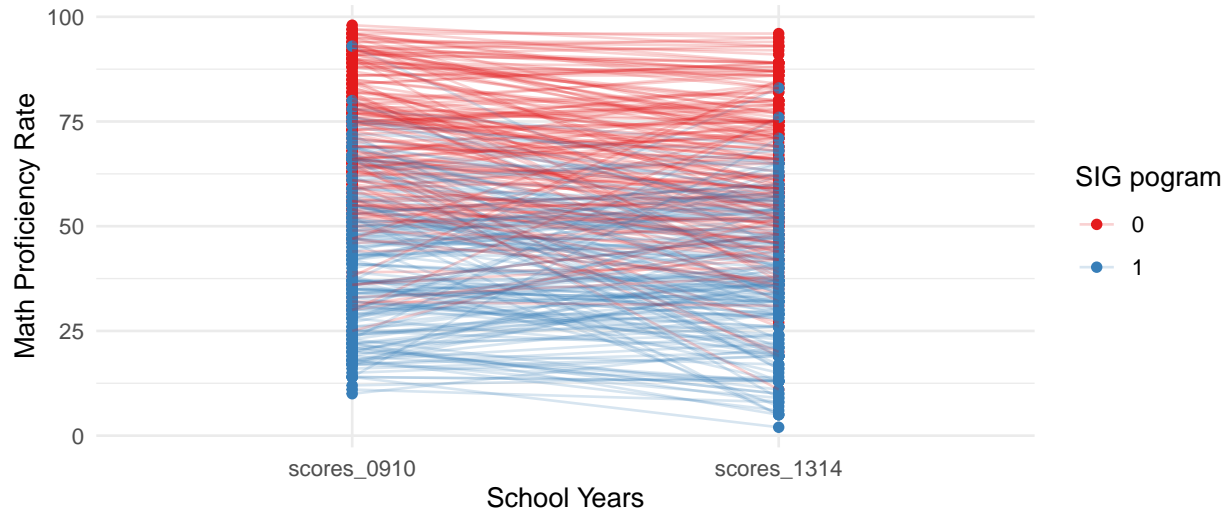

Figure 4. Proficiency rate of each school in both treatment and control groups

Linear Regression

Linear regression is applied to evaluate the effect of the SIG programs on proficiency rate across four school years. The equation used in the models is:

$$Y_{i,t} = \beta 0 + \beta 1(\text{Treatment group})_i + \beta 2(\text{Year})_t + \beta 3(\text{Treatment group} » \text{Year})_{i,t} + \epsilon_{i,t}$$

$\beta 3$ is how much the treatment group changes after the treatment event compared to how much the control group changes after the treatment event, which is the term we care about.

Model 1

Model 1 observes the change between school year 2009-10 and 2013-2014. Within Table 1, the average difference of the differences is 4.20. The treatment group in year 2013-2014 has a 4.2 percent increase in math assessment rate on average. However, the p-value of 3 is 0.06. This means that the effect of the SIG programs on the math proficiency rate would be insignificant at a 0.05 alpha level, with us failing to reject the null hypothesis. This may still be a relevant finding, however, as this P value is still very low and is close to significant if we used an alpha level of 0.05.

```
## Standard errors: OLS
## ---------------------------------------------------------
##                             Est.   S.E.   t val.      p
## ----------------------- -------- ------ -------- ------
## (Intercept)                72.15   1.29    55.97   0.00
## sig_program1              -30.23   1.82   -16.58   0.00
## year1                      -7.28   1.82    -4.00   0.00
## sig_program1:year1          3.11   2.58     1.21   0.23
## ---------------------------------------------------------
```

Table 1. OLS for diff-in-diff between 2009-10 and 2013-2014

According to a 2018 report by FutureEd, one of the reasons why more recently published federal reports may have reported no statistically significant improvements in student achievement at schools that implemented SIG programs is that these schools would need to have unrealistically large gains in achievement in order to consistently detect statistically significant improvement (Smith & Ginsburg, 2018). In turn, this may also play a role in our model as we were just shy of significance at a 0.05 alpha level.

Model 2: Model with another control group

To validate the model, we perform the diff-in-diff analysis with another comparison group to see if we would obtain similar estimates of the impact of the program. As shown in Table 2, the estimate of 2.97 is not too far away from 3.11 in Model 1.

```
## Standard errors: OLS
## ---------------------------------------------------------
##                             Est.   S.E.   t val.      p
## ----------------------- -------- ------ -------- ------
## (Intercept)                72.45   1.32    54.80   0.00
## sig_program1              -30.54   1.87   -16.33   0.00
## year1                      -7.14   1.87    -3.82   0.00
## sig_program1:year1          2.97   2.64     1.12   0.26
## ---------------------------------------------------------
```

Table 2. OLS for diff-in-diff between 2009-10 and 2013-2014 with another set of control groups

Model 3: Model with fake treatment group

Another validation method we used was to test the model on a fake treatment group that should not be affected by the program. As shown in Table 3, the estimate is now 0.01 and the p-value is insignificant.

```
## Standard errors: OLS
## ---------------------------------------------------------
```

```
##                              Est.   S.E.   t val.      p
## ----------------------- ------- ------ -------- ------
## (Intercept)               72.15   1.31    54.92   0.00
## sig_program1               0.07   1.86     0.04   0.97
## year1                     -7.28   1.86    -3.92   0.00
## sig_program1:year1         0.01   2.63     0.00   1.00
## ----------------------------------------------------------
```

Table 3. OLS for diff-in-diff between 2009-10 and 2013-2014 with fake treatment group
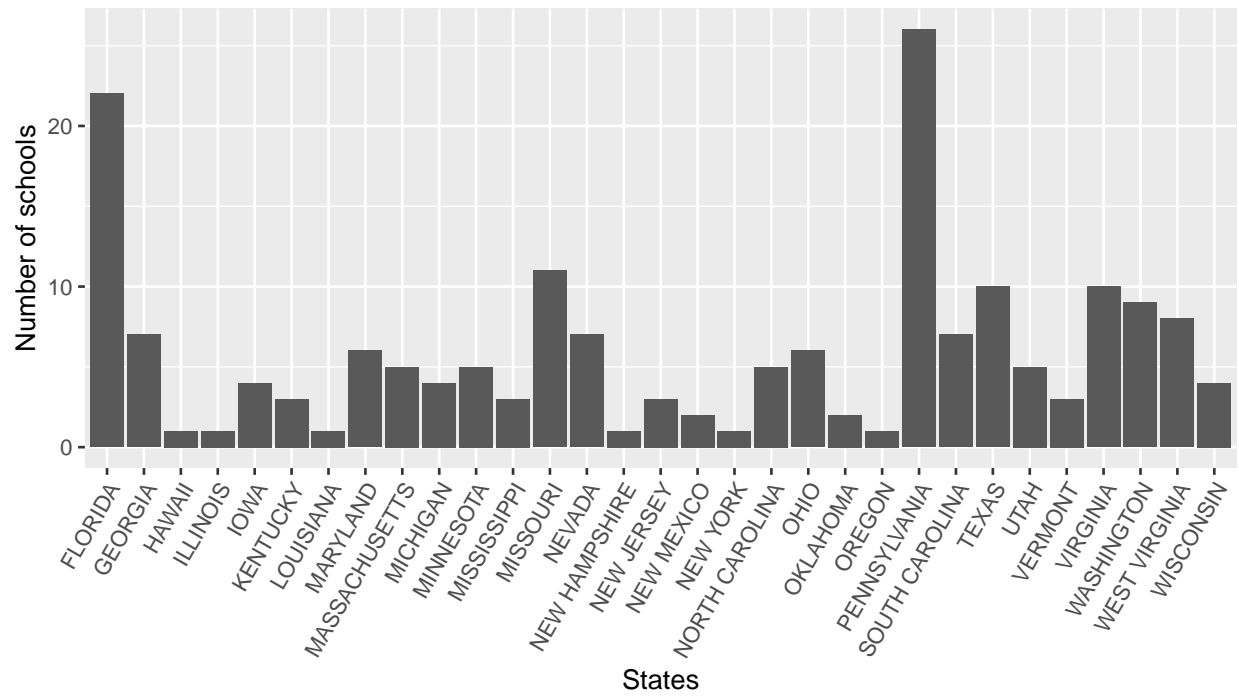
Appendix A

Figure 5. Distribution of schools which did not implement the SIG models across states

Appendix B

```r
knitr::opts_chunk$set(echo = FALSE, include = FALSE, cache = TRUE)## R Markdown
#Set up the environment
library(readxl)
library(dplyr)
library(tidyr)
library(ggplot2)
library(reshape2)
library(stringr)
library(knitr)
library(tidyverse)
library(broom)
library(huxtable)
library(jtools)
#setwd("~/Experimental Design for Data Science/ProblemSet4")
###Upload datasets
treated_schools <- read_excel("sy1011-1314.xlsx")
treated_schools <- janitor::clean_names(treated_schools)
math_assessment_09_10 <- read.csv("math-achievement-sch-sy2009-10.csv")
math_assessment_09_10 <- janitor::clean_names(math_assessment_09_10)
math_assessment_10_11 <- read.csv("math-achievement-sch-sy2010-11.csv")
math_assessment_10_11 <- janitor::clean_names(math_assessment_10_11)
math_assessment_11_12 <- read.csv("math-achievement-sch-sy2011-12.csv")
math_assessment_11_12 <- janitor::clean_names(math_assessment_11_12)
math_assessment_12_13 <- read.csv("math-achievement-sch-sy2012-13.csv")
math_assessment_12_13 <- janitor::clean_names(math_assessment_12_13)
math_assessment_13_14 <- read.csv("math-achievement-sch-sy2013-14.csv")
math_assessment_13_14 <- janitor::clean_names(math_assessment_13_14)
treated_all_cohorts <- filter(treated_schools,
                              sy201011sig_model != is.na(sy201011sig_model) &
                              sy201112sig_model != is.na(sy201112sig_model) &
                              sy201213sig_model != is.na(sy201213sig_model) &
                              sy201314sig_model != is.na(sy201314sig_model) &
                              ncessch_1011 == ncessch_1112 &
                              ncessch_1011 == ncessch_1213 &
                              ncessch_1011 == ncessch_1314 &
                              ncessch_1112 == ncessch_1213 &
                              ncessch_1112 == ncessch_1314 &
                              ncessch_1213 == ncessch_1314
                              )
treated_all_cohorts <- select(treated_all_cohorts, state, leaid_10_11,
                        leanm_1011, ncessch_1011, schnam_1011)
treated_all_cohorts$ncessch_1011 <- as.numeric(treated_all_cohorts$ncessch_1011)
colnames(treated_all_cohorts) <- c("state", "lea_id","lea_name","ncessch","school_name")
head(treated_all_cohorts)
treated_math_09_10 <- select(math_assessment_09_10, stnam, leaid,
                        leanm, ncessch, schnam09, all_mth00pctprof_0910)
treated_math_10_11 <- select(math_assessment_10_11, stnam, leaid,
                        leanm10, ncessch, schnam10, all_mth00pctprof_1011)
treated_math_11_12 <- select(math_assessment_11_12, stnam, leaid,
                        leanm, ncessch, schnam11, all_mth00pctprof_1112)
treated_math_12_13 <- select(math_assessment_12_13, stnam, leaid,
```

```r
                                    leanm, ncessch, schnam, all_mth00pctprof_1213)
treated_math_13_14 <- select(math_assessment_13_14, stnam, leaid,
                                    leanm, ncessch, schnam, all_mth00pctprof_1314)
merged_math <- merge(treated_math_09_10, treated_math_10_11, by= "ncessch")
merged_math <- select(merged_math, ncessch, schnam09, stnam.x,
                      leaid.x, leanm, all_mth00pctprof_0910, all_mth00pctprof_1011)
merged_math <- merge(merged_math, treated_math_11_12, by= "ncessch")
merged_math <- select(merged_math, ncessch, schnam09, stnam.x,
                      leaid.x, leanm.x, all_mth00pctprof_0910, all_mth00pctprof_1011,
                      all_mth00pctprof_1112)
merged_math <- merge(merged_math, treated_math_12_13, by= "ncessch")
merged_math <- select(merged_math, ncessch, schnam09, stnam.x,
                      leaid.x, leanm.x, all_mth00pctprof_0910, all_mth00pctprof_1011,
                      all_mth00pctprof_1112, all_mth00pctprof_1213)
merged_math <- merge(merged_math, treated_math_13_14, by= "ncessch")
merged_math <- select(merged_math, ncessch, schnam09, stnam.x,
                      leaid.x, leanm.x, all_mth00pctprof_0910, all_mth00pctprof_1011,
                      all_mth00pctprof_1112, all_mth00pctprof_1213, all_mth00pctprof_1314)
merged_math
### Create Treatment DataSet ###
### Create Treatment DataSet ###
results_treated <- merge(treated_all_cohorts, merged_math, by= "ncessch")
results_treated <- filter(results_treated,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_0910) )
results_treated <- filter(results_treated,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1011) )
results_treated <- filter(results_treated,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1112) )
results_treated <- filter(results_treated,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1213) )
results_treated <- filter(results_treated,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1314) )
results_treated <- select(results_treated, ncessch, schnam09, stnam.x,
                          leaid.x, leanm.x, all_mth00pctprof_0910,
                          all_mth00pctprof_1011, all_mth00pctprof_1112,
                          all_mth00pctprof_1213, all_mth00pctprof_1314)
colnames(results_treated) <- c("ncessch", "school_name", "state", "lea_id",
                               "lea_name", "scores_0910", "scores_1011",
                               "scores_1112", "scores_1213", "scores_1314")
results_treated
diff_in_diff_treated <- melt(results_treated, measure = 6:10)
diff_in_diff_treated[,7] <-
sapply(diff_in_diff_treated[,7], function(x) as.numeric(x))
colnames(diff_in_diff_treated) <- c("ncessch", "school_name", "state", "lea_id",
                               "lea_name", "cohort","percentage")
### Create Control Group ###
merged_math_cleaned <- merged_math
unique_treated_schools <- melt(treated_schools, measure = 10:13)
```

```r
unique_treated_schools <- unique(unique_treated_schools$value)
merged_math_cleaned <- filter(merged_math, !(ncessch %in%
                                 as.numeric(unique_treated_schools)))
merged_math_cleaned <- filter(merged_math_cleaned,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_0910) )
merged_math_cleaned <- filter(merged_math_cleaned,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1011) )
merged_math_cleaned <- filter(merged_math_cleaned,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1112) )
merged_math_cleaned <- filter(merged_math_cleaned,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1213) )
merged_math_cleaned <- filter(merged_math_cleaned,
                       !grepl("(([0-9]+[-][0-9]+)|([a-zA-Z]+[0-9]+)|([a-zA-z]+))",
                              all_mth00pctprof_1314) )
results_control <- data.frame()
for (i in unique(results_treated$state)) {

    count_state <- count(filter(results_treated, state == i))

    filter_by_state <- which(merged_math_cleaned$stnam.x == i)

    set.seed(178)
    sample_by_state <- merged_math_cleaned[sample(filter_by_state,
                                           as.numeric(count_state)),]

    results_control <- rbind(results_control, sample_by_state)

    }
colnames(results_control) <- c("ncessch", "school_name", "state", "lea_id",
                               "lea_name", "scores_0910", "scores_1011",
                               "scores_1112", "scores_1213", "scores_1314")
diff_in_diff_controled <- melt(results_control, measure = 6:10)
diff_in_diff_controled[,7] <-
sapply(diff_in_diff_controled[,7], function(x) as.numeric(x))
colnames(diff_in_diff_controled) <- c("ncessch", "school_name", "state", "lea_id",
                               "lea_name", "cohort","percentage")
results_treated_0910_1314 <- results_treated %>% select(scores_0910,scores_1314)
results_control_0910_1314 <- results_control %>% select(scores_0910,scores_1314)

diff_in_diff_treated_0910_1314 <- melt(results_treated_0910_1314, measure = 1:2)
diff_in_diff_treated_0910_1314 <- mutate(diff_in_diff_treated_0910_1314, sig_program = "treatment")

diff_in_diff_controled_0910_1314 <- melt(results_control_0910_1314, measure = 1:2)
diff_in_diff_controled_0910_1314 <- mutate(diff_in_diff_controled_0910_1314, sig_program = "control")

diff_in_diff_boxplot <- rbind(diff_in_diff_treated_0910_1314, diff_in_diff_controled_0910_1314)
diff_in_diff_boxplot$value <- as.numeric(diff_in_diff_boxplot$value)
boxplot(value ~ variable+sig_program, data = diff_in_diff_boxplot,
        xlab = "School Year",
```

```r
        ylab = "Math Proficiency Rate", main = "Control Group and Treatment Group Boxplot")
### Create diff in diff dataset with both treated and controlled groups ###
diff_in_diff_treated <- mutate(diff_in_diff_treated, sig_program = "1")
diff_in_diff_controled <- mutate(diff_in_diff_controled, sig_program = "0")
diff_in_diff_both <- rbind(diff_in_diff_controled, diff_in_diff_treated)
ggplot(results_treated, aes(x=state)) +
  geom_bar()+
  theme(axis.text.x = element_text(angle = 60, hjust = 1))+
  labs(x = "States", y= "Number of schools")
# Line chart to show the trends
results_control[,6:10] <- sapply(results_control[,6:10], function(x) as.character(x))
results_control[,6:10] <- sapply(results_control[,6:10], function(x) as.numeric(x))
results_control_plot <- melt(results_control, measure = 6:10)
control_plot <- aggregate(results_control_plot$value, list(results_control_plot$variable), mean)
control_plot <- as.data.frame(control_plot)
control_plot$x <- as.numeric(control_plot$x)

results_treated[,6:10] <- sapply(results_treated[,6:10], function(x) as.character(x))
results_treated[,6:10] <- sapply(results_treated[,6:10], function(x) as.numeric(x))
results_treated_plot <- melt(results_treated, measure = 6:10)
treated_plot <- aggregate(results_treated_plot$value, list(results_treated_plot$variable), mean)
treated_plot <- as.data.frame(treated_plot)
treated_plot$x <- as.numeric(treated_plot$x)
# Basic line plot with points

plot<-ggplot() +
  geom_line(data=control_plot, aes(x=Group.1, y=x, group=1), color = "blue") +
  geom_line(data=treated_plot, aes(x=Group.1, y=x, group=1), color = "red") +
  xlab('School Years') +
  ylab('Math Proficiency Rate')
print(plot)
diff_in_diff_both_0910_1314 <- diff_in_diff_both %>%
  filter(cohort == "scores_0910" | cohort == "scores_1314" )
diff_in_diff_both_0910_1314 %>%
  ggplot(aes(x = cohort,
             y = percentage,
             color = sig_program
             )) +
  geom_point() +
  geom_line(aes(group = ncessch), alpha = 0.2) +
  theme_minimal() +
  labs(x = "School Years",
       y = "Math Proficiency Rate",
       color = "SIG pogram") +
  scale_color_brewer(palette = "Set1")
cohort1 <- filter(diff_in_diff_both, cohort %in% c("scores_0910","scores_1314"))
cohort1$cohort <- if_else(cohort1$cohort == "scores_0910", "0","1")
cohort1$sig_program <- as.factor(cohort1$sig_program)
cohort1$cohort <- as.factor(cohort1$cohort)
cohort1 <-
  rename(cohort1,
    year = cohort
    )
```

```r
diff_in_diff_cohort1_model <- lm(percentage ~  sig_program*year,
                        data = cohort1)
summ(diff_in_diff_cohort1_model, model.info = FALSE, model.fit = FALSE)
# Construct a second control group
results_control2 <- data.frame()
for (i in unique(results_treated$state)) {

  count_state <- count(filter(results_treated, state == i))

  filter_by_state <- which(merged_math_cleaned$stnam.x == i)

  set.seed(344)
  sample_by_state <- merged_math_cleaned[sample(filter_by_state,
                                        as.numeric(count_state)),]

  results_control2 <- rbind(results_control2, sample_by_state)

}
colnames(results_control2) <- c("ncessch", "school_name", "state", "lea_id",
                        "lea_name", "scores_0910", "scores_1011",
                        "scores_1112", "scores_1213", "scores_1314")
diff_in_diff_controled2 <- melt(results_control2, measure = 6:10)

diff_in_diff_controled2[,7] <-
  sapply(diff_in_diff_controled2[,7], function(x) as.numeric(x))
colnames(diff_in_diff_controled2) <- c("ncessch", "school_name", "state", "lea_id",
                                "lea_name", "cohort","percentage")
diff_in_diff_controled2 <- mutate(diff_in_diff_controled2, sig_program = 0)

diff_in_diff_both2 <- rbind(diff_in_diff_controled2, diff_in_diff_treated)

cohort2 <- filter(diff_in_diff_both2, cohort %in% c("scores_0910","scores_1314"))
cohort2$cohort <- if_else(cohort2$cohort == "scores_0910", "0","1")
cohort2$sig_program <- as.factor(cohort2$sig_program)
cohort2$cohort <- as.factor(cohort2$cohort)
cohort2 <-
  rename(cohort2,
        year = cohort
  )
diff_in_diff_cohort2_model <- lm(percentage ~  sig_program*year,
                            data = cohort2)
summ(diff_in_diff_cohort2_model, model.info = FALSE, model.fit = FALSE)
# Construct a fake treatment group
results_treatment_fake <- data.frame()
for (i in unique(results_treated$state)) {

    count_state <- count(filter(results_treated, state == i))

    filter_by_state <- which(merged_math_cleaned$stnam.x == i)

    set.seed(391)
    sample_by_state <- merged_math_cleaned[sample(filter_by_state,
                                        as.numeric(count_state)),]
```

```
    results_treatment_fake <- rbind(results_treatment_fake, sample_by_state)


    }
colnames(results_treatment_fake) <- c("ncessch", "school_name", "state", "lea_id",
                                "lea_name", "scores_0910", "scores_1011",
                                "scores_1112", "scores_1213", "scores_1314")
diff_in_diff_treated_fake <- melt(results_treatment_fake, measure = 6:10)
diff_in_diff_treated_fake[,7] <-
sapply(diff_in_diff_treated_fake[,7], function(x) as.numeric(x))
colnames(diff_in_diff_treated_fake) <- c("ncessch", "school_name", "state", "lea_id",
                                "lea_name", "cohort","percentage")
diff_in_diff_treated_fake <- mutate(diff_in_diff_treated_fake, sig_program = 1)

diff_in_diff_both_fake <- rbind(diff_in_diff_controled, diff_in_diff_treated_fake)

cohort_fake <- filter(diff_in_diff_both_fake, cohort %in% c("scores_0910","scores_1314"))
cohort_fake$cohort <- if_else(cohort_fake$cohort == "scores_0910", "0","1")
cohort_fake$sig_program <- as.factor(cohort_fake$sig_program)
cohort_fake$cohort <- as.factor(cohort_fake$cohort)
cohort_fake <-
  rename(cohort_fake,
    year = cohort
    )
diff_in_diff_cohort_fake_model <- lm(percentage ~  sig_program*year,
                        data = cohort_fake)
summ(diff_in_diff_cohort_fake_model, model.info = FALSE, model.fit = FALSE)
ggplot(results_control, aes(x=state)) +
  geom_bar()+
  theme(axis.text.x = element_text(angle = 60, hjust = 1))+
  labs(x = "States", y= "Number of schools")
```

References

David Hugh-Jones (2020). huxable: Easily Create and Style Tables for LaTeX, HTML and Other Formats. R package version 4.7.1. https://cran.r-project.org/web/packages/huxtable

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., & James-Burdumy, S. (2017, January 27). School Improvement Grants: Implementation and Effectiveness Executive Summary (Rep.). Retrieved March 19, 2020, from U.S. Department of Education website: https://ies.ed.gov/ncee/pubs/20174013/pdf/20174012.pdf

Hadley Wickham and Lionel Henry (2019). tidyr: Tidy Messy Data. R package version 1.0.0. https://CRAN.R-project.org/package=tidyr

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2018). dplyr: A Grammar of Data Manipulation. R package version 0.7.6. https://CRAN.R-project.org/package=dplyr

Kutash, J., Nico, E., Gorin, E., Rahmatullah, S., & Tallant, K. (2010). The school turnaround field guide. FSG Social Impact Advisors. https://www.wallacefoundation.org/knowledge-center/Documents/The-School-Turnaround-Field-Guide.pdf

Le Floch, K. C. (2018, January 22). Did school improvement grants work anywhere? American Institutes for Research. https://www.air.org/resource/did-school-improvement-grants-work-anywhere

Public Impact, EdPlex, & Michael & Susan Dell Foundation. (2020). What is restart and why do it? School Restarts. https://www.schoolrestarts.org/what-is-restart-and-why-do-it/

R Core Team (2019). R: A language and environment for statistical computing.

R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

The U.S. Department of Education. (2013). School improvement grants national summary school year 2012-13. https://www2.ed.gov/programs/sif/data/school-impro-grants-nat-sum-sy1213.pdf

The U.S. Department of Education. (2015). SIG transformation model – required and permissible activities. https://www.alsde.edu/ofc/osl/SIG/SIG%20Transformation%20Model.pdf

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Yihui Xie (2019). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.25.

Yihui Xie (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963