

DARE 1

Anwesha Guha & Merly Klaas

1/11/2022

A. Data Management Tasks (1 point)

For these tasks, no write up is required. The code you submit will be sufficient.

A1. Convert the raw counts of enrollment by race/ethnicity into percentages (i.e., divide the enrollment count for each ethno-racial category by total enrollment). For programming efficiency, can you use a function to do this task?

```
dare1 <- dare1 %>%  
  mutate(across(c(10:15), ~ . / !! dare1$enroll * 100))
```

A2. Generate dichotomous policy predictor variables that take the value of 1 in state-year observations in which the policy is in place. Call them eval, class remove and suspension. They should take the value of 0 in years during which these policies were not in place.

```
dare1 <- dare1 %>%  
  mutate(eval = case_when(eval_year >= school_year ~ 1,  
    TRUE ~ 0)) %>%  
  mutate(class_remove = case_when(class_remove_year >= school_year ~ 1,  
    TRUE ~ 0)) %>%  
  mutate(suspension = case_when(suspension_year >= school_year ~ 1,  
    TRUE ~ 0))
```

Also, generate a running time variable (run time) that reflects how far or close the state-year observation is from the implementation of higher stakes teacher evaluation and a variable that permits the effects of the evaluation policy to vary (linearly) over time (evalXyear). How will you deal with states that never implement evaluation? Do that too.

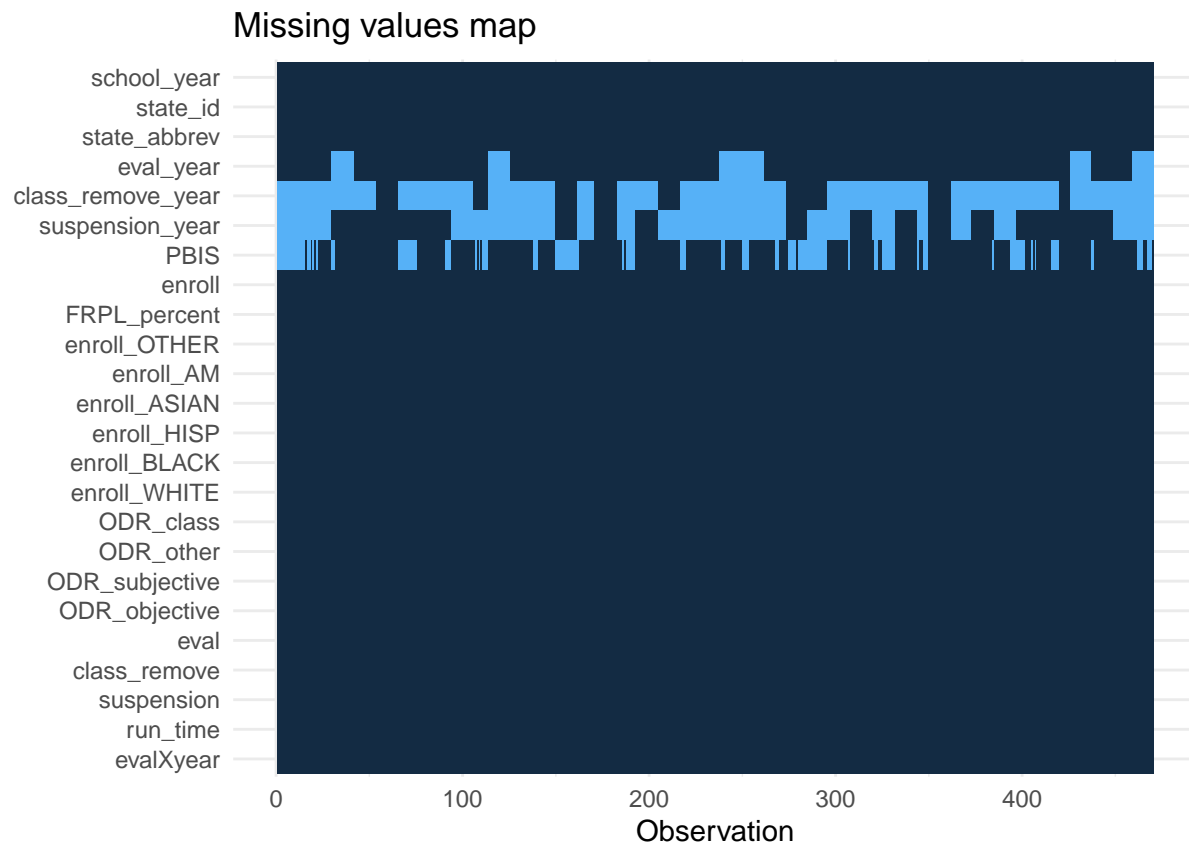
```
dare1 <- dare1 %>%  
  mutate(run_time = ifelse(is.na(eval_year), -99, school_year - eval_year)) %>%  
  # -99 for states that never implement evaluation  
  mutate(evalXyear = eval * run_time)
```

B. Understanding the Data and Descriptive Statistics (3 points)

For the following tasks, give your best attempt at completing the analysis and write-up. If you are unable to conduct the programming or analysis, describe what you are attempting to do and what your results would mean.

B1. Inspect your data. What sorts of missingness exist within the data file? What sorts of missingness should concern you? Which do not? In this assignment, please restrict your sample to state-years with non-missing outcomes.

```
dare1 %>%
  drop_na(ODR_class, ODR_objective, ODR_other, ODR_objective)%>%
  missing_plot()
```



```
dare1 %>% summary()
```

```
##   school_year      state_id      state_abbrev      eval_year
##   Min.   :2006      Min.    : 2.00      Length:516      Min.    :2011
##   1st Qu.:2009      1st Qu.:18.00      Class :character  1st Qu.:2013
##   Median :2012      Median :29.00      Mode  :character  Median :2014
##   Mean   :2012      Mean    :29.16                      Mean   :2014
##   3rd Qu.:2014      3rd Qu.:41.00                      3rd Qu.:2014
##   Max.   :2017      Max.    :56.00                      Max.    :2016
##                                     NA's    :72
##   class_remove_year suspension_year      PBIS      enroll
##   Min.   :2009      Min.    :2007      Min.    :0.0000      Min.    : 216
##   1st Qu.:2009      1st Qu.:2011      1st Qu.:0.0000      1st Qu.: 2891
##   Median :2012      Median :2014      Median :1.0000      Median : 9764
##   Mean   :2012      Mean    :2013      Mean    :0.7214      Mean   : 21897
##   3rd Qu.:2015      3rd Qu.:2016      3rd Qu.:1.0000      3rd Qu.: 26510
##   Max.   :2018      Max.    :2018      Max.    :1.0000      Max.    :207879
##   NA's    :408      NA's    :288      NA's    :175      NA's    :46
##   FRPL_percent      enroll_OTHER      enroll_AM      enroll_ASIAN
##   Min.   :0.07763      Min.    : 0.00000      Min.    : 0.0000      Min.    : 0.000
##   1st Qu.:0.44201      1st Qu.: 0.00000      1st Qu.: 0.3189      1st Qu.: 1.076
##   Median :0.53159      Median : 0.00000      Median : 0.5504      Median : 1.965
```

```
## Mean :0.54094 Mean : 0.32800 Mean : 3.1194 Mean : 3.091
## 3rd Qu.:0.62681 3rd Qu.: 0.00492 3rd Qu.: 1.2069 3rd Qu.: 3.826
## Max. :1.00000 Max. :20.81448 Max. :86.8996 Max. :17.611
## NA's :46 NA's :46 NA's :46 NA's :46
## enroll_HISP enroll_BLACK enroll_WHITE ODR_class
## Min. : 0.000 Min. : 0.000 Min. : 9.607 Min. :0.1612
## 1st Qu.: 3.760 1st Qu.: 2.860 1st Qu.: 47.575 1st Qu.:0.9673
## Median : 8.697 Median : 6.094 Median : 67.423 Median :1.4329
## Mean :13.744 Mean :11.663 Mean : 62.068 Mean :1.6872
## 3rd Qu.:18.147 3rd Qu.:18.487 3rd Qu.: 78.330 3rd Qu.:1.9747
## Max. :76.691 Max. :88.201 Max. :137.468 Max. :9.8629
## NA's :46 NA's :46 NA's :46 NA's :46
## ODR_other ODR_subjective ODR_objective eval
## Min. :0.1533 Min. :0.09597 Min. :0.04506 Min. :0.0000
## 1st Qu.:0.9565 1st Qu.:0.59837 1st Qu.:0.37276 1st Qu.:0.0000
## Median :1.4003 Median :0.89286 Median :0.52533 Median :1.0000
## Mean :1.5334 Mean :1.09670 Mean :0.60468 Mean :0.6124
## 3rd Qu.:1.8548 3rd Qu.:1.29252 3rd Qu.:0.76524 3rd Qu.:1.0000
## Max. :7.9305 Max. :6.84706 Max. :3.06346 Max. :1.0000
## NA's :46 NA's :46 NA's :46
## class_remove suspension run_time evalXyear
## Min. :0.000 Min. :0.0000 Min. : -99.00 Min. : -10.000
## 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.: -7.00 1st Qu.: -5.000
## Median :0.000 Median :0.0000 Median : -3.00 Median : -1.000
## Mean :0.126 Mean :0.2888 Mean : -15.57 Mean : -2.384
## 3rd Qu.:0.000 3rd Qu.:1.0000 3rd Qu.: 0.00 3rd Qu.: 0.000
## Max. :1.000 Max. :1.0000 Max. : 6.00 Max. : 0.000
##
```

```
dare1_clean <- dare1 %>%
  drop_na(ODR_class, ODR_objective, ODR_other, ODR_objective)
```

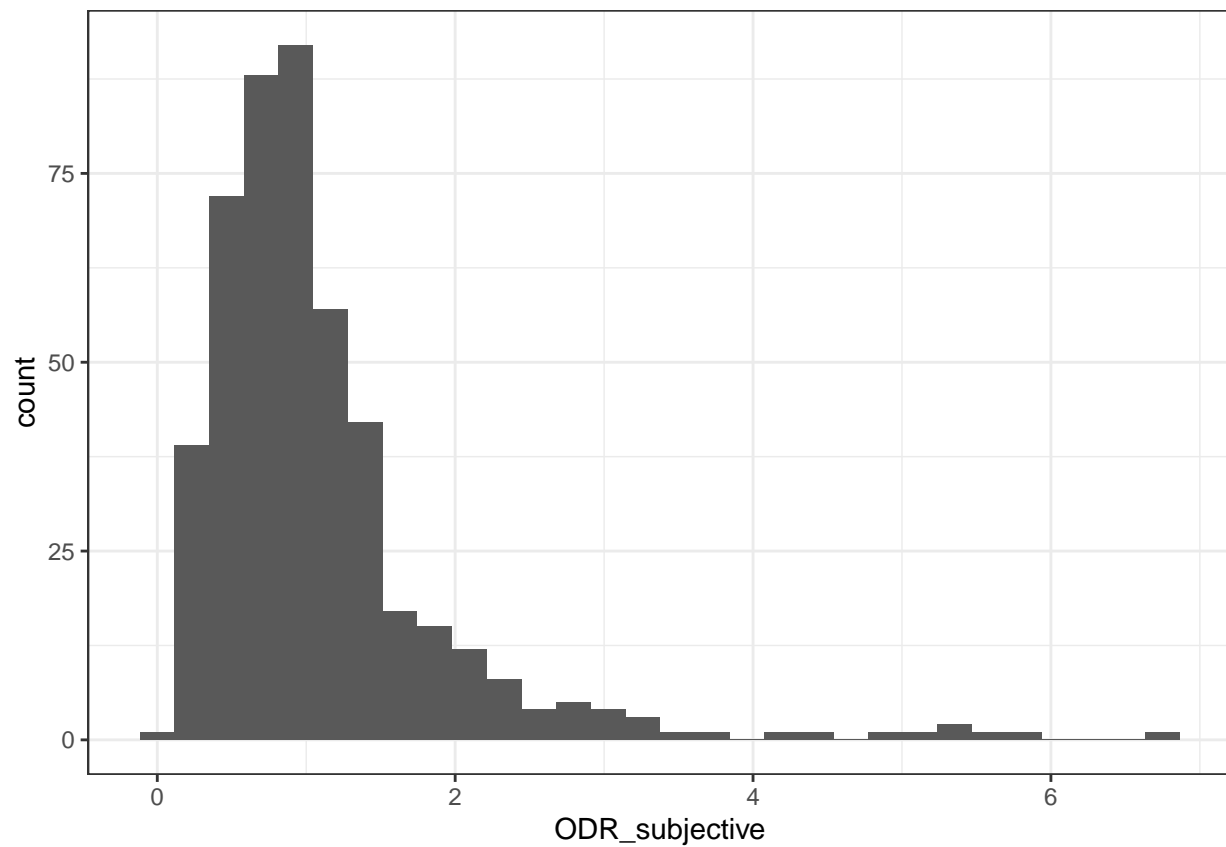
After excluding row with missing outcomes, there are only 470 observations left. Missing values found in these following variable: Var eval_year = 71, class_remove_year = 374, suspension_year = 259, PBIS = 129, based on the missingness pattern reflected in the plot above we see that missing data values do not relate to any other data in the dataset and there is no pattern to the actual values of the missing data themselves. Therefore we can conclude that this is Missing Completely at Random (MCAR). We should be concerned if there is specific pattern of the missingness.

B2. Graphically display the distribution of the outcome data. What do you notice about the distribution of outcomes? Are there any actions, transformations or sensitivity tests you would like to conduct based on this evidence?

```
# pivot_longer --> values to "ODR" could be another approach

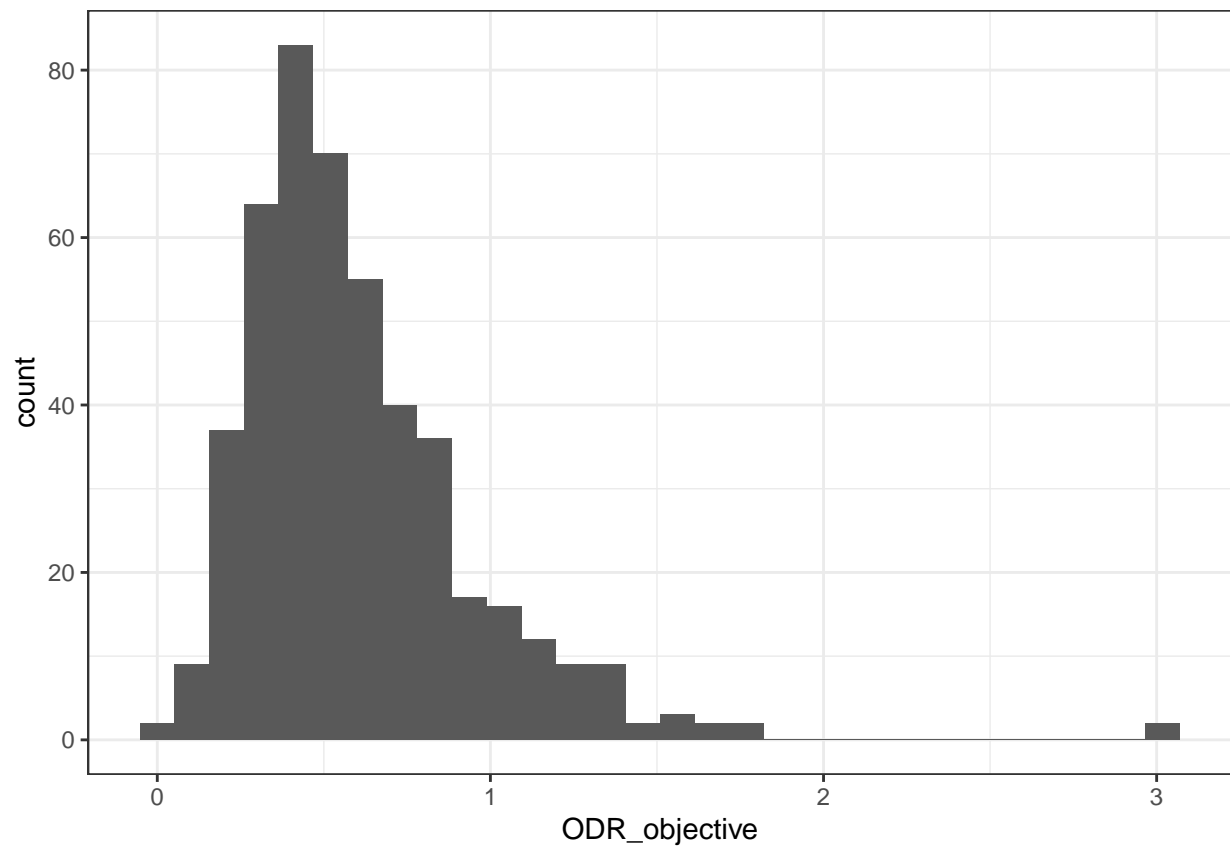
dare1_clean %>%
  ggplot(aes(ODR_subjective)) +
  geom_histogram() +
  theme_bw()
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



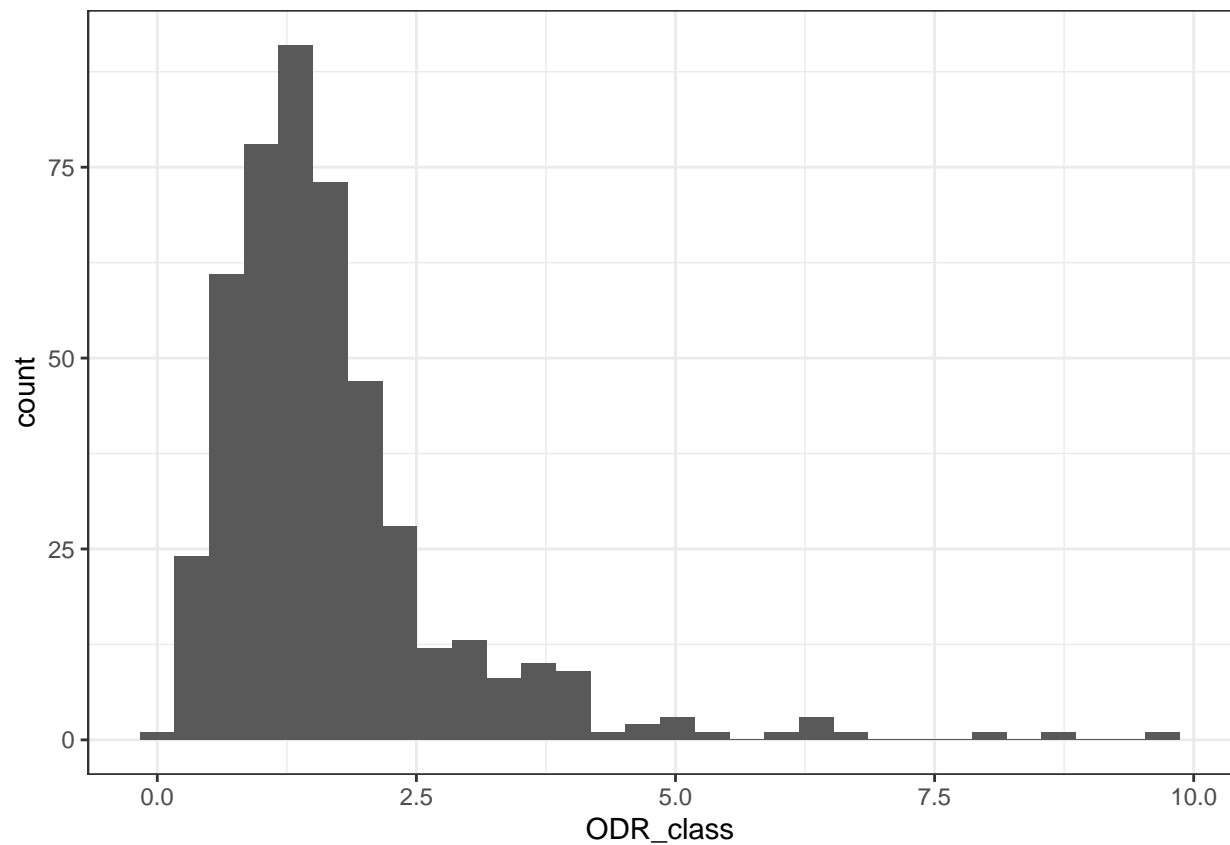
```
dare1_clean%>%  
  ggplot(aes(ODR_objective)) +  
  geom_histogram()+  
  theme_bw()
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



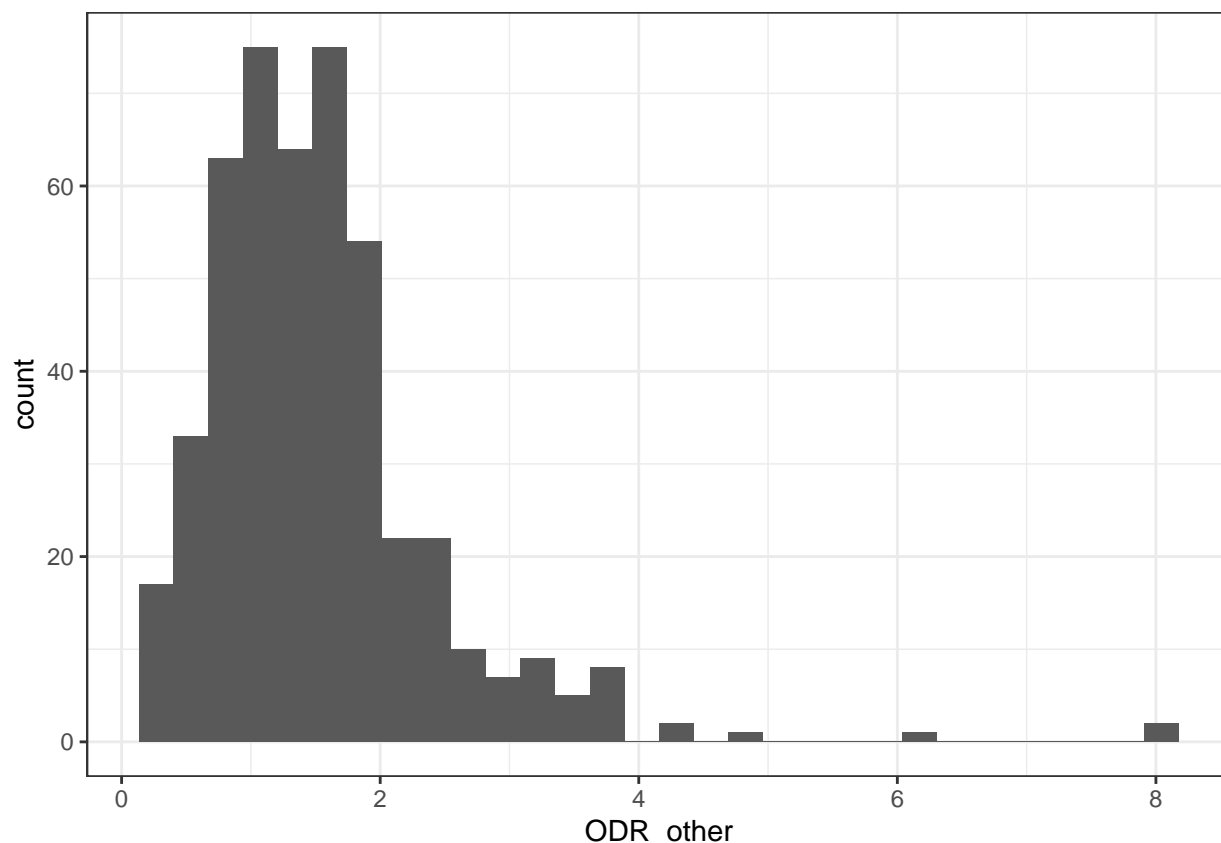
```
dare1_clean %>%  
  ggplot(aes(ODR_class)) +  
  geom_histogram()+  
  theme_bw()
```

'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.



```
dare1_clean %>%  
  ggplot(aes(ODR_other)) +  
  geom_histogram()+  
  theme_bw()
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



Each ODR outcome data is right-skewed. Based on these histograms, we would suggest 1) testing for normality (applying transformations as needed) and 2) removing outliers, particularly for ODR_objective and ODR_class.

B3. What is the analytic sample from which you will draw your inferences? To what population are you drawing these inferences? For this analytic sample, reproduce Column 1 of Table 1 from Liebowitz, Porter & Bragg (2022) to create a summary of descriptive statistics for the following data elements. All of these statistics (except for state-year and year enrollment) should be weighted by the state-year population:

- Mean state-year enrollment
- Mean year enrollment
- % low-income (FRPL)
- % Am. Indian/Alask. Native
- % Asian/PI
- % Black
- % Hispanic
- % White
- % state-year observations in which PBIS was successfully implemented
- Classroom ODR rate
- Other location ODR rate
- Subjective-Classroom ODR rate
- Objective-Classroom ODR rate

```
stargazer(attitude)
```

```
##
```

```
## % Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
```

```

## % Date and time: Mon, Jan 17, 2022 - 23:16:17
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccccc}
## \[-1.8ex]\hline
## \hline \[-1.8ex]
## Statistic & \multicolumn{1}{c}{N} & \multicolumn{1}{c}{Mean} & \multicolumn{1}{c}{St. Dev.} & \multicolumn{2}{c}{} \\
## \hline \[-1.8ex]
## rating & 30 & 64.633 & 12.173 & 40 & 58.8 & 71.8 & 85 \\
## complaints & 30 & 66.600 & 13.315 & 37 & 58.5 & 77 & 90 \\
## privileges & 30 & 53.133 & 12.235 & 30 & 45 & 62.5 & 83 \\
## learning & 30 & 56.367 & 11.737 & 34 & 47 & 66.8 & 75 \\
## raises & 30 & 64.633 & 10.397 & 43 & 58.2 & 71 & 88 \\
## critical & 30 & 74.767 & 9.895 & 49 & 69.2 & 80 & 92 \\
## advance & 30 & 42.933 & 10.289 & 25 & 35 & 47.8 & 72 \\
## \hline \[-1.8ex]
## \end{tabular}
## \end{table}

```

```
stargazer(dare1, header= FALSE, type = 'latex')
```

```

##
## \begin{table}[!htbp] \centering
##   \caption{}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lcccccc}
## \[-1.8ex]\hline
## \hline \[-1.8ex]
## Statistic & \multicolumn{1}{c}{N} & \multicolumn{1}{c}{Mean} & \multicolumn{1}{c}{St. Dev.} & \multicolumn{2}{c}{} \\
## \hline \[-1.8ex]
## school\_year & 516 & 2,011.500 & 3.455 & 2,006 & 2,008.8 & 2,014.2 & 2,017 \\
## state\_id & 516 & 29.163 & 14.760 & 2 & 18 & 41 & 56 \\
## eval\_year & 444 & 2,013.541 & 1.446 & 2,011.000 & 2,013.000 & 2,014.000 & 2,016.000 \\
## class\_remove\_year & 108 & 2,012.333 & 3.349 & 2,009.000 & 2,009.000 & 2,015.000 & 2,018.000 \\
## suspension\_year & 228 & 2,012.895 & 3.455 & 2,007.000 & 2,011.000 & 2,016.000 & 2,018.000 \\
## PBIS & 341 & 0.721 & 0.449 & 0.000 & 0.000 & 1.000 & 1.000 \\
## enroll & 470 & 21,897.410 & 32,425.870 & 216.000 & 2,890.750 & 26,509.500 & 207,879.000 \\
## FRPL\_percent & 470 & 0.541 & 0.150 & 0.078 & 0.442 & 0.627 & 1.000 \\
## enroll\_OTHER & 470 & 0.328 & 1.222 & 0.000 & 0.000 & 0.005 & 20.814 \\
## enroll\_AM & 470 & 3.119 & 8.959 & 0.000 & 0.319 & 1.207 & 86.900 \\
## enroll\_ASIAN & 470 & 3.091 & 3.195 & 0.000 & 1.076 & 3.826 & 17.611 \\
## enroll\_HISP & 470 & 13.744 & 14.047 & 0.000 & 3.760 & 18.147 & 76.691 \\
## enroll\_BLACK & 470 & 11.663 & 12.370 & 0.000 & 2.860 & 18.487 & 88.201 \\
## enroll\_WHITE & 470 & 62.068 & 21.141 & 9.607 & 47.575 & 78.330 & 137.468 \\
## ODR\_class & 470 & 1.687 & 1.182 & 0.161 & 0.967 & 1.975 & 9.863 \\
## ODR\_other & 470 & 1.533 & 0.905 & 0.153 & 0.956 & 1.855 & 7.931 \\
## ODR\_subjective & 470 & 1.097 & 0.857 & 0.096 & 0.598 & 1.293 & 6.847 \\
## ODR\_objective & 470 & 0.605 & 0.351 & 0.045 & 0.373 & 0.765 & 3.063 \\
## eval & 516 & 0.612 & 0.488 & 0 & 0 & 1 & 1 \\
## class\_remove & 516 & 0.126 & 0.332 & 0 & 0 & 0 & 1 \\
## suspension & 516 & 0.289 & 0.454 & 0 & 0 & 1 & 1 \\
## run\_time & 516 & $-15.570 & 33.808 & $-99 & $-7 & 0 & 6 \\
## evalXyear & 516 & $-2.384 & 2.813 & $-10 & $-5 & 0 & 0

```



```
## \hline \\[[-1.8ex]
## \end{tabular}
## \end{table}
```

```
#Mean State-Year Enrollment
dare1_clean %>%
  group_by(state_abbrev) %>% #use abbrev for readability
  summarise(mean_state = mean(enroll))
```

A tibble: 43 x 2

state_abbrev mean_state 1 AK 230. 2 AR 3859. 3 AZ 2983. 4 CA 111710. 5 CO 26989. 6 CT 26782. 7 FL 2165. 8 GA 21169. 9 IA 19158. 10 ID 2768 # ... with 33 more rows

```
#Mean Year Enrollment
dare1_clean %>%
  group_by(school_year) %>%
  summarise(mean_year = mean(enroll))
```

A tibble: 12 x 2

school_year mean_year 1 2006 10504. 2 2007 13881. 3 2008 16379. 4 2009 20562. 5 2010 22764. 6 2011 24405. 7 2012 25130. 8 2013 24339. 9 2014 26065. 10 2015 26949. 11 2016 23948. 12 2017 22309

```
#Mean State-Year Enrollment
dare1_clean %>%
  group_by(state_abbrev, school_year) %>%
  summarise(mean_state = mean(enroll))
```

'summarise()' has grouped output by 'state_abbrev'. You can override using the '.groups' argument.

A tibble: 470 x 3

Groups: state__abbrev [43]

state_abbrev school_year mean_state 1 AK 2010 229 2 AK 2011 224 3 AK 2012 221 4 AK 2013 255 5 AK 2014 216 6 AK 2016 229 7 AK 2017 234 8 AR 2006 782 9 AR 2007 3131 10 AR 2008 3491 # ... with 460 more rows

```
#Summary statistics for demographic information and outcome variables.

dare1_clean %>%
  select(
    `low-income (FRPL)` = FRPL_percent,
    `Am. Indian/Alask. Native` = enroll_AM,
    `Asian/PI` = enroll_ASIAN,
    `Black` = enroll_BLACK,
```

```

`% Hispanic` = enroll_HISP,
`% White` = enroll_WHITE,
`% Schools by Year Implementing PBIS` = PBIS,
`Classroom ODR Rate` = ODR_class,
`Other location ODR Rate` = ODR_other,
`Subjective-Classroom ODR rate` = ODR_subjective,
`Objective-Classroom ODR rate` = ODR_objective) %>% drop_na() %>%
tbl_summary(statistic = list(all_continuous() ~ "{mean} ({sd})") %>%
  modify_footnote(
    all_stat_cols() ~ "Mean (SD) per school"
  ) %>%
  modify_caption("**Table 1. Summary Statistics**")

```

```

## Table printed with 'knitr::kable()', not {gt}. Learn why at
## http://www.danielsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include 'message = FALSE' in code chunk header.

```

Table 1: Table 1. Summary Statistics

Characteristic	N = 341
% low-income (FRPL)	0.52 (0.13)
% Am. Indian/Alask. Native	1.44 (3.70)
% Asian/PI	3.4 (3.2)
% Black	11 (11)
% Hispanic	13 (12)
% White	65 (18)
% Schools by Year Implementing PBIS	246 (72%)
Classroom ODR Rate	1.65 (0.91)
Other location ODR Rate	1.55 (0.74)
Subjective-Classroom ODR rate	1.06 (0.63)
Objective-Classroom ODR rate	0.61 (0.30)

Describe the characteristics of your sample as you would report these statistics in an academic paper. How are the characteristics of the sample you will be using for this replication exercise different from the sample in Liebowitz, Porter & Bragg (2022)? How, if at all, do you anticipate this will affect your results?

This sample primarily differs at level of detail. This sample is at the year and state level while the Liebowitz, Porter & Bragg include school and grade level, as well. As a result, the characteristics of the above variables will look different: for example, one school may have much higher FRPL than another, but that variation may look different year to year.

B4. Optional Extension Plot the average classroom (ODR class) and classroom-subjective ODRs (ODR subjective) by how close the state/year observation is to the implementation of the teacher evaluation policy for the states that implemented evaluation reform. (Note: this is similar to Figure 2 in the original paper). What do you notice about the raw outcome data plotted against the secular trend? Are there any actions, transformations or sensitivity tests you would like to conduct based on this evidence? Why do we stress plotting these raw averages only for states that implemented evaluation reform? How would including these states alter the interpretation of this figure?

C. Replication and Extension (6 points)

For the following tasks, give your best attempt at completing the analysis and write-up. If you are unable to conduct the programming or analysis, describe what you are attempting to do and what your results would

mean.

C1. Estimate the effects of the introduction of higher-stakes teacher evaluation reforms on Office Disciplinary Referrals. In one of your models, assume that the effects are constant and in another relax this assumption to allow the effects to differ (linearly) over time. Present these difference-in-differences estimates in a table and the associated writeup as you would report these results in an academic paper. Do you notice any important differences in these results and those reported in the original paper? If so, how would you consider addressing them (it is not necessary at this point for you to actually conduct the analysis, just describe approaches you might take)?

For classroom ODRs: Assume effects are constant

```
library(fixest)
mod_class_constant <- feols(ODR_class ~ eval |
  state_id + school_year, #default clustering on state id
  data = dare1,
  weights = dare1$enroll)

## NOTE: 46 observations removed because of NA values (LHS: 46, Weights: 46).
```

```
summary(mod_class_constant)

## OLS estimation, Dep. Var.: ODR_class
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##      Estimate Std. Error  t value Pr(>|t|)
## eval 0.035771    0.060818  0.588161  0.55957
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 41.9      Adj. R2: 0.801764
##              Within R2: 0.001163
```

Allow effects to differ over time

```
mod_class_time <- feols(ODR_class ~ eval |
  state_id + school_year,
  data = dare1,
  vcov = ~school_year^state_id,
  weights = dare1$enroll)

## NOTE: 46 observations removed because of NA values (LHS: 46, Weights: 46).
```

```
summary(mod_class_time)

## OLS estimation, Dep. Var.: ODR_class
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (school_year^state_id)
##      Estimate Std. Error  t value Pr(>|t|)
## eval 0.035771    0.045199  0.791413  0.4291
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 41.9      Adj. R2: 0.801764
##               Within R2: 0.001163
```

For subjective ODRs:

Assume effects are constant

```
mod_subj_constant <- feols(ODR_subjective ~ eval |
  state_id + school_year,
  data = dare1,
  weights = dare1$enroll)
```

```
## NOTE: 46 observations removed because of NA values (LHS: 46, Weights: 46).
```

```
summary(mod_subj_constant)
```

```
## OLS estimation, Dep. Var.: ODR_subjective
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##      Estimate Std. Error t value Pr(>|t|)
## eval  0.03298   0.043522  0.757784  0.45281
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 29.3      Adj. R2: 0.791287
##               Within R2: 0.002023
```

Allow effects to differ over time

```
mod_subj_time <- feols(ODR_subjective ~ eval |
  state_id + school_year,
  data = dare1,
  vcov = ~school_year^state_id,
  weights = dare1$enroll)
```

```
## NOTE: 46 observations removed because of NA values (LHS: 46, Weights: 46).
```

```
summary(mod_subj_time)
```

```
## OLS estimation, Dep. Var.: ODR_subjective
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (school_year^state_id)
##      Estimate Std. Error t value Pr(>|t|)
## eval  0.03298   0.029894  1.10321  0.2705
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 29.3      Adj. R2: 0.791287
##               Within R2: 0.002023
```

Table 2: Table X. Effects of Teacher Eval Reforms on Discipline Referrals

	1	2	3	4
eval	0.036 (0.061)	0.036 (0.045)	0.033 (0.044)	0.033 (0.030)
Num.Obs.	470	470	470	470
R2	0.825	0.825	0.815	0.815

Notes: 1 - Class, Constant Effects; 2 - Class, Time Effects, 3 - Subj. Constant Effects, 4 - Subj. Time Effects

```
results <- list()
results[["1"]] <- mod_class_constant
results[["2"]] <- mod_class_time
results[["3"]] <- mod_subj_constant
results[["4"]] <- mod_subj_time

row <- tribble(~term,           ~'Class, Constant Effects', ~'Class, Time Effects', ~'Subj. Constant Effects', ~'Subj. Time Effects')
attr(row, 'position') <- c(7)

modelsummary(results,
  title = "Table X. Effects of Teacher Eval Reforms on Discipline Referrals",
  stars=T,
  estimate = "{estimate}{stars}",
  gof_omit= "Adj|Pseudo|Log|Within|AIC|BIC|FE|Std",
  fourparttable= T,
  notes = c("Notes: 1 - Class, Constant Effects; 2 - Class, Time Effects, 3 - Subj. Constant Effects, 4 - Subj. Time Effects"),
  type='pdf')
```

Warning: In version 0.8.0 of the 'modelsummary' package, the default significance markers produced by the 'p' argument are not working. This warning is displayed once per session.

Differences between the values in the table above and the paper can be come from controls not being accounted for in this model. Since the authors of the paper mention preferring a simpler model, a simpler model was used here. Secondly, these results are at the state-year level, which is a different sample than at the school-year level. As a result, differences are expected and can be remedied with more data (or a more advanced analysis).

C2. Liebowitz et al. (2022) conduct a broad set of robustness checks. For this DARE assignment, you will conduct two (2). First test whether the main results you present in Question C1 are robust to the introduction of potentially simultaneous discipline policy reforms. Present the table and associated write-up as you would report these results in an academic paper. Then select an additional robustness check (either from the paper or not) and present evidence on whether your findings are sensitive to this test.

The first set of robustness check is test the effect of evaluation policy implementation on rates on suspension from the Civil Right Data Collection (B11 - B13 of Figure 4)

For ODR Class

```
#Robustness check with CRDC (B11)
rc_b11 <- feols(ODR_class ~ suspension |
  state_id + school_year, #default clustering on state id
  data = dare1_clean,
```

```

weights = dare1_clean$enroll)
summary(rc_b11)

```

```

## OLS estimation, Dep. Var.: ODR_class
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##           Estimate Std. Error   t value Pr(>|t|)
## suspension -0.069408   0.083888 -0.827386  0.41269
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 41.9      Adj. R2: 0.802307
##              Within R2: 0.0039

```

```

#Robustness check with CRDC and Controls (B12)

```

```

rc_b12 <- feols(ODR_class ~ suspension + FRPL_percent + enroll_OTHER + enroll_AM + enroll_HISP + enroll_BLACK + enroll_WHITE,
                state_id + school_year,
                data = dare1_clean,
                weights = dare1_clean$enroll)
summary(rc_b12)

```

```

## OLS estimation, Dep. Var.: ODR_class
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##           Estimate Std. Error   t value Pr(>|t|)
## suspension  -0.053887   0.084717 -0.636092 0.5281680
## FRPL_percent -0.027840   0.395350 -0.070418 0.9441949
## enroll_OTHER -0.005647   0.021916 -0.257644 0.7979399
## enroll_AM     0.016950   0.024893  0.680926 0.4996557
## enroll_HISP  -0.000645   0.007778 -0.082870 0.9343485
## enroll_ASIAN -0.057915   0.023983 -2.414858 0.0201743 *
## enroll_BLACK  0.030102   0.010863  2.770956 0.0082916 **
## enroll_WHITE  0.005572   0.004948  1.125922 0.2665912
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 40.6      Adj. R2: 0.8105
##              Within R2: 0.061285

```

```

#Robustness check with CRDC, controls, and Time (B13)

```

```

rc_b13 <- feols(ODR_class ~ suspension + FRPL_percent + enroll_OTHER + enroll_AM + enroll_HISP + enroll_BLACK + enroll_WHITE,
                state_id + school_year,
                data = dare1_clean,
                weights = dare1_clean$enroll)
summary(rc_b13)

```

```

## OLS estimation, Dep. Var.: ODR_class
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)

```

```
##           Estimate Std. Error   t value Pr(>|t|)
## suspension    -0.049651   0.087273 -0.568921 0.5724397
## FRPL_percent   0.017980   0.388742  0.046251 0.9633297
## enroll_OTHER  -0.008510   0.021061 -0.404061 0.6882184
## enroll_AM      0.012607   0.028456  0.443042 0.6600089
## enroll_HISP   -0.003277   0.010160 -0.322551 0.7486359
## enroll_ASIAN  -0.054135   0.026296 -2.058643 0.0457663 *
## enroll_BLACK   0.029763   0.010876  2.736620 0.0090596 **
## enroll_WHITE   0.005682   0.004939  1.150313 0.2565225
## run_time      -0.015746   0.030840 -0.510578 0.6123197
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 40.6      Adj. R2: 0.810682
##                Within R2: 0.064489
```

For ODRs Subjective

```
#Robustness check with CRDC (B11)
rc_b11s <- feols(ODR_subjective ~ suspension |
                 state_id + school_year, #default clustering on state id
                 data = dare1_clean,
                 weights = dare1_clean$enroll)
summary(rc_b11s)
```

```
## OLS estimation, Dep. Var.: ODR_subjective
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##           Estimate Std. Error   t value Pr(>|t|)
## suspension -0.007453   0.058637 -0.127102 0.89947
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 29.3      Adj. R2: 0.790883
##                Within R2: 9.199e-5
```

```
#Robustness check with CRDC and Controls (B12)
rc_b12s <- feols(ODR_subjective ~ suspension + FRPL_percent + enroll_OTHER + enroll_AM + enroll_HISP + enroll_BLACK + enroll_WHITE |
                 state_id + school_year,
                 data = dare1_clean,
                 weights = dare1_clean$enroll)
summary(rc_b12s)
```

```
## OLS estimation, Dep. Var.: ODR_subjective
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##           Estimate Std. Error   t value Pr(>|t|)
## suspension    0.014078   0.063985  0.220019 0.8269224
## FRPL_percent  -0.007931   0.254774 -0.031131 0.9753124
## enroll_OTHER  -0.018570   0.017833 -1.041306 0.3036909
## enroll_AM     0.027864   0.013830  2.014800 0.0503569 .
## enroll_HISP   0.003294   0.005387  0.611401 0.5442294
```

#Robustness check with CRDC, controls, and Time (B13)

```
## OLS estimation, Dep. Var.: ODR_subjective
## Observations: 470
## Fixed-effects: state_id: 43, school_year: 12
## Standard-errors: Clustered (state_id)
##
##           Estimate Std. Error   t value Pr(>|t|)
## suspension    0.018022   0.064916   0.277617 0.7826681
## FRPL_percent   0.034728   0.245331   0.141557 0.8881069
## enroll_OTHER  -0.021236   0.017199  -1.234702 0.2238048
## enroll_AM      0.023820   0.015699   1.517306 0.1366811
## enroll_HISP    0.000843   0.006986   0.120623 0.9045651
## enroll_ASIAN  -0.036436   0.018615  -1.957318 0.0569776 .
## enroll_BLACK   0.023356   0.008332   2.803175 0.0076263 **
## enroll_WHITE   0.005326   0.003931   1.354890 0.1826960
## run_time      -0.014660   0.021005  -0.697938 0.4890622
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 28.1      Adj. R2: 0.804183
##                Within R2: 0.081735
```

#Other (B1)

OLS estimation, Dep. Var.: ODR_other Observations: 470 Fixed-effects: state_id: 43, school_year: 12
Standard-errors: Clustered (state_id) Estimate Std. Error t value Pr(>|t|) eval -0.04855 0.037355 -1.29969
0.2008 — Signif. codes: 0 ‘**0.001**’ 0.01 ‘0.05’ 0.1 ‘.’ 1 RMSE: 26.8 Adj. R2: 0.896358 Within R2:
0.005222

#Other and controls (B2)

16


```

      state_id + school_year,
      data = dare1_clean,
      weights = dare1_clean$enroll)
summary(rc_b2)

```

OLS estimation, Dep. Var.: ODR_other Observations: 470 Fixed-effects: state_id: 43, school_year: 12
 Standard-errors: Clustered (state_id) Estimate Std. Error t value Pr(>|t|)
 eval -0.038656 0.046233 -0.836121 0.407819
 FRPL_percent 0.061635 0.297257 0.207347 0.836741
 enroll_OTHER -0.017391 0.021072 -0.825316 0.413856
 enroll_AM 0.002127 0.012003 0.177202 0.860202
 enroll_HISP 0.003719 0.005812 0.639876 0.525728
 enroll_ASIAN -0.008396 0.021224 -0.395569 0.694425
 enroll_BLACK 0.016501 0.006219 2.653176 0.011208 *
 enroll_WHITE 0.004592 0.003668 1.251791 0.217573
 — Signif. codes: 0 ‘**0.001**’ ‘0.01’ ‘0.05’ ‘0.1’ ‘1’ RMSE: 26.4 Adj. R2: 0.89794 Within R2: 0.036929

```

#Other, controls, and time (B3)
rc_b3 <- rc_b2 <- feols(ODR_other ~ eval + FRPL_percent + enroll_OTHER + enroll_AM + enroll_AM + enroll_HISP,
      state_id + school_year,
      data = dare1_clean,
      weights = dare1_clean$enroll)
summary(rc_b3)

```

OLS estimation, Dep. Var.: ODR_other Observations: 470 Fixed-effects: state_id: 43, school_year: 12
 Standard-errors: Clustered (state_id) Estimate Std. Error t value Pr(>|t|)
 eval -0.046307 0.052382 -0.884024 0.381716
 FRPL_percent 0.071396 0.303894 0.234936 0.815400
 enroll_OTHER -0.018051 0.021328 -0.846318 0.402172
 enroll_AM 0.001342 0.012364 0.108553 0.914074
 enroll_HISP 0.003213 0.006392 0.502673 0.617820
 enroll_ASIAN -0.007552 0.022370 -0.337595 0.737350
 enroll_BLACK 0.016308 0.006515 2.503110 0.016286 *
 enroll_WHITE 0.004693 0.003711 1.264728 0.212942
 run_time -0.003883 0.016017 -0.242402 0.809649
 — Signif. codes: 0 ‘**0.001**’ ‘0.01’ ‘0.05’ ‘0.1’ ‘1’ RMSE: 26.4 Adj. R2: 0.897727 Within R2: 0.037286

```

#Objective (B4)
rc_b4 <- feols(ODR_objective ~ eval |
      state_id + school_year, #default clustering on state id
      data = dare1_clean,
      weights = dare1_clean$enroll)
summary(rc_b4)

```

OLS estimation, Dep. Var.: ODR_objective Observations: 470 Fixed-effects: state_id: 43, school_year: 12
 Standard-errors: Clustered (state_id) Estimate Std. Error t value Pr(>|t|) eval -0.010117 0.016461 -0.614613 0.54213 — Signif. codes: 0 ‘**0.001**’ ‘0.01’ ‘0.05’ ‘0.1’ ‘1’ RMSE: 10.7 Adj. R2: 0.888442
 Within R2: 0.00143

```

#Objective and controls (B5)
rc_b5 <- feols(ODR_objective ~ eval + FRPL_percent + enroll_OTHER + enroll_AM + enroll_AM + enroll_HISP,
      state_id + school_year,

```

```

      data = dare1_clean,
      weights = dare1_clean$enroll)
summary(rc_b5)

```

OLS estimation, Dep. Var.: ODR_objective Observations: 470 Fixed-effects: state_id: 43, school_year: 12
Standard-errors: Clustered (state_id) Estimate Std. Error t value Pr(>|t|)
eval -0.004918 0.018187 -0.270435 0.788150
FRPL_percent 0.036497 0.119482 0.305459 0.761526
enroll_OTHER -0.001820 0.006672 -0.272762 0.786373
enroll_AM -0.002449 0.006751 -0.362800 0.718574
enroll_HISP 0.000679 0.002207 0.307692 0.759837
enroll_ASIAN -0.001799 0.007370 -0.244080 0.808357
enroll_BLACK 0.005165 0.001944 2.656704 0.011108 *
enroll_WHITE 0.001255 0.001312 0.956997 0.344045
— Signif. codes: 0 ‘**0.001**’ 0.01 ‘0.05’ 0.1 ‘0.1’ 1 RMSE: 10.6 Adj. R2: 0.888813 Within R2: 0.02154

```

#Objective, controls, and time (B6)
rc_b6 <- rc_b2 <- feols(ODR_objective ~ eval + FRPL_percent + enroll_OTHER + enroll_AM + enroll_AM + enroll_WHITE,
      state_id + school_year,
      data = dare1_clean,
      weights = dare1_clean$enroll)
summary(rc_b6)

```

OLS estimation, Dep. Var.: ODR_objective Observations: 470 Fixed-effects: state_id: 43, school_year: 12
Standard-errors: Clustered (state_id) Estimate Std. Error t value Pr(>|t|)
eval -0.003323 0.018866 -0.176145 0.861027
FRPL_percent 0.034462 0.123933 0.278070 0.782323
enroll_OTHER -0.001682 0.006557 -0.256581 0.798755
enroll_AM -0.002286 0.006769 -0.337687 0.737281
enroll_HISP 0.000784 0.002453 0.319777 0.750723
enroll_ASIAN -0.001975 0.007935 -0.248855 0.804686
enroll_BLACK 0.005205 0.002015 2.583252 0.013359 *
enroll_WHITE 0.001234 0.001304 0.946253 0.349432
run_time 0.000809 0.005512 0.146841 0.883960
— Signif. codes: 0 ‘**0.001**’ 0.01 ‘0.05’ 0.1 ‘0.1’ 1 RMSE: 10.6 Adj. R2: 0.888551 Within R2: 0.021638

```

compare_performance(rc_b11, rc_b12,rc_b13, rc_b11s, rc_b12s, rc_b13s, rc_b1, rc_b2, rc_b3, rc_b4, rc_b5)
print_md()

```

```

## Warning: When comparing models, please note that probably not all models were fit from
## same data.

```

Table 3: Comparison of Model Performance Indices

Name	Model	R2	R2 (adj.)	RMSE	Sigma	AIC weights	BIC weights	Performance-Score
rc_b4	fixest	0.90	0.89	0.19	11.38	0.996	1.000	96.81%
rc_b5	fixest	0.90	0.89	0.19	11.36	0.002	< 0.001	63.94%
rc_b2	fixest	0.90	0.89	0.19	11.38	< 0.001	< 0.001	63.86%
rc_b6	fixest	0.90	0.89	0.19	11.38	< 0.001	< 0.001	63.86%
rc_b3	fixest	0.91	0.90	0.48	28.34	< 0.001	< 0.001	49.66%
rc_b1	fixest	0.91	0.90	0.48	28.53	< 0.001	< 0.001	48.69%

Name	Model	R2	R2 (adj.)	RMSE	Sigma	AIC weights	BIC weights	Performance-Score
rc_b13s	fixest	0.83	0.80	0.54	30.21	< 0.001	< 0.001	18.18%
rc_b12s	fixest	0.83	0.80	0.54	30.26	< 0.001	< 0.001	17.82%
rc_b11s	fixest	0.81	0.79	0.54	31.21	< 0.001	< 0.001	12.87%
rc_b13	fixest	0.84	0.81	0.75	43.61	< 0.001	< 0.001	7.21%
rc_b12	fixest	0.84	0.81	0.75	43.63	< 0.001	< 0.001	7.06%
rc_b11	fixest	0.83	0.80	0.75	44.56	< 0.001	< 0.001	3.53%

C3. Write a discussion paragraph in which you present the substantive conclusions of your results about the effects of the introduction of higher-stakes teacher evaluation on ODRs.

According to this analyses, there is insufficient evidence to suggest that relationship between the introduction of higher-stakes teacher evaluation reforms on Office Disciplinary Referrals is significant.

C4. Optional Extension Use an event-study approach to this difference-in-differences research design to estimate the effects of the introduction of higher-stakes teacher evaluation reforms on Office Disciplinary Referrals (ODRs). Present these findings in an event-study graph. Present the figure and associated write-up as you would report these results in an academic paper. Do you notice any important differences in these results and those reported in the original paper? If so, how would you consider addressing them (At this point, it is not necessary for you to actually conduct the analysis. Just describe approaches you might take.)?

C5. Optional Extension Use one (or more) approaches to present the extent to which the successful implementation of Positive Behavioral Intervention and Supports (PBIS) framework moderating the effects of the introduction of higher-stakes teacher evaluation policies. Present these difference-in-differences estimates and associated write-up as you would report these results in an academic paper. Do you notice any important differences in these results and those reported in the original paper? If so, how would you consider addressing them?