# Relationship Between School Staffing and Student Suspensions, Referrals, and Arrests

Manuel V. Cano[†,1], Errol Kaylor[†,1], Havisha Khurana[†,1], Merly Klass[†,1], & Cassie N. Malcom[†,1]

[1] University of Oregon

Author Note

Department of Educational Policy, Methodology, and Leadership. College of Education

[†] All authors contributed equally to this work.

Abstract

School staffing choices impact student outcomes. Prior research has explored the association between novice teachers and law enforcement officers on students' disciplinary actions. This study uses publicly available data on schools from the 2017-18 Civil Rights Data Collection (n = 97,632). We had two research questions that aimed to answer how school staffing affects school climate. To study the relationship between novice teachers and student rates of suspensions, we used a 3-level hierarchical linear model (HLM) with schools nested within districts and districts nested within states. We also used a multiple regression with one continuous and two categorical predictors to find the association between law enforcement officials and student arrests and referrals. In both cases, the relationship was statistically significant even after controlling for characteristics. We also noted that males and Black students receive more disciplinary actions. Our results indicate that new teachers may benefit from additional classroom management training. In addition, alternative methods of encouraging proper student behavior should be considered.

*Keywords:* novice teachers, suspension, law enforcement officers, arrests and referrals, school staffing

Relationship Between School Staffing and Student Suspensions, Referrals, and Arrests

# 1 Introduction

## 1.1 Prior Literature

Previous research has shown that school staffing choices can have an impact on student outcomes. In particular, novice teachers and law enforcement officials located on school campuses may impact negatively the number of disciplinary actions taken with students of various subgroups. Often, poor classroom management by novice teachers leads to them sending more students to the school office for a form of disciplinary action (e.g., referrals, suspensions, and even arrests) (Williams, Johnson, Dangerfield-Persky, & Mayakis, 2020). Less research has been done on how law enforcement officials impact student disciplinary outcomes, but there have been some studies that suggest there are both benefits (i.e., deterring student to student assaults) and costs (i.e., more students reported for “non-serious violent crimes”) (James & McCallion, 2013).

The relationship between student disciplinary outcomes and number of new teachers has several potential mechanisms. In addition to poor classroom management, schools with a higher level of new teachers may lack in senior leadership, giving new teachers fewer opportunities to improve their teaching. Our study also looks for a similar relationship between law enforcement presence and student discipline, where higher LEO presence affects how students act and teachers respond.

## 1.2 Research Questions

Our team developed two research questions in order to determine if school staffing choices are related to student disciplinary outcomes.

**RQ1. What is the relationship between school level rates of novice teachers and school level rates of student suspensions?**

a. Descriptively, how do rates of exclusionary discipline differ by gender and race/ethnicity?

**RQ2. What is the relationship between the number of full time equivalent (FTE) law enforcement officials (both law enforcement officers and security guards) and the total combined number of student referrals and arrests?**

The rest of this document summarizes the methods, results, and ends with a discussion of each research question.

# 2 Methods

## 2.1 Sample and Exclusion Criteria

This analysis was conducted using publicly available data from the 2017-18 civil rights data collection. Each biennium, schools across the United States are required to respond and report key information about their school across a range of topics. This study uses the data in the following areas:

1. School Support (count of enrolled novice teachers, sworn law enforcement officers (LEOs), and security guards)
2. Referrals & Arrests (count of referrals and arrests of students by subgroup, which includes gender and disability status)
3. Suspensions and Expulsions (count of suspensions and expulsions of students by subgroup)
4. Enrollment (count of enrolled students by subgroup)

5. School characteristics (type: magnet, alternate, charter, and special education)
6. Data sets are organized by state, school district (LEA), and school level

Table 1

*Summary Statistics*

| **Characteristic** | **N = 96,853** |
|---|---|
| White Students Enrollment (% | 51 (33) |
| Hispanic Students Enrollment (%) | 24 (28) |
| Black Students Enrollment (%) | 15 (23) |
| Asian Students Enrollment (%) | 3.8 (8.3) |
| Native American/Alaskan Students Enrollment (%) | 1.72 (8.50) |
| Hawaiian and Pacific Islander Students Enrollment (%) | 0.35 (2.27) |
| Two or more races Student Enrollment (%) | 3.8 (4.2) |
| Students with Disability Status Enrollment (%) | 15 (14) |
| Novice Teacher Enrolled (%) | 12 (14) |
| Law Enforcement Officials | 34 (26) |
| Suspension Rate | 53 (101) |
| Explusion Rate | 1 (6) |
| Number of Arrest | 1 (6) |
| Number of Referral | 2 (11) |

The sample in this study included all K-12 schools across the United States. Overall, the 2017-18 civil rights data collection contains 97632 schools who serve 50511682students in grades kindergarten (K) through 12th during the 2017-18 school year. For research question 1, schools that had missing enrollment information, such as missing counts of total enrollment or enrollment by race/ethnicity, were excluded from the analysis. Moreover, schools that did not report number of suspensions, referrals, arrests, number of novice teachers, were excluded from the analysis. For research question 1, the final sample included 95754 K-12 schools. For research question 2, school districts that had missing values for referrals and arrests had those values coded as NA and were excluded, leaving a final sample of 127,292 districts. Table 1 summarizes information on the study sample for both research questions.

We used R [Version 4.0.4; R Core Team (2021)] and the R-packages *broom.mixed* [Version 0.2.7; Bolker & Robinson (2021)], *car* [Version 3.0.12; Fox & Weisberg (2019); Fox, Weisberg, & Price (2020)], *carData* [Version 3.0.4; Fox et al. (2020)], *dplyr* [Version 1.0.7; Wickham, François, Henry, & Müller (2021)], *forcats* [Version 0.5.1; Wickham (2021a)], *ggplot2* [Version 3.3.5; Wickham (2016)], *ggridges* [Version 0.5.3; Wilke (2021)], *ggstance* [Version 0.3.5; Henry, Wickham, & Chang (2020)], *ggthemes* [Version 4.2.4; Arnold (2021)], *gtsummary* [Version 1.5.0; Sjoberg, Whiting, Curry, Lavery, & Larmarange (2021)], *here* [Version 1.0.1; Müller (2020)], *janitor* [Version 2.1.0; Firke (2021)], *jtools* [Version 2.1.4; Long (2020)], *lm.beta* [Version 1.5.1; Behrendt (2014)], *lme4* [Version 1.1.27.1; Bates, Mächler, Bolker, & Walker (2015)], *Matrix* [Version 1.3.2; Bates & Maechler (2021)], *papaja* [Version 0.1.0.9997; Aust & Barth (2020)], *performance* [Version 0.8.0; Lüdecke, Ben-Shachar, Patil, Waggoner, & Makowski (2021)], *psych* [Version 2.1.9; Revelle (2021)], *purrr* [Version 0.3.4; Henry & Wickham (2020)], *readr* [Version 2.1.1; Wickham & Hester (2021)], *reshape2* [Version 1.4.4; Wickham (2007)], *rio* [Version 0.5.29; Chan, Chan, Leeper, & Becker (2021)], *sjmisc* [Version 2.8.9; Lüdecke (2018)], *sjPlot* [Version 2.8.10; Lüdecke (2021)], *skimr* [Version 2.1.3; Waring et al. (2021)], *stringr* [Version 1.4.0; Wickham (2019)], *tibble* [Version 3.1.6; Müller & Wickham (2021)], *tidyr* [Version 1.1.4; Wickham (2021b)], and *tidyverse* [Version 1.3.1; Wickham et al. (2019)] for all our analyses.

## 2.2 Outcome Measures and Predictors of Interest

The main predictor variable for research question one was the percent of teachers that are novice teachers at a school. A teacher was considered novice if they were in their first or second year of teaching. This variable was calculated by dividing the total FTE of novice teachers at a school by the total FTE of full time teachers. Across schools in our sample, the mean rate of novice teachers was 12.02, and the numbers ranged from 0 to 100. There was wide variation in the rate of novice teachers within and across states. In our sample, Florida had the highest rate of novice teachers while North Carolina had the lowest rate (see Figure 1 in the appendix).

The main predictor of interest for research question two was the number of law enforcement officials (sworn law enforcement officers and security guards). This variable was calculated by summing the total FTE of sworn law enforcement officers and security guards to get a count.

The analysis had two outcomes of interest that broadly examined school climate, one for each research question. For research question 1, the main outcome of interest was the rate of students who received a suspension. More specifically, this variable was defined as the number of suspended students per 1,000 students at a school. Across schools in our sample, the mean rate of suspensions was 98.55, and the numbers ranged from 0 to 10500. There was wide variation in the rate of suspensions across and within states. Utah had the lowest rate of suspensions while South Carolina had the highest rate of suspensions (see Figure 2 in the appendix).

For research question 2, the outcome measure of interest was the total number of disciplinary actions. The total number of disciplinary actions was determined by adding the number of student referrals and arrests for each district.

### 2.3 Analytic Approach

Descriptive statistics and regression analyses were used to answer the two research questions. To answer the first research question, we used a 3-level hierarchical linear model (HLM) with schools nested within districts and districts nested within states. The final model is represented by the following sets of equations.

$$Y_{jks} = \pi_{0ks} + X_j + \pi_{1ks} * noviceteachers + \varepsilon_{jks}$$

$$\pi_{ks} = \beta_{00s} + r_{0jk}$$

$$\pi_{1ks} = \beta_{01s} + r_{1jk}$$

$$\beta_{00s} = \gamma_{000} + \mu_{00k}$$

$$\beta_{01s} = \gamma_{100}$$

Where the outcome (Y) of school (j) in district (k) in state (s) is predicted by a set of school level enrollment covariates (X_j) and a school level rate of novice teachers. This model allows the intercept to vary by district and school and also includes a random slope for the rate of novice teachers across districts. The y100 coefficient represents the relationship between school level rate of novice teachers and study outcomes and addresses the first research question. This model was selected through a model comparison test that compared this 3 level HLM to other model specifications and was found to be the best performing model (based on the AIC score).

To answer the second research question, we used a multiple regression with one continuous and two categorical predictors. The continuous predictor was total number of law enforcement officials and the categorical variables included gender and disability status. Gender was coded so that **female** was 0 and disability status was coded so that students **with** a disability were 0.

# 3 Results

## 3.1 Research Question 1 - Rates of Novice Teachers and Suspensions

Rates of novice teachers was positively and statistically significantly associated with higher rates of suspensions. Across schools in the United States, the model estimated average rate of suspension was 104 students per 1000 students at typical schools (e.g. not an alternative, charter, or magnet school) with average rates of total enrollment and average rates of enrollment by race/ethnicity, EL students, and special education enrollment. A one percentage point increase in the rate of novice teachers increased the number of students suspended per 1000 students by 1. Similar estimates were found across model specifications including a simpler HLM model with no added covariates and a 3 level random intercept and slopes model with interactions between rates of novice teachers and rate of black and special education students.

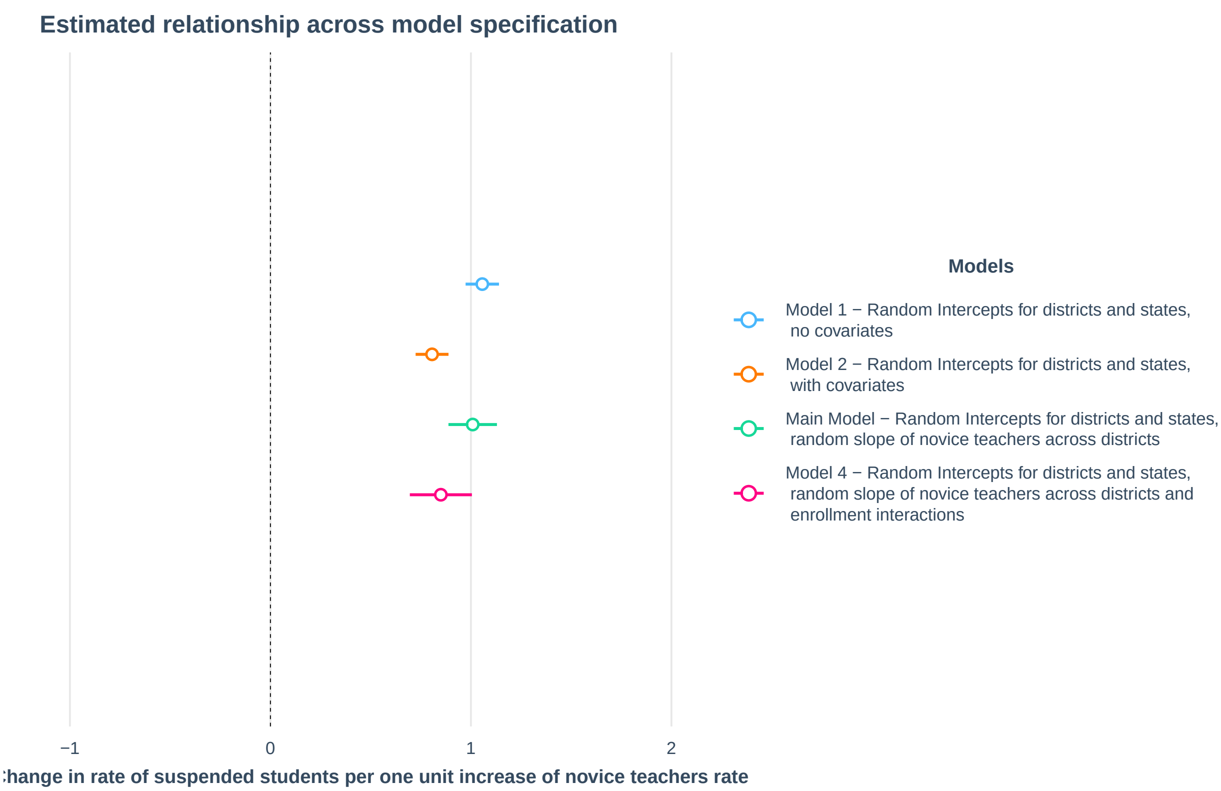


*Figure 1*

## 3.2 Research Question 1 - Descriptive Results

Figure 2 and figure 3 about Suspensions rate by Ethnicity in Preschool and K-12 show that **Black Students** were suspended more compared to their peers in Preschool and K-12.

In Figure 4 and Figure 5 when gender was included, both **Black Male** and **Black Female** students were suspended more compared to their peers in Preschool and K-12. Overall, **Male** students were suspended more than female students across all school level.

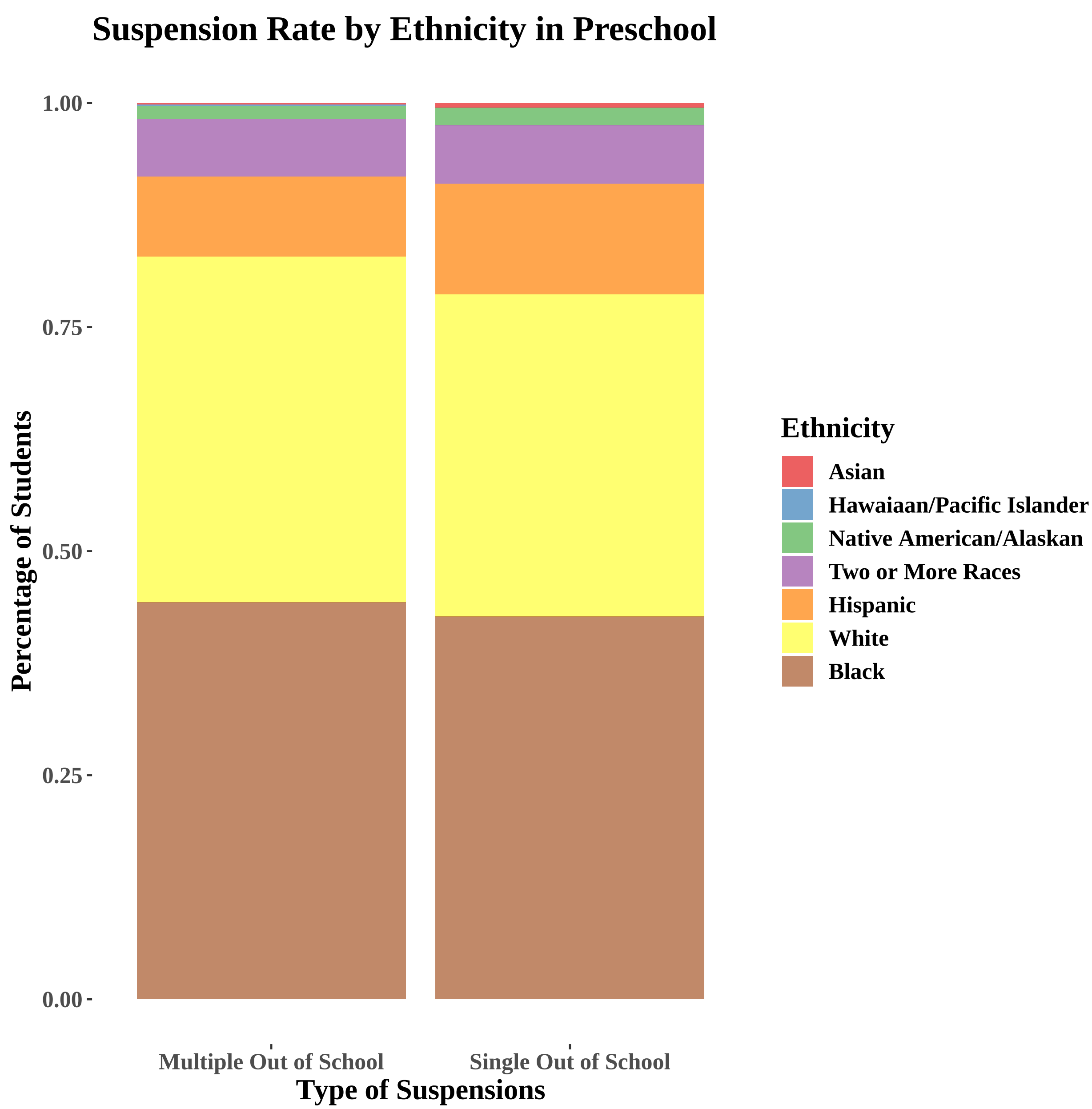


*Figure 2*

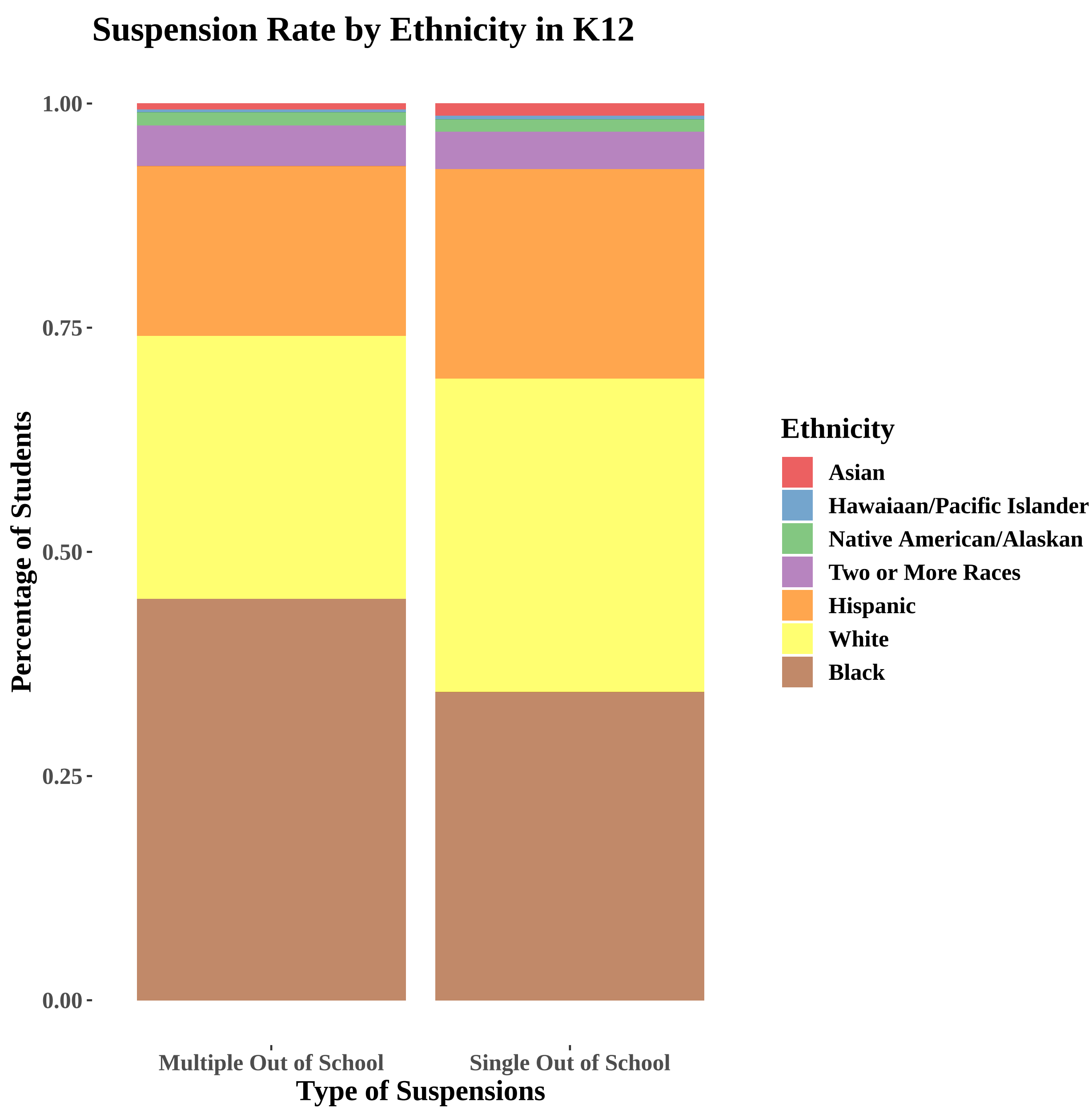
Suspension Rate by Ethnicity in K12
1.00
0.75
0.50
0.25
0.00
Percentage of Students
Multiple Out of School
Single Out of School
Type of Suspensions
Ethnicity
Asian
Hawaiaan/Pacific Islander
Native American/Alaskan
Two or More Races
Hispanic
White
Black

*Figure 3*

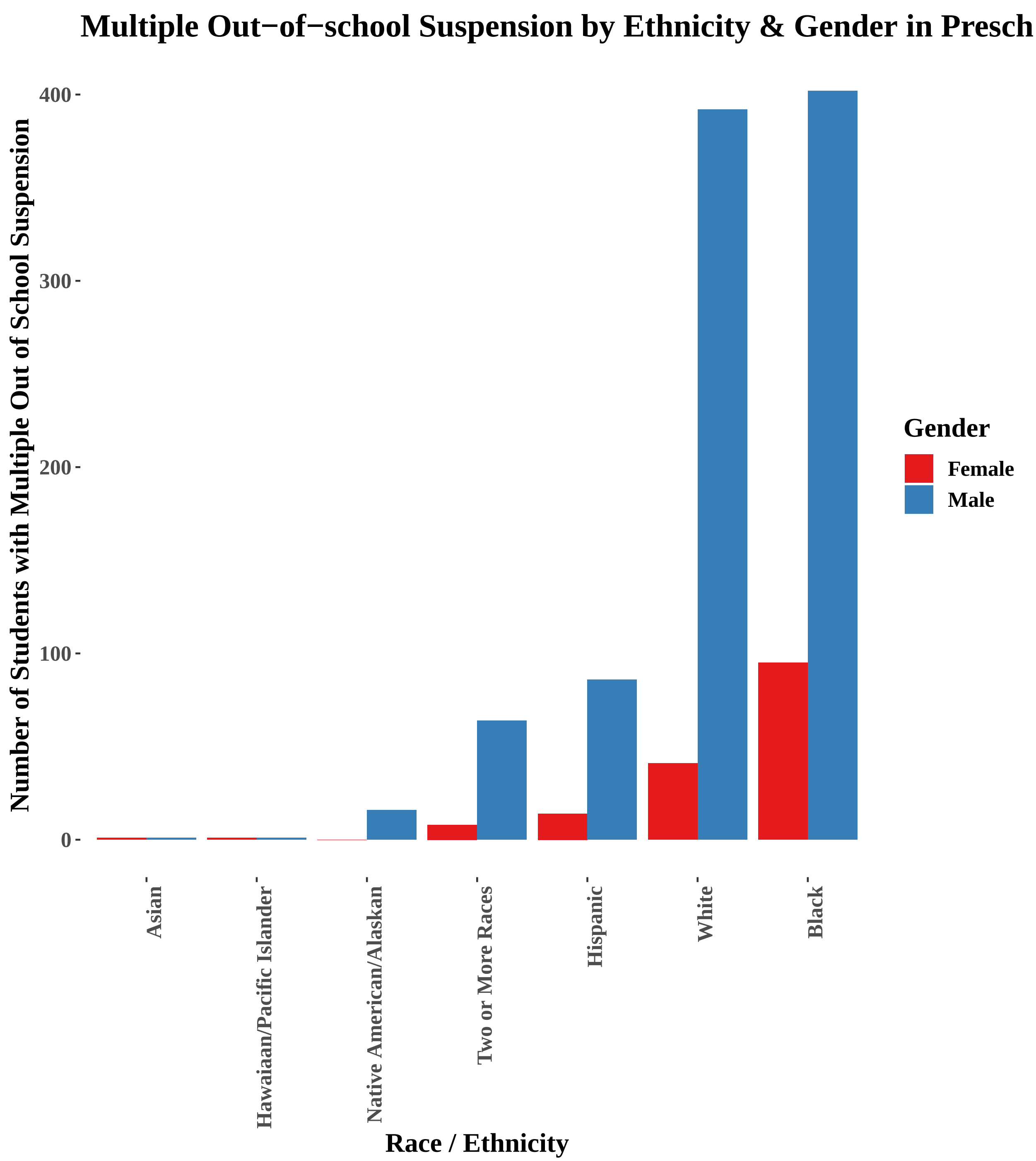
Multiple Out−of−school Suspension by Ethnicity & Gender in Presch
Number of Students with Multiple Out of School Suspension
400
300
200
100
0
Gender
Female
Male
Asian
Hawaiaan/Pacific Islander
Native American/Alaskan
Two or More Races
Hispanic
White
Black
Race / Ethnicity

*Figure 4*

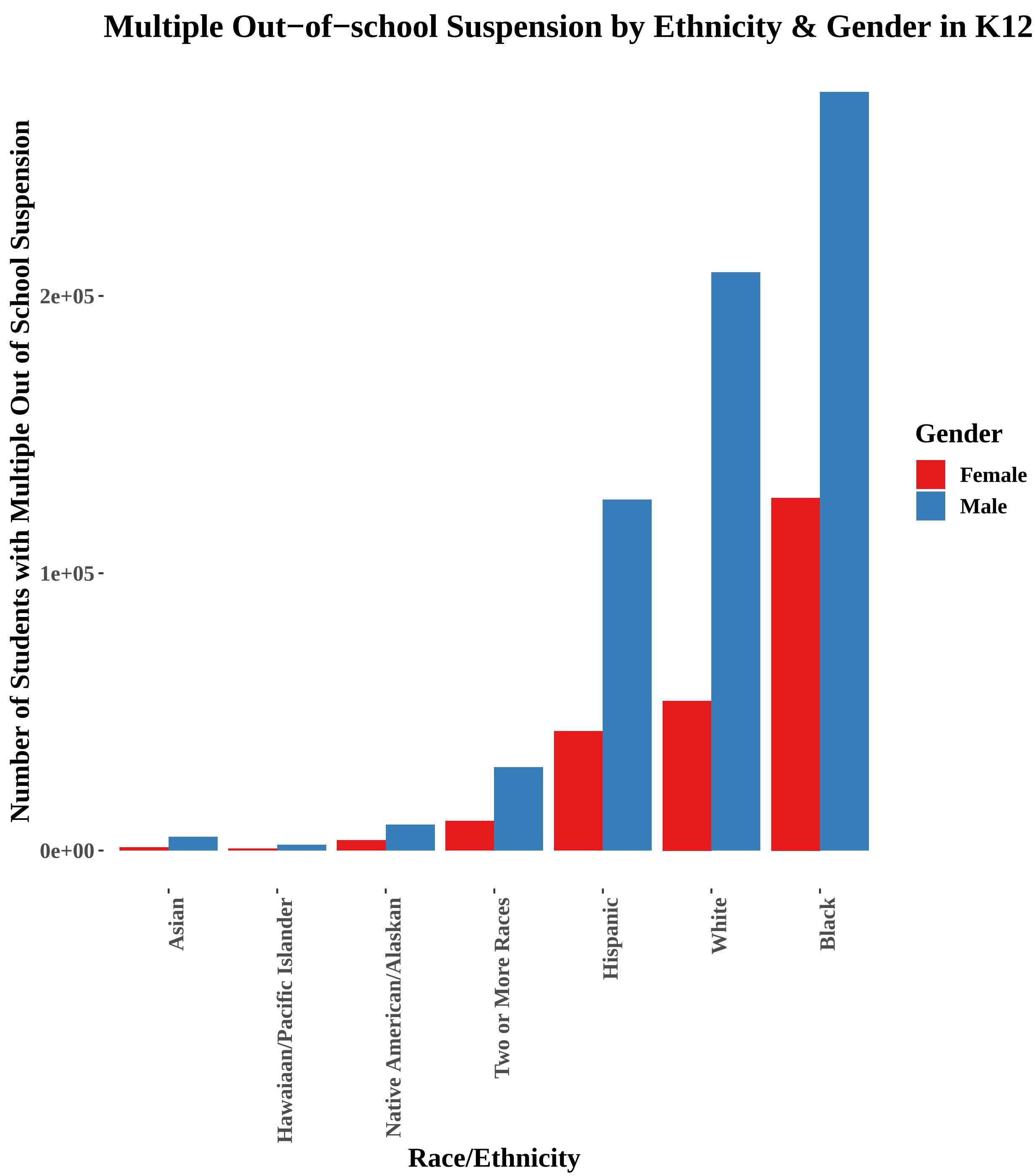
Multiple Out−of−school Suspension by Ethnicity & Gender in K12
Number of Students with Multiple Out of School Suspension
2e+05
1e+05
0e+00
Asian
Hawaiaan/Pacific Islander
Native American/Alaskan
Two or More Races
Hispanic
White
Black
Race/Ethnicity
Gender
Female
Male

*Figure 5*

## 3.3 Research Question 2 - Descriptive Results and Analysis

Figure 6 displays the total number of student arrests and referrals per state broken down by student gender.

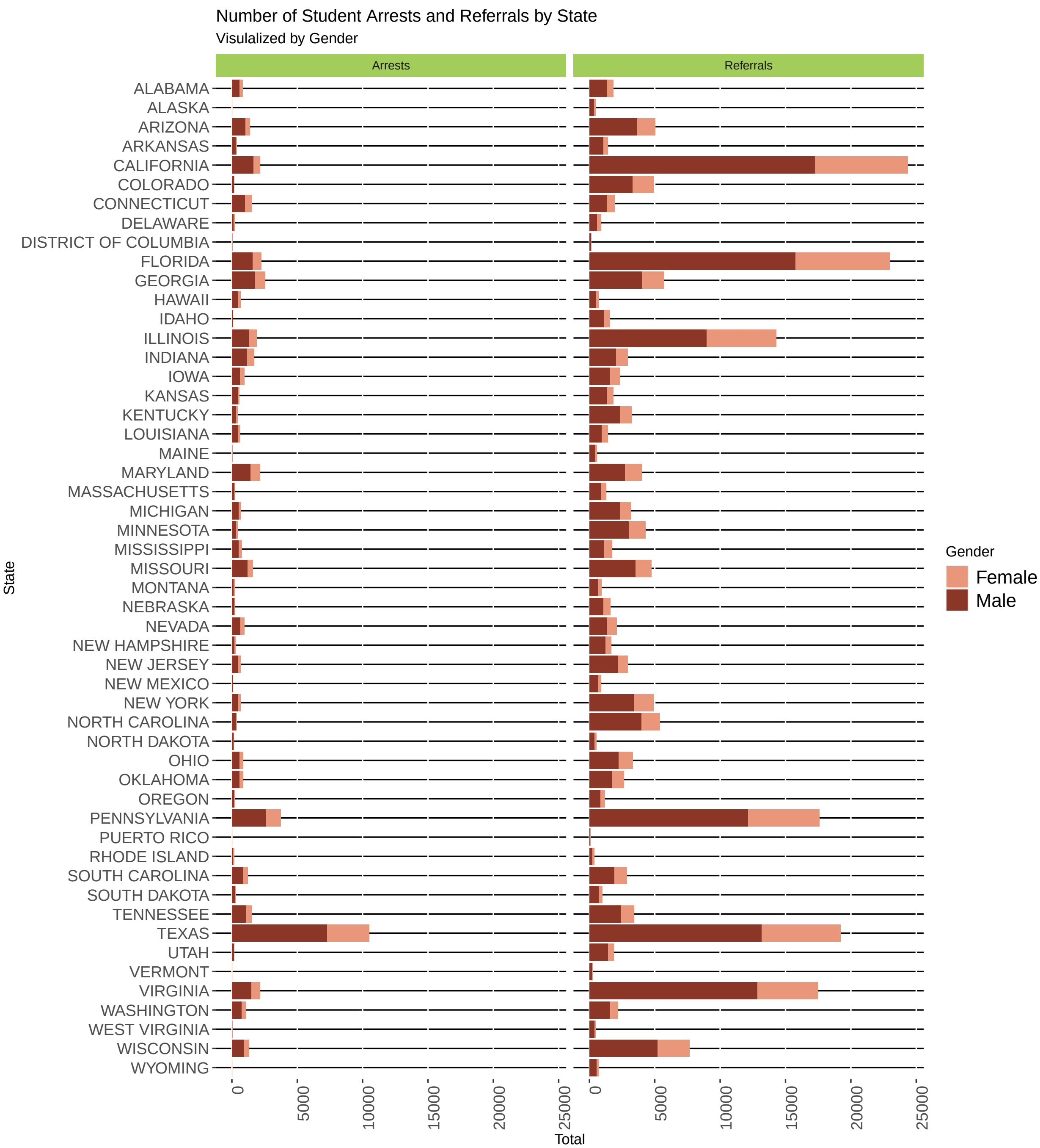


*Figure 6*

Figure 7 displays the total number of student arrests and referrals per state broken down by student disability status.

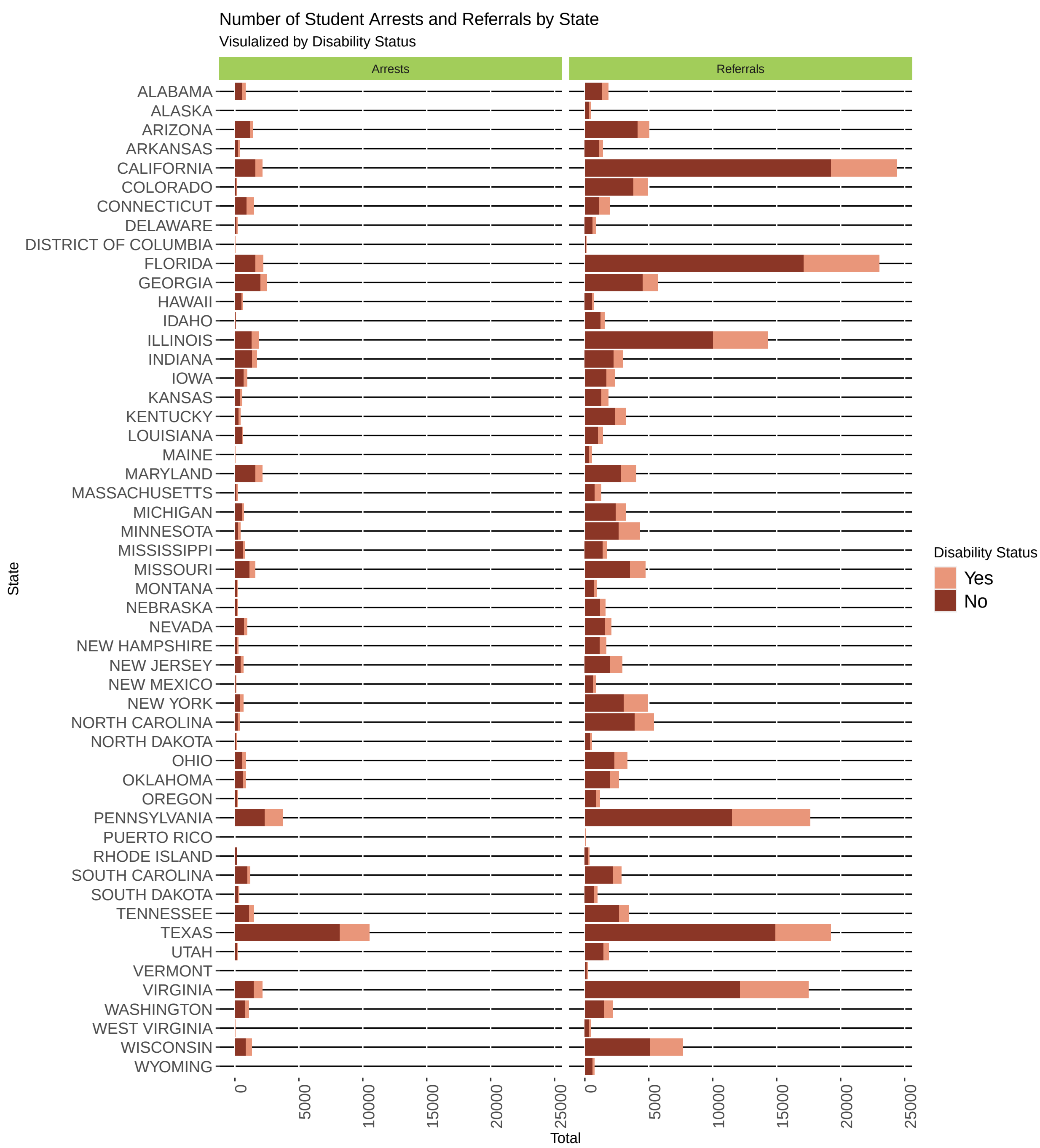


*Figure 7*

As shown in Table 2, for every one unit increase in FTE law enforcement officials the

Table 2

*Regression results*

| predictor | $b$ | 95% CI | $t(127288)$ | $p$ |
|---|---|---|---|---|
| (Intercept) | -0.34 | $[-0.47, -0.21]$ | -5.00 | $< .001$ |
| Total Law | 0.23 | [0.23, 0.24] | 63.83 | $< .001$ |
| Gender (M) | 1.01 | [0.86, 1.16] | 12.92 | $< .001$ |
| Status (wodis) | 1.14 | [0.98, 1.29] | 14.57 | $< .001$ |

number of student disciplinary actions increase by 0.23. After adjusting for student gender and disability status the number of students with disciplinary actions between male and female students remained significant, ($p<0.001$). The model indicated that the male students had 1.01 more disciplinary actions. The model also indicated that after adjusting for student gender and disability status the number of students with disciplinary actions between students with and without a disability remained significant, ($p<0.001$). This model explained 3.4% of the variance.

## 4 Discussion

The present study aimed to identify relationships between novice teachers and student suspension rates, and law enforcement official concentration and referrals and arrests. In both cases, our analysis finds significant correlations between staffing choices and student disciplinary outcomes. These correlations are still significant after controlling for school demographics and school status as well.

### 4.1 New Teachers and Suspensions

Our analysis shows a meaningful relationship between presence of new teachers and student discipline, where more new teachers correlate with more disciplinary actions. This supports our hypothesis that at schools with higher percentages of novice teachers, student disciplinary rates tend to be higher. Our results indicate that new teachers may benefit from additional classroom management training, or for administrators to schedule classroom management trainings when large numbers of new teachers are hired at the same time.

### 4.2 Law Enforcement Officials and Expulsions/Referrals

The significant relationship between count of security presence at a school and number of arrests and referrals that occur at a given school should encourage administrators to investigate alternative methods to reducing school violence and ensuring safety of the students.

### 4.3 Limitations and Future Research

The current study has some shortcomings in regards to applicability of our findings - the school level data does not give a total count of suspensions, only the number of students who had suspensions during the study time. As this data does not break down to specific classrooms, we cannot validate whether more suspensions are happening in classrooms with new teachers specifically. Additionally, our analyses do not examine LEO presence as a proportion of total school staff, but rather as a simple count. Additionally, our analyses includes Juvenile Justice facilities, which tend to have much higher concentrations of law enforcement officers than traditional public schools. A future analysis would benefit from anaylyzing these two groups separately, as well as investigating the cross-section of schools with high LEO presence and novice teacher percent. As our LEO analysis does not look at proportional data, further analysis there is necessary to understand the true significance of LEO and security guard presence on student disciplinary outcomes.

A valuable follow up to this study would be to further contextualize the teacher and LEO employment data with overall school size, funding, and other available data. These analyses would help determine whether student disciplinary outcomes are rising as a result of an unusual or misallocation of school resources, or due to the mechanisms discussed earlier.

## 5 References

Arnold, J. B. (2021). *Ggthemes: Extra themes, scales and geoms for 'ggplot2'.* Retrieved from https://CRAN.R-project.org/package=ggthemes

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown.* Retrieved from https://github.com/crsh/papaja

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., & Maechler, M. (2021). *Matrix: Sparse and dense matrix classes and methods.* Retrieved from https://CRAN.R-project.org/package=Matrix

Behrendt, S. (2014). *Lm.beta: Add standardized regression coefficients to lm-objects.* Retrieved from https://CRAN.R-project.org/package=lm.beta

Bolker, B., & Robinson, D. (2021). *Broom.mixed: Tidying methods for mixed models.* Retrieved from https://CRAN.R-project.org/package=broom.mixed

Chan, C., Chan, G. C., Leeper, T. J., & Becker, J. (2021). *Rio: A swiss-army knife for data file i/o.*

Firke, S. (2021). *Janitor: Simple tools for examining and cleaning dirty data.* Retrieved from https://CRAN.R-project.org/package=janitor

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (Third). Thousand Oaks CA: Sage. Retrieved from https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fox, J., Weisberg, S., & Price, B. (2020). *carData: Companion to applied regression data sets.* Retrieved from https://CRAN.R-project.org/package=carData

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools.* Retrieved from https://CRAN.R-project.org/package=purrr

Henry, L., Wickham, H., & Chang, W. (2020). *Ggstance: Horizontal 'ggplot2' components.* Retrieved from https://CRAN.R-project.org/package=ggstance

James, N., & McCallion, G. (2013). School resource officers: Law enforcement officers in schools. Congressional Research Service Washington, DC.

Long, J. A. (2020). *Jtools: Analysis and presentation of social scientific data.* Retrieved from https://cran.r-project.org/package=jtools

Lüdecke, D. (2018). Sjmisc: Data and variable transformation functions. *Journal of Open Source Software*, *3*(26), 754. https://doi.org/10.21105/joss.00754

Lüdecke, D. (2021). *sjPlot: Data visualization for statistics in social science.* Retrieved from https://CRAN.R-project.org/package=sjPlot

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Müller, K. (2020). *Here: A simpler way to find your files.* Retrieved from https://CRAN.R-project.org/package=here

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames.* Retrieved from https://CRAN.R-project.org/package=tibble

R Core Team. (2021). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research.* Evanston, Illinois: Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Sjoberg, D. D., Whiting, K., Curry, M., Lavery, J. A., & Larmarange, J. (2021). Reproducible summary tables with the gtsummary package. *The R Journal*, *13*, 570–580. https://doi.org/10.32614/RJ-2021-053

Waring, E., Quinn, M., McNamara, A., Arino de la Rubia, E., Zhu, H., & Ellis, S. (2021). *Skimr: Compact and flexible summaries of data.* Retrieved from https://CRAN.R-project.org/package=skimr

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12), 1–20. Retrieved from http://www.jstatsoft.org/v21/i12/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data*. Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Hester, J. (2021). *Readr: Read rectangular text data*. Retrieved from https://CRAN.R-project.org/package=readr

Wilke, C. O. (2021). *Ggridges: Ridgeline plots in 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggridges

Williams, J. A., Johnson, J. N., Dangerfield-Persky, F., & Mayakis, C. G. (2020). Does employing more novice teachers predict higher suspensions for black students? A hierarchical multiple regression analysis. *The Journal of Negro Education*, *89*(4), 448–458.

# Appendix

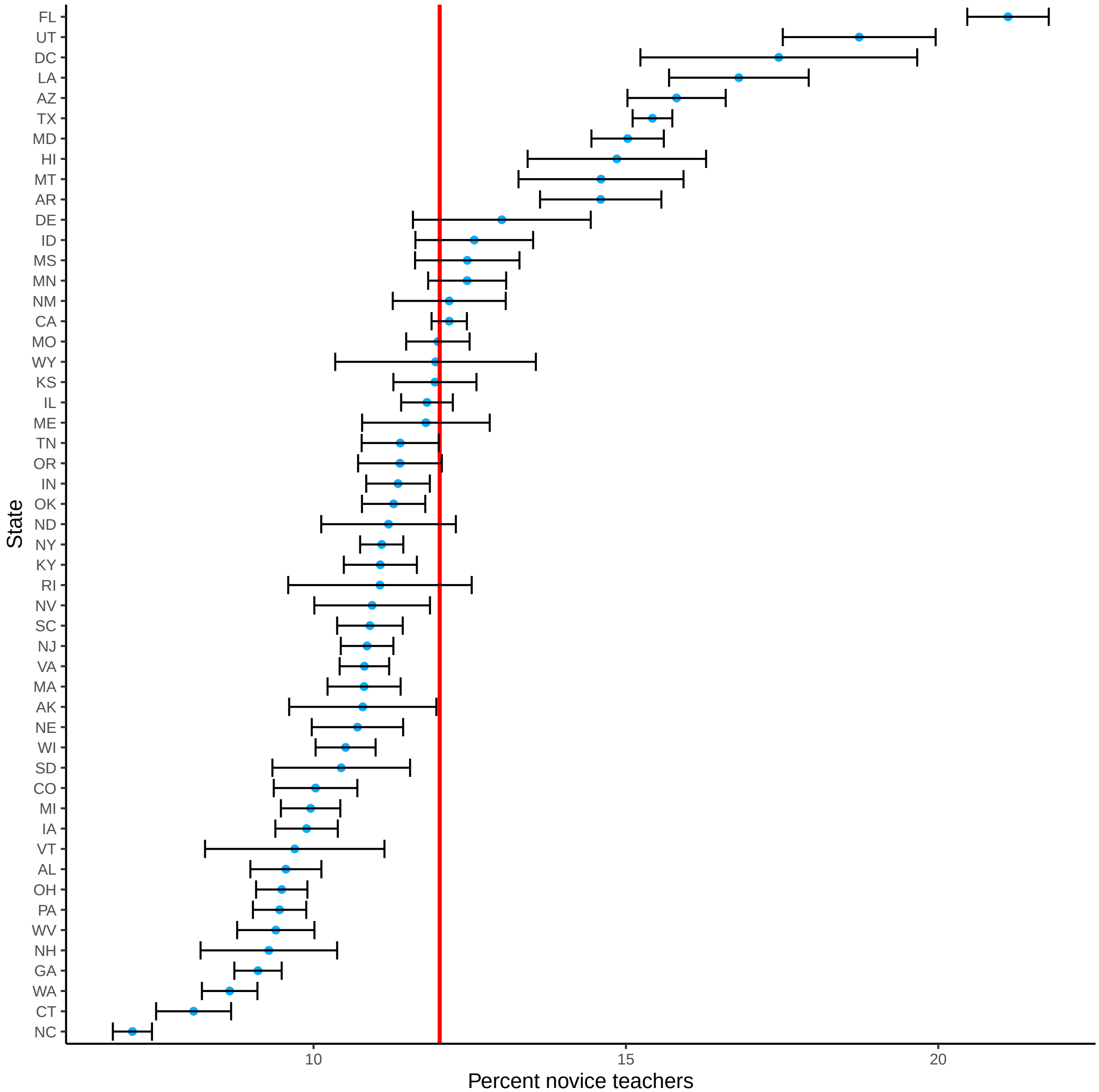


*Figure 1*

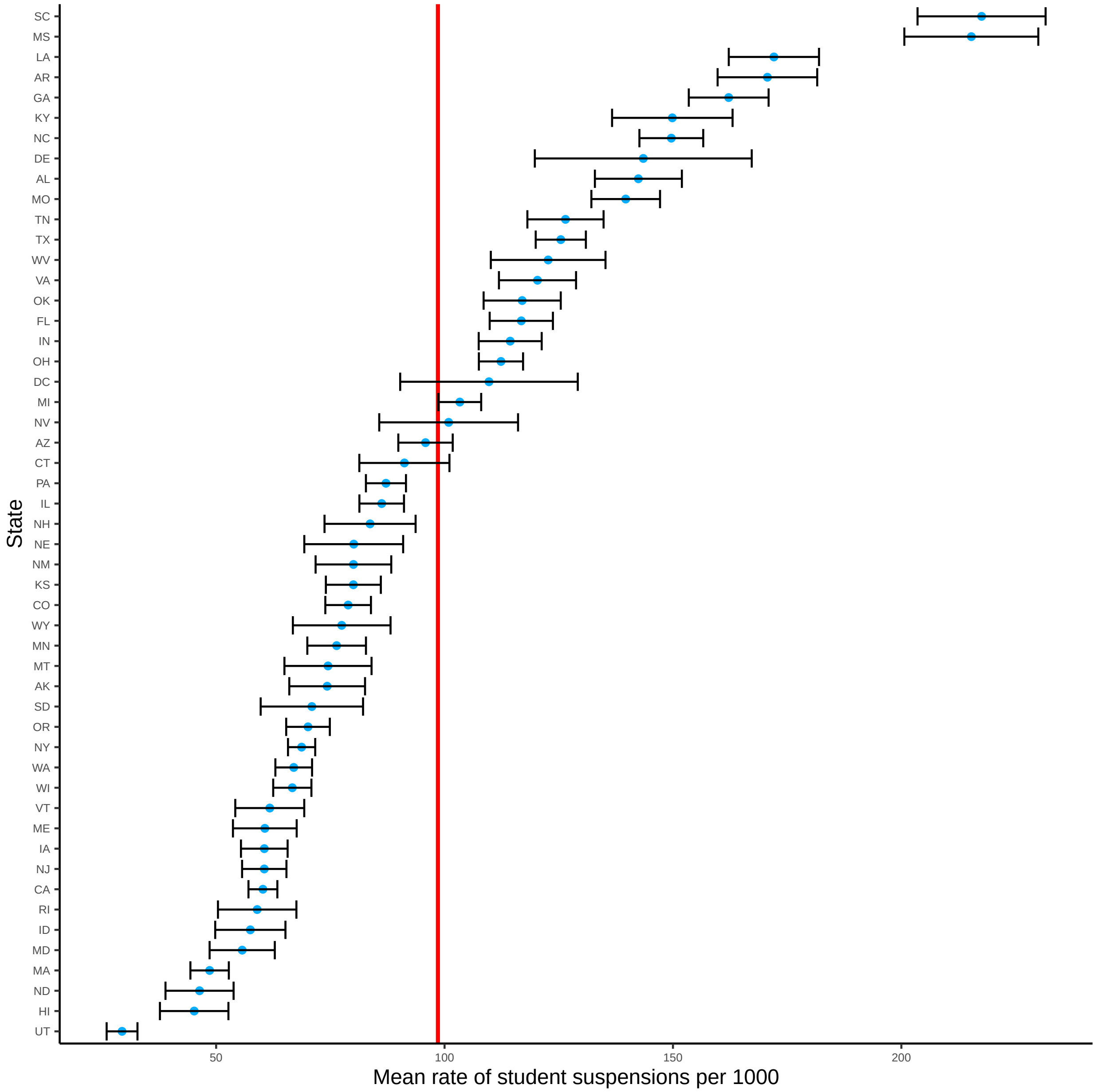
SC
MS
LA
AR
GA
KY
NC
DE
AL
MO
TN
TX
WV
VA
OK
FL
IN
OH
DC
MI
NV
AZ
CT
PA
IL
NH
NE
NM
KS
CO
WY
MN
MT
AK
SD
OR
NY
WA
WI
VT
ME
IA
NJ
CA
RI
ID
MD
MA
ND
HI
UT
State
50
100
150
200
Mean rate of student suspensions per 1000

*Figure 2*