TBD*

TBD

Rachael Lam

08 April 2021

Abstract

TBD

*Code and data are available at: https://github.com/rachaellam/edu-suspension-analysis.git.

Contents

1 Introduction

Zero tolerance policies that began as a mechanism to lower crime rates percolated into school policies and practices. In 1994, the Gun Free Schools Act compelled states who received federal funding to pass a state law that required local educational agencies to expel any student that was discovered to have brought a weapon to school. Although this act was intended to reduce violent crimes, it instead allowed schools to increase the use of out of school suspensions (OOS) for non-violent offenses such as "vulgar language, dress code violations, tardiness, or generic insubordination" (Baker-Smith 2018). The negative consequences of these exclusionary practices, such as OOS and in school suspensions (ISS), have been well documented in numerous studies. Students who encounter disciplinary action are more likely to have lower academic achievements and higher involvement in the criminal justice system (Baker-Smith 2018). One study found that instances of delinquency or crime only occurred after an OOS, validating the existence of the school-to-prison pipeline. Additionally, while ISS are seen as a less punitive form of punishment, students who received ISS had lower GPAs and higher dropout rates (Jabbari and Johnson Jr 2019). These negative repercussions do not only affect students who directly receive disciplinary sentences but also indirectly impact students who attend high-suspension rate schools (Hinze-Pifer and Sartain 2018). Suspensions can lead to an increase in fear of crime and teacher attrition (Hinze-Pifer and Sartain 2018), which can all negatively affect the performance of students and staff.

Although OOS and ISS lead to a substantial number of disadvantageous impacts for all students, they disproportionately affect students of colour. Racial disparities in exclusionary punishments begin as early as preschool where black students experience more severe and frequent punishments (Skiba, Mediratta, and Rausch 2016). Black students are 3.5 times more likely to receive OOS than white students (Skiba, Mediratta, and Rausch 2016) and in one randomized experiment, results indicated that teachers were more likely to punish a student who committed a second offense if they were black (Baker-Smith 2018). This disciplinary inequality affects all students of colour, as Indigenous and Lantinx students share similar experiences. Race, gender, sexual orientation and disability all increase a student's chance of being subject to exclusionary practices (Skiba, Mediratta, and Rausch 2016).

Socioeconomic circumstances also play a crucial role in determining OOS and ISS rates. It was found that schools with a larger percentage of students from low socioeconomic backgrounds were associated with high rates of school suspensions (Lee et al. 2011). Additionally, schools with majority white students had more financial resources than schools with majority of students of colour (Lee et al. 2011). Due to the significant function of wealth in disciplinary action, it is necessary to further explore this relationship. In this paper, I used data from the Civil Rights Data Collection (CRDC) for the 2015-2016 school year to investigate the connection between school expenditures and suspension rates. After a discussion of the data and the model used, I find that, while there is an association between teacher salary and non-personnel expenditures and ISS rates, there is no association between non-personnel expenditures and number of days missed due to OOS. Additionally, the results are obscured when running the model on the ten most populous states. Unfortunately, causality cannot be found with the available data, but I will discuss what can be done to address causality with the appropriate data. Finally, I will end with a discussion on what can be done to lower OOS and ISS rates and how that will positively affect all students by drawing on existing literature.

2 Data

To attempt an analysis on school expenditures and disciplinary punishment, I manipulated public school data from the 2015-2016 school year (**data?**). These data are collected by the Civil Rights Data Collection (CRDC) for use by the U.S. Department of Education and other policy and research agencies. The dataset was gathered in 2017 and last updated on September 28, 2018. The raw data includes 96,360 schools and 1,836 school attributes including school name, location and type; grades offered between preschool and grade 12; population of the school segmented by race and gender; type of classes and sports available; total instances of bullying, absenses, offenses, suspension and expulsions; special education and gifted and talented programs; and financial expenditures of personnel and non-personnel. Using `R` (R Core Team 2020), `tidyverse` (Wickham et al. 2019), `tidyr` (Wickham 2021), `devtools` (Wickham, Hester, and Chang 2020) and `dplyr` (Wickham et al. 2021), I cleaned and extracted the necessary data to complete an exploritory analysis and modelling.

To begin, I removed any school that was labelled alternative, charter, magnet, special education, or juvenile justice facility as I concentrated on schools that were publicly funded and adhered to the same rules and regulations. I then filtered schools that only offer grades 9-12, excluded schools that offer partial grades (ex. grades 10-12) and eliminated schools that offer preschool to grade 8 in order to limit the scope and examine similar schools. The analysis would be unbalanced if schools that only offer two grades were being compared to schools that offer four grades. Unfortunately, Hawaii and the District of Columbia were not included in the dataset after filtering by grade because there are no schools that only offer grades 9-12. In these states, schools either offer more grades, fewer grades or are ungraded, according to the dataset. Additionally, I removed columns that focused on special education and conducted the analysis without racial attributes, although this could be another area of analysis in the future especially as it is well documented that race and disabilities have an adverse affect on suspension rates. Finally, I removed any attributes regarding classes, sports, absences, offenses and bullying.

Before adding calculations of the data, I removed any schools that did not report their financial records. Schools that reported -5 signified that there were missing values but also an action plan for providing the necessary data for the subsequent school year. I then began combining columns to produce new variables that sum the number of ISS and OOS as number of enrollment and suspensions were categorized by gender. Additionally, I calculated new variables that would yield an standardizing affect on schools. For example, schools that have many ISS could just have more students in general. To adjust for this, I found the number of ISS per 100 students. I repeated this equation for the number of days missed due to OOS and number of instances of single OOS (SINGOOS) and multiple OOS (MULTOOS). I included on the number of days missed into the analysis as it contributes to the severity of school suspensions rather than just the number given. Finally, I determined the amount each school spends per student by dividing the total spent on personnel and non-personnel by the total enrolled. I also applied to calculation to the total spent on type of personnel to better understand where a school spends the majority of funds. These types of personnel include teachers, administration, support services and instructional aid. In all expenditure variables, I focused on dollar amounts without federal funding as federal funding is a minority of total funding and it is determined by a number of factors. Comparatively, dollar amounts without federal funding are largely determined by state and local funding which prioritizes location and the demographics of each neighbourhood.

2.1 Missing Data

It is crucially important to discuss the data that were purposefully excluded and that were under or unreported by the contributing schools. The decision to combine columns regardless of race, gender of disability has a variety of consequences. Namely, it could provide an additional analysis of the circumstances and address a larger portion of the hypothesis and theory. For example, a school with a high count of days missed due to OOS could have severely punished only one minority student in a majority white school population, highlighting extreme instances of discriminatory disciplinary practices. This type of analysis will not be reflected in the model and could be a property of causality. It should be noted that it is a necessary and urgent area of study, as many studies indicate that demographics are a substantial factor in suspension rates.

In addition, schools that underreport or do not report data at all, could severely affect the results. Data quality is dependent on truthful and reliable self-reporting. In the raw data, there are schools that disclose a $1,000 spend on personnel salary in the 2015-2016 school year and ones that disclose a spend of over four hundred million dollars. These outliers can lessen or inflate the extent of association between expenditures and suspensions. Additionally, schools that do not report any data raise questions of biases. There could be a type of school that is not reporting their numbers and could potentially be concealing some significant figures.

3 Model

In order to understand if a relationship exists between the amount spent per student and the number of suspensions (ISS and OOS), I performed multiple linear regression. This model was appropriate for an exploratory analysis of quantitative data, allowing for further investigation into the relationship and how that may be conducted in the future. Through multiple linear regression, I am unable to decisively conclude results but rather explore if any relationship exists and to the potential strength. I will be using four standard linear regression models, each with separate dependent variables and two predictors as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

In the first model(1), Y is the number of ISS per 100 students, X_1 is the amount spent per student on personnel without federal funding X_2 is the amount spent per student on non-personnel with federal funding. Additionally, β_0 represents the predicted value of Y when X is 0 and β_1 and β_2 are the expected change to Y when X_1 and X_2 increase. This model is repeated for three other dependent variables (Y): single OOS (2), multiple OOS (3) and number of days missed due to OOS (4).

4 Results

The model was run on four independent variables to explore if there is an association with the amount a schools spends on personnel and non-personnel and different types of school suspensions. While ISS are seen as less punitive forms of punishment, they still can lead to missed educational opportunities, ostracization and a decrease in attentiveness. In addition, the amount of OOS, both single and multiple, can inform our understanding of the severity of punishment, especially when observing the number of days missed. There is also access to the type of personnel and the correlated spend per student. Using this data, we can see if there is any association between salary of teachers, administration and support staff and the number of ISS and OOS. In the next two section, I will discuss the results of all models.

4.1 ISS and OOS Models

The first model showed some relationship between instances of ISS and both X_1, spend per student on personnel, and X_2, spend per student on non-personnel. To summarize the results shown in Table 1, a one dollar increase in spend per student on personnel is associated with a -0.00005 decrease in instances of ISS controlling for all the other variables in the model. Although the decrease of the dependent variable seems ineffectual as it is so small, a one dollar increase is a comparatively small increase. Interestingly, a one dollar increase in spend per student on non-personnel is associated with a 0.00001 increase in instances of ISS.

Both the second and third model reveals no association between instances in single OOS and multiple OOS, respectively, and spend per student on personnel (X_1) and non-personnel (X_2). These models only explain that the data is not unusual under this model, rather than proof of a decisive causal relationship. Finally, model four displays some association between number of days missed due to OOS and spend per student on personnel, but no association with spend per student on non-personnel. These results are contradictory, as I would assume if we were to see an association between days missed due to OOS and spend, we would also see some sort of association between the number of single and multiple OOS. An increase in OOS should correspondingly lead to an increase in days missed.

Unfortunately, in all models, the residual standard percentage error is well over 100%, signifying a poor model fit for the data. As a result, the model leads to inconclusive results and it is indeterminable whether or not school spend has any affect on the number of ISS and OSS with the data provided.

The question here is if an increase in spend per student is associated with a decrease in in schools suspensions or the severity of school suspensions (number of days missed). Although the model suggests that there is some relationship, it is not a representation of a causal relationship. There are many factors that could

Table 1

	ISS	SINGOOS	MULTOOS	DAYS M
	7.86583 ***	3.90181 ***	2.12017 ***	38
	(7.60059 – 8.13107)	(3.78240 – 4.02122)	(2.01808 – 2.22227)	(36.45328 – 4
R_STUDENT_WOFED	-0.00005 **	-0.00001	-0.00001	-
	(-0.00009 – -0.00002)	(-0.00003 – 0.00000)	(-0.00002 – 0.00000)	(-0.00055 – -
R_STUDENT_NPE_WOFED	0.00001 **	-0.00000	0.00000	
	(0.00000 – 0.00002)	(-0.00000 – 0.00000)	(-0.00000 – 0.00000)	(-0.00008
	8059	8059	8059	
	0.00216	0.00024	0.00023	
	-29218.71655	-22787.11406	-21524.57083	-44
	58445.43310	45582.22812	43057.14167	89

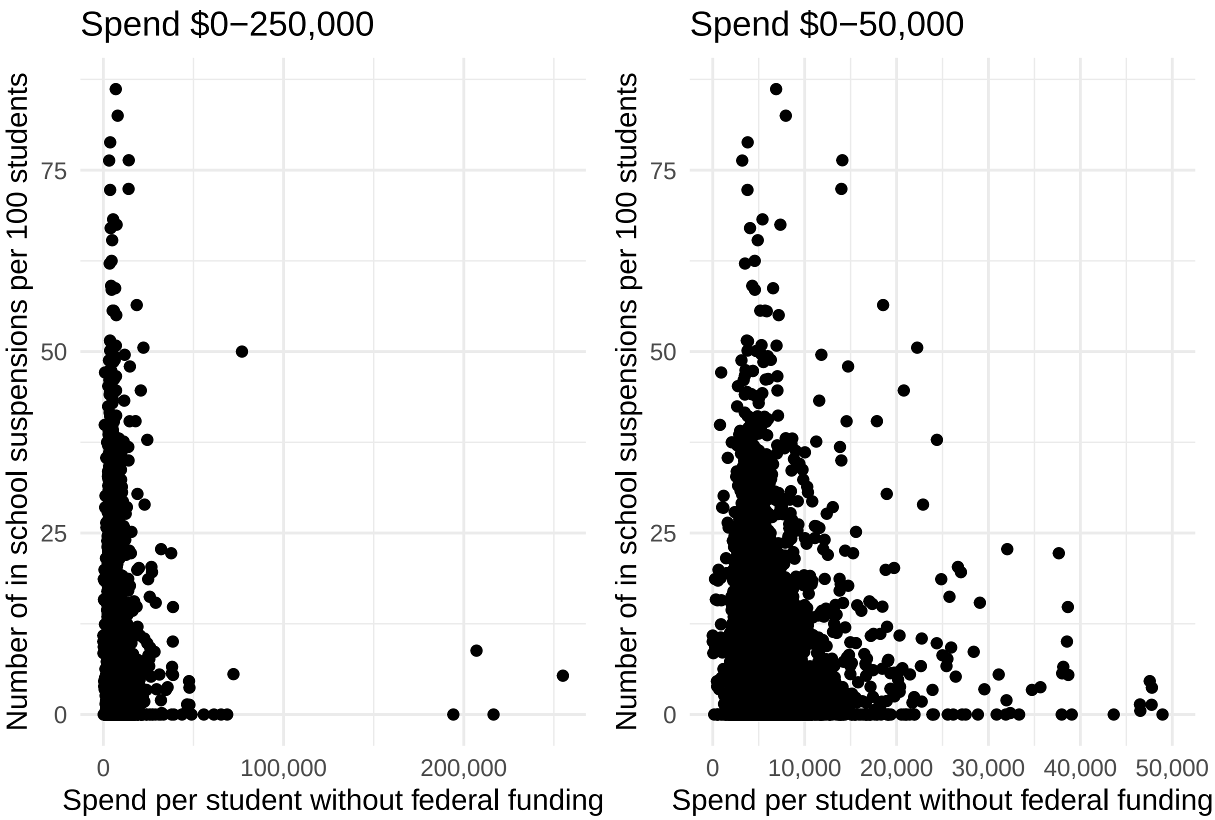
1; ** $p < 0.01$; * $p < 0.05$.

lead to high suspension rates in schools. Studies have been done that determine high-suspending schools are usually schools from in low-income communities and have a higher population of minority students.

Funding of schools relies on the majority of state and local funding. Local funding is collected from property taxes and other taxes such as sales and income. This automatically leads to a discrepancy as low-income communities collect less property tax (cheaper housing) and less sales and income tax (less income = less spending). Very few states attempt to equalize school funding, and even in states that do, there is still a large discrepancy with the amount of resources schools receive.

initial findings: there is some relationship between number of days missed due to suspensions and salary expenditures of all school personnel there is some relationship between in school suspensions and number and salary expenditures of all school personnel there is some relationship between number of days missed due to suspensions and salary of teachers but no relationship between days missed and salary of other personnel (aid, support services, administration) there is some relationship between number of in school suspensions and number of teachers. this is probably due to more teachers = more students = more suspensions there is no relationship between number of in school suspensions and number of teachers per student… odd looking at the 10 most populous states, half of them have a relationship between days missed and salary of all school personnel but the other half have no relationship. Broadening to the entire population, there is a strong relationship so there is something to explore there.

Prior to conducting multiple linear regression, I plotted the applicable data to view any preliminary patters using `ggplot2` (Wickham 2016) for graphs and `scales` (Wickham and Seidel 2020) to prepare readable numbers. Figure **??** is an example of the initial graphs. It captures the amount spent per student without federal funding on the x-axis and the number of ISS per 100 students of the y-axis. Both plots represent the same data but the figure to the right is concentrated on the school spend between \$0 and \$50,000 to remove outliers and get a better understanding of the shape of the data.



5 Discussion

5.1 Causality

speak to causality, what you would do, and why you can't do it Not able to do diff in diff so there could be some factors causing suspensions rates that are not discussed in the paper. States could vary drastically, especially in political values, thus outlooks on criminality, punishment and tolerance. create a study that gives certain schools an additional sum of money and see how it affects suspension rates.

6 Conclusion

Rebecca Hinze-Pifera and Lauren Sartain: https://doi.org/10.1080/0161956X.2018.1435051 - difficult to identify causal impact because of the many factors that go into deciding student suspension including personal biases - data quality problems due to schools under reporting - suspensions cause increase fear of crime, teacher attrition, anger, disconnect from school, impacts of missing school - negative impacts for not only students who are suspended but students who attend schools with high suspension rates - suspension inequality (race, gender, LGBTQ, special needs, socioeconomic factors)

Jason Jabbari, Odis Johnson Jr: https://doi-org.myaccess.library.utoronto.ca/10.1177%2F0042085920902256 - zero tolerance policies 1990s - exclusionary practices - urban communities with heightened police surveillance - school to prison pipeline - study that found students who attend high-suspension schools are negatively affected, even if they do not themselves receive a suspension. - mention a study that found students who received ISS suffered a lower GPA and higher dropout rates even though ISS has been seen as a less punitive form of punishment.

Talisha Lee, Dewey Cornell, Anne Gregory, and Xitao Fan: doi:10.1353/etc.2011.0014 - references studies that have been done that discuss the characteristics of a school with drop out rates - schools wit a higher percentage of low income students have higher drop out rates - higher rate of students from low socioeconomic backgrounds and a greater percentage of minority students were associated with high rates of school suspensions - schools with majority white students had more financial resources than schools with high percentage of minority students - drop out rates for rural areas tend to be lower than urban areas - reliant on schools to report drop out rates/retention rates but doesn't account for student mobility/declining school populations

E. Christine Baker-Smith: https://doi.org/10.1080/0161956X.2018.1435043 - exclusionary discipline related to lower levels of educational attainment, involvement in the criminal justice system, poor labour market outcomes - gun free schools act (1994) required any state "receiving federal funds to have a state law in effect requiring local educational agencies to expel, for at least one year, any student who is determined to have brought a weapon to school" - this act allowed schools to adopt harsher exclusionary practices against violent offenses, but it also brough an increased use of out of school suspensions for non-violent offences such as vulgar language, dress code violations, tardiness, or generic insubordination. - zero tolerance policies - a randomized experiment found that teachers were more likely to punish a student who committed a second offense if they were black

Russell J. Skiba Kavitha Mediratta M. Karega Rausch: Inequality in School Discipline Research and Practice to Reduce Disparities - racial disparities (black students are 3.5 times more likely to receive out of school suspensions than white students) - begins as early as preschool - experience harsher punishments - 1/3 of black high schools students receive suspensions in a given year compared to 1 of 10 white students - broadens to other ethnic minorities such as Indigenous students and Latinxs - disabilities combined with race increase chances of experiencing exclusionary practices. - intersectionality - gender, race, sexual orientation and disability - controlling for socioeconomic status, Black students still encounter harsher disciplinary action than white counterparts - strong relationships between students, teachers and administrators were associated with lower suspension rates - loss of educational opportunity from out of school suspensions led to higher rates of dropouts - out of school suspensions propel contact with the incarceration system. A study found that delinquency or crime only occurred after suspension (school to prison pipeline)

References

Baker-Smith, E Christine. 2018. "Suspensions Suspended: Do Changes to High School Suspension Policies Change Suspension Rates?" *Peabody Journal of Education* 93 (2): 190–206.

Hinze-Pifer, Rebecca, and Lauren Sartain. 2018. "Rethinking Universal Suspension for Severe Student Behavior." *Peabody Journal of Education* 93 (2): 228–43.

Jabbari, Jason, and Odis Johnson Jr. 2019. "The Collateral Damage of in-School Suspensions: A Counterfactual Analysis of High-Suspension Schools, Math Achievement and College Attendance." *Urban Education*, 0042085920902256.

Lee, Talisha, Dewey Cornell, Anne Gregory, and Xitao Fan. 2011. "High Suspension Schools and Dropout Rates for Black and White Students." *Education and Treatment of Children*, 167–92.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Skiba, Russell J, Kavitha Mediratta, and M Karega Rausch. 2016. *Inequality in School Discipline: Research and Practice to Reduce Disparities*. Springer.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2021. *Tidyr: Tidy Messy Data*. https://CRAN.R-project.org/package=tidyr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Jim Hester, and Winston Chang. 2020. *Devtools: Tools to Make Developing r Packages Easier*. https://CRAN.R-project.org/package=devtools.

Wickham, Hadley, and Dana Seidel. 2020. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.