What Happens When High School Teachers Hit Their Students?

By Max Schotz

ABSTRACT: Corporal punishment is still used in many Southern, public high schools to discipline students. This paper uses two-stage least squares instrumental variable analysis to examine the effects that this corporal punishment usage has on academic performance and antisocial behavior within the student body. I find that if a school goes from using no corporal punishment to using the average level for a school that employs the practice, the rate of out-of-school suspensions goes down by 3.77 percentage points (.44 standard deviations), while the SAT participation rate goes down by 15.34 percentage points (1.10 standard deviations). These results hold across gender and race. While this paper finds evidence that the use of corporal punishment can decrease antisocial behavior within the student body, this reduction comes at the price of a severe decline in academic performance. Because of this, the use of corporal punishment in high schools is strongly recommended against.

Acknowledgements

Thank you to the Haverford Economics department for giving me the tools over the past four years to analyze the issue of corporal punishment and write a thesis that encompasses all of this work. Thanks also to Professor Richard Ball for helping me define the questions I would be asking and for directing me to the best methodologies for answering these questions. I am especially appreciative of Professor Anne Preston for providing an incredible amount of her time and expertise to help turn my proposal into a completed thesis. And, of course, thank you to my parents and my sister for tolerating me for the past 22 years.

Table of Contents

Introduction:

Though a majority of states in the U.S. have done away with the use of corporal punishment in schools over the past few decades, there still exists a wide swath of America where physical punishment is used to discipline students. This phenomenon is not isolated to elementary schools, either; there are many high schools where corporal punishment is still considered a common practice for disciplining children. A large amount of research has been done on the effects of household corporal punishment on children, but little has been done to look at the effects of physical punishment in the classroom. And in high school—an academic environment where students' futures are directly at stake—there is a need for an examination of corporal punishment's effects on the social and academic well-being of students.

Using data collected by the Office of Civil Rights and the U.S. department of education for the 2013-2014 academic year, this paper utilizes two-stage instrumental variable regression and other methodologies to examine this issue. Whereas prior research in this field predominantly analyzes panel data for hundreds of children over time, I am analyzing a large (3,834 observations) cross-section of public high schools located in states where in-school corporal punishment is still legal. While previous research looks into the effect that corporal punishment has on individual children, I am looking at the ramifications of physical punishment on the student bodies of these high schools. I aim to answer two questions with my analysis: first, does the use of corporal punishment in a high school increase the frequency of antisocial behavior within the student body? Previous research finds that the use of corporal punishment in the household causes antisocial behavior in children (Grogan-Kaylor (2004), Straus, Sugarman, & Giles-Sims (1997)). While the use of physical punishment like spanking is controversial within homes, it is even more controversial in schools—where it is not the parents doing the

disciplining; it is teachers. To answer this question, I use the out-of-school suspension rate for each school as a measure of antisocial behavior in the student body, as previous literature suggests a vast majority of these suspensions are related to antisocial behavior (Mendez & Knoff (2003)). In line with previous research, I posit that an increase in corporal punishment usage causes more antisocial behavior in the student body.

I also examine whether the use of corporal punishment affects academic achievement, which is why high schools were chosen for this analysis as opposed to elementary or middle schools. Previous literature finds an inverse connection between antisocial behavior and academic achievement (McEvoy & Welker (2000), Maguin & Loeber (1996)). So, if the use of corporal punishment in high schools in fact causes more antisocial behavior, I would like to know if this has ramifications on the academic achievement of students. Besides the effect that an increase in antisocial behavior can have on academic success, I also believe that the use of corporal punishment disengages students and creates a toxic learning environment. This is another channel by which corporal punishment can adversely impact academic achievement. I use the SAT participation rate of each school as a measure of academic success, because this variable signals academic standing in high school and college matriculation.

Corporal punishment is a complex point of contestation in many pockets of the United States due to the long-standing tradition of its use, the perceived lack of viable alternatives, or a combination of these and many other factors. However, the analysis put forth in this paper allows the issue of corporal punishment in high schools to be analyzed while setting aside the moral, ethical, and cultural factors that often derail discussions and pollute academic research. In creating a work devoid of the many biases that have plagued previous research, I hope to make a substantive contribution to research on the way that corporal punishment affects students.

Literature Review/Background:

Much of the prior work in this area consists of appeals to the morality of the reader, and these works read much more like closing arguments than scientific studies. Lots of the empirical work done in the field attempts to adopt these ethical arguments while finding scientific backing for them. However, a broad examination of the research done in this area shows that academics are *not* united in trying to answer the question at hand before their data do. Indeed, there has been research done which posits a place for corporal punishment in the development of children.

Grogan-Kaylor (2004) looks at whether there is a connection between corporal punishment and antisocial behavior in children by utilizing time-series data from the National Longitudinal Survey of Youth. The author uses ordinary least squares and fixed-effects analysis to find that antisocial behavior is caused by the variable "corporal punishment used in the past week" at a significance level of less than one percent. Additionally, he finds that this effect is non-linear, as corporal punishment's effect on antisocial behavior is tied to whether it is used or not—not the magnitude or frequency of its use. A common issue in this field of study is the issue of endogeneity, because children who misbehave more frequently could be physically punished more frequently. As a result, it can be unclear whether corporal punishment causes antisocial behavior or antisocial behavior causes corporal punishment. Grogan-Kaylor controls for this issue in OLS regressions of the 1998 wave of NLSY data by controlling for a child's prior level of antisocial behavior. Additionally, in analysis of panel data from the NLSY, the author uses fixed effects analysis to control for time-invariant aspects of children's households that could have an effect on antisocial behavior. No other controls for endogeneity were included. This paper presents a compelling argument due to its large sample size of over 1,000 children, and it

is one of the most prominent studies put forth by the group of academicians and researchers who find themselves opposing the use of corporal punishment.

Straus, Sugarman, & Giles-Sims (1997) also explore the connection between corporal punishment and antisocial behavior by using analysis of variance on longitudinal survey data from 807 mothers. They find that corporal punishment has a long-term effect of increasing antisocial behavior in children. They expand beyond Grogan-Kaylor's study by finding that when corporal punishment is used into a child's adolescence, it can be connected to depression, physical assaults, and alienation upon adulthood. The authors recognize the same endogeneity problem that Grogan-Kaylor specifies, and they highlight the inability of previous research to control for the two-way causation problem presented by corporal punishment and antisocial behavior. Similar to Grogan-Kaylor, the authors control for the child's level of antisocial behavior at time 1 in analyzing the causes of the child's antisocial behavior in time 2. Straus, Sugarman, & Giles also recognize that the use of corporal punishment could have collinearity with other parental styles that have adverse effects on antisocial behavior in children. To control for other aspects of parental involvement that impact antisocial behavior, the authors include measures of emotional warmth and cognitive stimulation in their analyses.

These are strong findings, but their accuracy is somewhat cast into doubt by the obvious anti-spanking tone of the paper and by the findings of Straus & Stewart (1999). Straus and Stewart do not examine whether corporal punishment has adverse developmental effects on children. Rather, they look at the prevalence and severity of corporal punishment. Their main findings could be summarized as follows: Corporal punishment is everywhere, and it is bad, so we must work to get rid of it. Obviously this work should serve to compliment Straus's earlier work, but instead it makes clear that he has a blatant agenda which has likely tinged his research.

When his earlier work is reexamined in this light, the bias is evident from a methodological point of view as well: there is no distinction in his 1997 study between severe forms of corporal punishment and lighter forms. It could be that severe forms of corporal punishment have adverse effects on children, while lighter forms of punishment have no such ill effect.

In fact, Larzelere & Kuhn (2005) find that this distinction makes a large difference in the findings that researchers come to. In their meta-analysis of research on corporal punishment and alternative methods, the authors find that decomposing corporal punishment into different degrees of severity leads to concrete findings that contradict the earlier work of Straus and his associates. Larzelere & Kuhn conclude that adverse effects on children occur not from physical punishment itself; rather, children are harmed when corporal punishment is used too severely (i.e. more drastic than customary spanking) or when corporal punishment is the *primary* mode of disciplining children. Additionally, they find that customary spanking is actually associated with less defiance and antisocial behavior than a majority of alternative disciplinary tactics. Larzelere & Kuhn highlight the same issue of endogeneity that prior research does, and they control for it in a similar manner. When analyzing effect sizes, the authors only use data from studies where initial child misbehavior is taken into account. By controlling for initial antisocial behavior and decomposing corporal punishment into subcategories, Larzelere & Kuhn address the two largest methodological pitfalls of prior research in the area.

Fuller (2009) takes a decidedly much more qualitative path to come to a similar conclusion that Larzelere & Kuhn do—that corporal punishment can have a place in a responsible parent's household. This paper examines the case study of Sweden, which was the first country to outlaw spanking in 1979 in the hopes that it would perpetuate a culture of nonviolence in households. Fuller argues that the ban has had the opposite effect, with assaults

against children having actually gone up since the ban was passed (though this could simply be due to increased reporting). He uses a table to compare the frequency of criminal assaults against 7 to 14 year old children in 1984 versus 1994, and he finds increases across the board (no matter the age of the assaulter). The author also includes a table that looks at the frequency of cases of in-home and general abuse of children ages 0 to 6 from 1981 to 1999. Over the years, he notes a clear trend of increases in abuse of both types, which further supports his argument that Sweden's ban on corporal punishment was counterproductive. It should be noted, however, that Fuller presents this data in the form of number of cases and not instances per capita, so this data could be misleading due to population growth. On a grander scale, Fuller believes that the generations of children who have grown up in Sweden since the passing of the ban lack the discipline of previous generations, and this is even reflected in the outside world's perception of Swedish children (he cites travel brochures that warn prospective tourists of the behavior of Swedish youths). The author also heavily criticizes previous works that have come to unfavorable conclusions about corporal punishment, saying that they typically do not control for enough factors or they have the tendency to lump together forms of physical abuse with legitimate child-rearing methods.

A majority of the research in this field deals with children on an individual level, examining how household corporal punishment impacts their well-being. Most of these authors also have access to panel data, whereas I am dealing with a cross-section of American high schools for the 2013-14 school year. To contextualize my research, it is essential to examine much more qualitative cultural case studies on corporal punishment in high schools, because the same level of quantitative analysis has not yet been applied to this field.

Carr (2014) poses a simple question: “Why are Black Students Facing Corporal Punishment in Public Schools?” The author looks at the case study of Holmes County, Mississippi, where 99 percent of school children are black. In this town, corporal punishment is a way of life that has raised countless generations, and its use is surprisingly uncontroversial. In fact, parents in the town deem corporal punishment as completely necessary to raise children, as do many teachers who find its use essential in the classroom. This culture follows children from the moment they start preschool until they graduate high school. And though the author notes that corporal punishment is used disproportionately often on black children at a national level, its acceptance at a local level could be much more tied in with racial dynamics than implicit bias on the part of the teacher. In Holmes County, most educators who use the paddle are black. So, even though it is black children who are disproportionately targeted by the practice, arguments against corporal punishment can easily be construed as attacks on the rights of the black community to self-govern. Herein lies a complicated dynamic where it is very difficult to disentangle discrimination from a long-accepted culture of using corporal punishment to raise children.

The work most relevant to my research is another article that is decidedly less academic than the earlier articles. Rather, its importance lies in the fact that it uses the same dataset I am going to use, and it examines racial discrepancies in corporal punishment in a manner similar to the way I am going to. Sparks & Harwin (2016) use the 2013-14 Civil Rights Data Collection Survey to find 109,000 cases of students being paddled, swatted, or otherwise physically punished during the 2013-2014 school year. They also used the Education Department’s Common Core of Data for 2013-14 for data on poverty levels in school districts. They find the vast majority of these cases in the south. The article uses no regression analysis to control for other factors, but they nonetheless come to conclusions that deserve more exploration. They find

that low-income students are more likely to attend schools that use corporal punishment than higher-income students on a nationwide basis. Additionally, white students make up 60 percent of the student population but only 50 percent of those who experience corporal punishment. Conversely, black students make up 22 percent of students at schools where corporal punishment is used, but they find themselves on the receiving end of this practice in 38 percent of cases nationwide. The article also points out that there is little guidance on how and when corporal punishment should be administered.

To measure antisocial behavior at a student body level, I am using the percentage of students subjected to at least one out-of-school suspension in the 2013-2014 school year. This decision is born out of previous research that substantiates the strong connection between out-of-school suspensions and antisocial behavior. Mendez & Knoff (2003) conduct an analysis of out-of-school suspensions in a large (142 schools) and diverse school district in Florida for the 1996-1997 school year. They find that 18.90% of male high school students experienced at least one out-of-school suspension, while 8.84% of female high school students did. Additionally, 39.46% of black males were subjected to at least one out-of-school suspension, with 20.69% of black females having the same fate. Across all grades, Mendez & Knoff find that antisocial behavior accounts for at least 82% of out-of-schools suspensions in their sample. This antisocial behavior presented itself in the form of disobedience or insubordination, disruption, fighting, inappropriate behavior, noncompliance, profanity, disrespect, and other forms of misbehavior. In the high schools within their sample, the authors find that 3.65% of all students were suspended during the 1996-1997 school year for violence against persons, 1.62% for violence against property, 4.28% for substance possession, and 26.74% for disobedience. Mendez & Knoff's findings

reinforce my choice of out-of-school suspensions as a variable for student body-level antisocial behavior.

There is also previous research which suggests that elevated levels of antisocial behavior could hamper the academic achievement of students. McEvoy & Welker find in their 2000 meta-analysis of antisocial behavior and academic performance that academic performance is consistently inversely related to antisocial behavior, though causality is difficult to establish. Maguin & Loeber (1996) find that poor academic performance is connected with the onset and frequency of delinquency, and better academic performance is associated with lower levels of misbehavior. However, they too can not establish causality. McEvoy & Welker believe that while a student's antisocial conduct is in part a byproduct of poor academic performance, it is equally likely that the antisocial conduct can contribute to poor academic performance. They do, however, hypothesize a "contagion" effect that explains why antisocial behavior could impact academic performance within schools adversely. This hypothesis stipulates that because classes contain many students—each of whom are impacted by the behavior of other students—the misbehavior of one student can adversely impact the academic performance of the class as a whole. This "contagion" effect is an effective argument for why higher levels of antisocial behavior amongst students could cause lower academic achievement at the student body level.

For this paper, I am building upon previous findings of a causal relationship between corporal punishment usage and antisocial behavior, but I am expanding my analysis of this issue by examining it at the school (not individual) level. By looking at corporal punishment through this lens, I am also able to examine how corporal punishment affects high school students' academic success, which is an area that has been neglected by previous research. This neglect is especially surprising considering the prevalence of corporal punishment in some southern states.

As shown below in Table 1, a surprisingly large portion of public high schools in some states still use corporal punishment. The areas represented below are the east south central and west south census census blocks (eight states in total), and this sample was chosen because a vast majority of cases of in-school corporal punishment occur in these eight states.

Table 1: The percentage of high schools that allow corporal punishment in each state

State	Percentage
Mississippi	45.96%
Alabama	40.22%
Oklahoma	37.35%
Arkansas	30.00%
Texas	22.53%
Tennessee	16.81%
Louisiana	11.63%
Kentucky	0.58%
All states	**24.62%**

Note. Data from the Office of Civil Rights and National Center of Education Statistics.

Especially in the South, the issue of in-school corporal punishment is one where the fates of many students are at stake, and I intend for my paper to give this issue the attention it deserves. Whether a decrease in academic achievement is brought on directly by corporal punishment usage or indirectly via an increase in antisocial behavior (which would then cause lower academic achievement), my analysis parses out this causality. In the process, I contribute to previous research by examining just how corporal punishment affects teenagers where it matters most—in the classroom.

Data:

The first source of data for this paper is the 2013-14 Civil Rights Data Collection (CRDC) Public Use Data File. The Office of Civil Rights conducts a nationwide study of every school from pre-kindergarten through the twelfth grade every two years, recording data from one academic school year. All of the data from this data-collection for the 2013-14 school year was then aggregated into the CRDC Public Use Data File that I am using for this paper.

I took this dataset (which initially contained data on over 92,000 schools) and narrowed it down to public, non-charter, non-special-educational high schools in the 15 states where corporal punishment is still legal. In all, my final dataset has 3,834 observations, and each one is an individual high school. I eliminated observations for which there were missing or possibly erratic values for variables of interest in my study. These possibly erratic values were flagged by the Office of Civil Rights, and their occurrence was extremely infrequent. There is a possible problem that a school which had in-school arrests or SAT testing data that presented it in an unfavorable light could omit this data from its CRDC submission and thus be omitted from my study. I am not worried about this concern for several reasons. First, the data is submitted by district administrators—not individual schools—so it is unlikely that data is omitted from CRDC submission to protect a particular school. Second, if a district administrator fails to submit data for any of the parameters that the Office of Civil Rights requires, the administrator is required to submit a detailed action plan to ensure that this data will be submitted in the next round of data collection (2 years later). So, there is little incentive to cheat by data omission, and this incentive is diminished further for even the most short-sighted administrators by the Office of Civil Rights's requirement of action plan submission and execution upon the next wave of data collection.

I also take data from the National Center for Education Statistics (NCES) Common Core of Data Public Elementary/Secondary School Universe Survey for the 2013-14 school year. The National Center for Education Statistics is a subdivision of the Department of Education that collects and analyzes data related to education in the United States. The Common Core of Data is a program that collects data covering the universe of public schools in the United States for each school year. From this dataset, I obtained location information for each of the schools in my data set and socioeconomic information for the student body. I used location variables from this dataset to construct a variable that identifies whether a school is located in a city, suburb, town, or rural area. The socioeconomic variable I created using this dataset is the percentage of students who are receiving free or partially subsidized lunches.

My final dataset consists of 56 variables, with 49 school-level variables constructed from the CRDC Public Use Data File and 7 variables made using the NCES Common Core of Data for the 2013-14 school year. 29 variables were used in the main body of regressions, and descriptive statistics for these variables are included in Table 2.

Table 2: Descriptive Statistics for Variables Used in Main Regressions						
	Variable	Obs	Mean	Std. Dev.	Min	Max
1	*Corp punish dummy*	3832	0.15	0.36	0.00	1.00
2	*Student teacher ratio*	3750	19.36	38.68	0.00	916.67
3	*Percent white*	3834	0.52	0.30	0.00	1.00
4	*Percent black*	3834	0.22	0.26	0.00	1.00
5	*Percent asian*	3834	0.02	0.04	0.00	0.59
6	*Percent hispanic*	3834	0.19	0.24	0.00	1.00
7	*Percent other*	3834	0.05	0.10	0.00	1.00
8	*Male SAT participation*	3833	0.19	0.12	0.00	1.00
9	*Female SAT participation*	3833	0.23	0.14	0.00	1.00
10	*White SAT participation*	3786	0.22	0.15	0.00	1.00
11	*Black SAT participation*	3561	0.22	0.18	0.00	1.00
12	*All student SAT participation*	3833	0.21	0.13	0.00	1.00
13	*Male corp punish rate*	3833	0.01	0.05	0.00	0.81
14	*Female corp punish rate*	3834	0.00	0.02	0.00	0.40
15	*White corp punish rate*	3826	0.01	0.03	0.00	0.40
16	*Black corp punish rate*	3757	0.01	0.05	0.00	1.00
17	*All student corp punish rate*	3833	0.01	0.03	0.00	0.44
18	*Male suspension rate*	3833	0.10	0.10	0.00	0.99
19	*Female suspension rate*	3833	0.06	0.07	0.00	0.83
20	*White suspension rate*	3787	0.06	0.08	0.00	1.00
21	*Black suspension rate*	3562	0.12	0.13	0.00	1.00
22	*All student suspension rate*	3833	0.08	0.08	0.00	0.90
23	*Socioeconomic status*	3826	0.53	0.22	0.00	1.00
24	*Ln(funding per student)*	3823	8.28	0.69	0.00	10.82
25	*Ln(overall enrollment)*	3834	6.59	0.98	0.69	8.59
26	*city*	3834	0.21	0.41	0.00	1.00
27	*suburb*	3834	0.23	0.42	0.00	1.00
28	*town*	3834	0.18	0.39	0.00	1.00
29	*rural*	3834	0.38	0.49	0.00	1.00

For the entire sample of the 15 states where corporal punishment is legal, 15 percent of public high schools allow the usage of corporal punishment (row 1). The maximum for student-teacher ratio is 916 (row 2), which is certainly an outlier. However, I ran regressions omitting all schools where the student-teacher ratio is over 100, and the results of my regressions did not change. Additionally, the distribution of student-teacher ratio is normal when omitting these

outliers. The average SAT participation rate for schools in my sample is 21 percent (row 12), while the out-of-school suspension rate is eight percent (row 22). Funding per student and overall enrollment for each school are turned into natural logarithms due to the wide dispersion of values between schools (rows 24 & 25). These descriptive statistics show the rate of corporal punishment usage in this sample to be one percent on average (row 17), but this figure is five percent when including only schools that utilize the practice. Socioeconomic status is the percentage of students at a given school who receive free or assisted lunches (row 23).

Included in the full dataset are dummy variables for a school's region, whether or not it uses corporal punishment, funding per student (in dollars), student-teacher ratio, and many other school-level variables. Additionally, I created a variable using data from the Historical American Lynching Data Collection Project that documents the number of lynchings recorded in the same zip code as each school in the dataset. A list of all 56 variables and variable meanings can be found in Appendix A. Variables that pertain to enrollment, AP course enrollment, SAT participation, corporal punishment, and suspensions are all expressed as percentages of the student population. Furthermore, these variables are broken down by race (including white, black, and all) and gender. So for example, *corp_punish_rate_black* is the percentage of black students at a given school who were subjected to corporal punishment at least once in the 2013-2014 school year. Missing observations, mean, standard deviation, minimum, and maximum for all variables are shown in Appendix B.

Model/Methodology:

This paper utilizes two-stage least squares instrumental variable regression to assess the effect that corporal punishment usage has on students. The first model I run concerns corporal punishment's effects on antisocial behavior in students. Without the use of instrumental variables, this model is as follows:

$$susp\_rate\_all_i = \beta_0 + \beta_1 * corp\_punish\_rate\_all_i + \beta_2 * racial\_composition_i + \beta_3 * student\_teacher\_ratio_i + \beta_4 * \ln(fundingperstudent)_i + \beta_5 * \ln(enrollment\_all)_i + \beta_6 * SESall_i + e_i$$

The percentage of students subjected to out-of-school suspension is the dependent variable of interest in this regression, because it is a good measure of the level of antisocial behavior in the student body. The variable *corp_punish_rate_all*, which is the percentage of students who are subjected to corporal punishment at a given school, is the independent variable of interest. Controls for the antisocial behavior model include racial composition, which represents the variables percent white, percent black, percent asian, and percent hispanic to account for the racial composition of each school. Funding per student is the state and local funding a school receives yearly per student, while student teacher ratio is a school's student-teacher ratio as measured by full-time-equivalent count. *Enrollment_all* measures the total enrollment of a school, and *SESall* is the portion of the student body that is eligible for free or reduced-price lunches. This is incorporated as a barometer of the socioeconomic status of the student body.

The model as presented has the issue of two-way causation, because corporal punishment could increase the frequency of out-of-school suspensions. However, more out-of-school suspensions could also be associated with higher levels of corporal punishment because there are more "problem students," or certain students could act up more frequently after being suspended.

These concerns mean that the uninstrumented model displayed above is likely to overestimate the coefficient of *corp_punish_rate_all*. To deal with the issue of two-way causation, this paper uses a two-stage least squares instrumental variable model that takes the following form:

First Stage: $corp\_punish\_rate\_all_i = \beta_0 + \beta_1 * racial\_composition_i + \beta_2 * student\_teacher\_ratio_i + \beta_3 * \ln(fundingperstudent)_i + \beta_4 * \ln(enrollment\_all)_i + \beta_5 * SESall_i + \beta_6 * location_i + e_i$

Second Stage: $susp\_rate\_all_i = \alpha_0^{IV} + \alpha_1^{IV} * corp\_punish\_rate\_all_i + \alpha_2^{IV} * racial\_composition_i + \alpha_3^{IV} * student\_teacher\_ratio_i + \alpha_4^{IV} * \ln(fundingperstudent)_i + \alpha_5^{IV} * \ln(enrollment\_all)_i + \alpha_6^{IV} * SESall_i + \varepsilon_i$

The dependent variable, independent variable of interest, and controls in the second stage are the same as the previous model, but in the first stage *corp_punish_rate_all* is instrumented on the 3 (one omitted) variables that document the school's proximity to a city (*urban, suburb, town, rural*). I chose these variables as instruments because I hypothesize the urban/rural dichotomy to be one of the main determinants of whether or not a school utilizes corporal punishment—and the degree to which it does (Carr (2014), Sparks & Harwin (2016)). I think these instruments are relevant to *corp_punish_rate_all*, but I also hypothesize that these dummy variables are exogenous, because I do not believe they have significant independent effects on out-of-school suspensions after instituting the controls included in this model. I empirically test the hypotheses of relevance and exogeneity of these instruments.

To test for the relevance of instrumental variables, the first stage of the aforementioned model is run, and I test whether *city, suburb, town,* and *rural* are jointly equal to zero. This test yields an F value which signals the explanatory power these variables have with respect to the corporal punishment variable of interest. If this F value is above 10, the instruments are relevant enough in explaining *corp_punish_rate_all* to use. If the F value is below 10, these instruments

are not relevant. The results of all instrument relevance tests are presented in Appendix C and in regression tables in the Results section.

To test for the exogeneity of instrumental variables, I calculate the residuals from the second stage of the antisocial behavior model. I then regress these residuals on the instruments *city, suburb, town,* and *rural* along with all other controls included in the second stage of the regression. Then, an F-test is run on whether *city, suburb, town,* and *rural* are jointly equal to zero. The F value obtained from this test is used to calculate a J value, where J=mF (m = the number of instruments). J has a chi-squared distribution with degrees of freedom equal to the degree of overidentification. For this test, a lower J value (and higher p value) is preferred, because this signals that the instruments cannot explain variation in the dependent variable of interest and are therefore exogenous. Using a significance level of 5 percent, I test the following hypothesis: H_0: J ≥ chi-squared critical value; H_A: J < chi-squared critical value. If the p value for any J statistic is below .05, I cannot reject the null hypothesis, and the instruments must therefore be treated as endogenous. If the null hypothesis is rejected in favor of the alternative hypothesis, the instruments are exogenous. All exogeneity tests conducted are included in Appendix D and in regression tables in the Results section.

The second two-stage least squares model addresses corporal punishment's effect on SAT participation. Uninstrumented, this model is as follows:

$$SAT\_participation\_all_i = \beta_0 + \beta_1 * corp\_punish\_rate\_all_i + \beta_2 * racial\_composition_i + \beta_3 * student\_teacher\_ratio_i + \beta_4 * \ln(fundingperstudent)_i + \beta_5 * \ln(enrollment\_all)_i + \beta_6 * SESall_i + e_i$$

The dependent variable of interest is *SAT_participation_all*, which is the portion of students at a given school who take the SAT. As mentioned earlier, this measures both the academic achievement of the student body *and* the ambition of students to go to college. This

uninstrumented model has similar two-way causation problems to the antisocial behavior model. The use of corporal punishment could cause academic achievement to decline, but students could also be getting punished because they are lower achievers academically. To deal with this issue of two-way causation, I instrument *corp_punish_rate_all* on the geographic variables *city, suburb, town,* and *rural* in the same manner that I did in the antisocial behavior model. This academic achievement model is shown below:

First Stage: $corp\_punish\_rate\_all_i = \beta_0 + \beta_1 * racial\_composition_i + \beta_2 * student\_teacher\_ratio_i + \beta_3 * \ln(fundingperstudent)_i + \beta_4 * \ln(enrollment\_all)_i + \beta_5 * SESall_i + \beta_6 * location_i + e_i$

Second Stage: $SAT\_participation\_all_i = \alpha_0^{IV} + \alpha_1^{IV} * corp\_punish\_rate\_all_i + \alpha_2^{IV} * racial\_composition_i + \alpha_3^{IV} * student\_teacher\_ratio_i + \alpha_4^{IV} * \ln(fundingperstudent)_i + \alpha_5^{IV} * \ln(enrollment\_all)_i + \alpha_6^{IV} * SESall_i + \varepsilon_i$

corp_punish_rate_all is again instrumented on the dummy variables *city, suburb, town,* and *rural* (3 with 1 omitted). I posit these instruments to be relevant for the same reasons that they were relevant for the previous model. They are also exogenous most of the time, which I attribute to the suite of controls included in these models. It was a concern that SAT participation would be heavily tied to a school's proximity to a city, but this connection did not show up frequently in my models when controls were included. Relevance testing of instruments for this model is included in Appendix C, and exogeneity testing can be found in Appendix D.

This dataset also includes gender and race-based breakdowns of each variable, which allow me to run the out-of-school suspension and SAT participation regressions on smaller subgroups of the student population. I run these regressions for males, females, white students, and black students to see if certain groups are affected more than others. As an example, the

following model addresses the effect of corporal punishment on the SAT participation rate of black students:

First Stage: $corp\_punish\_rate\_black_i = \beta_0 + \beta_1 * racial\_composition_i + \beta_2 * student\_teacher\_ratio_i + \beta_3 * \ln(fundingperstudent)_i + \beta_4 * \ln(enrollment\_all)_i + \beta_5 * SESall_i + \beta_6 * location_i + e_i$

Second Stage: $SAT\_participation\_black_i = \alpha_0^{IV} + \alpha_1^{IV} * corp\_punish\_rate\_black_i + \alpha_2^{IV} * racial\_composition_i + \alpha_3^{IV} * student\_teacher\_ratio_i + \alpha_4^{IV} * \ln(fundingperstudent)_i + \alpha_5^{IV} * \ln(enrollment\_all)_i + \alpha_6^{IV} * SESall_i + \varepsilon_i$

Additionally, I run the antisocial behavior and SAT participation models with a corporal punishment dummy variable instead of corporal punishment usage rates. *Corp_punish_dummy* is 1 if a school allows the use of corporal punishment to discipline students, and it is 0 if corporal punishment is not allowed. All other aspects of these models are the same as their corresponding *corp_punish_rate_all* variants: I run two-stage least squares regressions using *city*, *suburb*, *town*, and *rural* as instruments. I run these regressions using the dummy variable to find out if the coefficients on *corp_punish_rate_all* in the original model reflect the differences within schools that use corporal punishment or whether they show the differences between schools that allow corporal punishment and those that don't. Relevance and exogeneity testing for instruments are once again found in Appendixes C and D.

I utilize all of these models on three different samples. The first sample consists of the 15 states where corporal punishment usage in schools is allowed (AZ, WY, TX, TN, OK, MS, AR, LA, MO, AL, GA, FL, NC, SC, and KY). This is the largest sample I analyze, because making the sample any bigger would only add more schools with increasingly heterogeneous cultural factors that I can't control for. The second sample consists of the west south central and east south central census blocks, which is an eight state grouping where a vast majority of in-school

corporal punishment in the United States takes place. These eight states are Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas. The third sample includes these eight states along with Georgia, which falls outside of the aforementioned census blocks but still has many schools that use corporal punishment. The regressions for the nine state sample are included in Appendix E.

Regressions were also run on a fourth sample that consists only of schools that utilize corporal punishment. This analysis was run to find out if increased usage of corporal punishment has academic or social ramifications at a school that already uses corporal punishment. In all regressions for both antisocial behavior and academic performance, I was unable to reject the null hypothesis that the instruments were irrelevant in explaining corporal punishment usage. These instrumental variable relevance test results are shown in Appendix C.

My analysis is centered around the first two models fleshed out in this section: the *susp_rate_all* instrumental variable (IV) regression and the *SAT_participation_all* IV regression. I conduct the same regressions on male, female, white, and black subsets of the student population to find out if certain subgroups of the student population are affected by corporal punishment more than others. Finally, the corporal punishment dummy variable regressions serve as an important robustness check and also as an indicator of what is driving the coefficients on *corp_punish_rate_all* in the antisocial behavior and academic performance models.

Robustness Checks

I ran several specifications in addition to the aforementioned models. Using data from the Historical American Lynching Data Collection Project, I created a variable for the number of lynchings that occurred in the same zip code as each school. The idea was that this variable

might capture certain social factors that the other controls in my model wouldn't be able to. Unfortunately, when I used this variable as an instrument for *corp_punish_rate_all*, I could not reject the null hypothesis that the lynching variable was irrelevant in explaining *corp_punish_rate_all*.

I also ran two-stage antisocial behavior regressions with in-school suspensions as the dependent variable instead of out-of-school suspensions. I then added the in-school suspension rate and out-of-school suspension rate for each school together and ran regressions with that variable as the dependent variable. The idea behind this robustness check is that while out-of-school suspensions are related to antisocial behavior a vast majority of the time, it could be that not all antisocial behavior is captured in out-of-school suspensions. Certain schools could use in-school suspensions in the place of out-of-school suspensions, and this could taint the effectiveness of out-of-school suspensions in capturing antisocial behavior. Unfortunately, the instruments were endogenous for both the in-school suspension regressions and the combined in and out-of-school suspension regressions.

Next, I ran regression discontinuity models for antisocial behavior and SAT participation. This entailed pairing up each school that used corporal punishment with the nearest school that didn't use it. This controls for localized cultural variables that I can't control for otherwise with my dataset. Once each pair of schools was put together as one entity, I ran entity fixed effects regressions. The model for antisocial behavior is as follows:

$$susp\_rate\_all_i = \beta_0 + \beta_1 * corp\_punish\_rate\_all_i + \beta_2 * racial\_composition_i + \beta_3 * student\_teacher\_ratio_i + \beta_4 * \ln(fundingperstudent)_i + \beta_5 * \ln(enrollment\_all)_i + \beta_6 * SESall_i \beta_7 * E_i + \varepsilon_{i,E}$$

E_i represents entity fixed effects in this model. I also ran the same model using the corporal punishment dummy. These regressions were run on three samples: schools pairing within 5 miles of each other (41 pairs), schools pairing within 7.5 miles of each other (105 pairs), and schools pairing within 10 miles of each other (181 pairs). Unfortunately, regression discontinuity did not yield any significant or consistent results. I attribute this to small sample sizes and the fact that each of these samples was not necessarily representative of the population. While the average rate of corporal punishment usage hovers around 5% for schools that use the practice, this rate was much lower in all three regression discontinuity samples. For the 41 pairs of schools within five miles of each other, the average rate of corporal punishment usage for schools that use the practice was 2.98%. For the sample of schools within 7.5 miles of each other, this rate bumped up to 4.10%. Finally, for the sample of schools within 10 miles of each other, the average value of *corp_punish_rate_all* in schools that allowed corporal punishment was 4.45%. These figures are noticeably different from the population mean (5%). Furthermore, the average value of *corp_punish_rate_all* in these samples only converged to the mean as schools got farther away from each other, meaning that location-specific omitted variables became less adequately controlled-for. Fuzzy regression discontinuity models where I included the instruments *city, suburb, town*¸ and *rural* similarly yielded no meaningful results.

Additionally, I originally thought Advanced Placement course participation rates to be a good measure of academic success for many of the same reasons that SAT participation is a good measure. However, instruments were endogenous for all regressions run with Advanced Placement course participation rates as the dependent variable. Overall, while none of the robustness checks I ran contributed to confirming the results I found with my main regressions, they also did not deny or cast doubt on what I find in my main body of results.

Results:

Antisocial Behavior Analysis

To analyze the effect of corporal punishment usage on levels of antisocial behavior within high schools, I use two-stage least squares instrumental variable regression with a school's out-of-school suspension rate as the dependent variable and the rate of corporal punishment usage as the independent variable of interest. I then also run the same model on the male, female, white, and black subsets of the student population. The results of these regressions for the 15 state sample are shown below in Table 3.

Table 3: Antisocial Behavior Regressions (15 State Sample)

	VARIABLES	(1) *susp_rate_all*	(2) *susp_rate_male*	(3) *susp_rate_female*	(4) *susp_rate_white*	(5) *susp_rate_black*
(1)	*corp_punish_rate_all*	-0.0283				
		(0.325)				
(2)	*corp_punish_rate_male*		0.0724			
			(0.277)			
(3)	*corp_punish_rate_female*			-0.346		
				(0.488)		
(4)	*corp_punish_rate_white*				-0.119	
					(0.664)	
(5)	*corp_punish_rate_black*					-0.0386
						(0.403)
(6)	F test for relevance	17.16	16.49	13.44	16.02	18.97
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(7)	J test for exogeneity	3.48	2.70	5.79	.69	14.67
		(p=.17)	(p=.26)	(p=.05)	(p=.71)	(p=.0007)
(8)	Observations	3,740	3,740	3,739	3,699	3,483
(9)	R-squared	0.359	0.320	0.358	0.124	0.114

Note. The coefficients for each *corp_punish_rate_all* variable are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments were endogenous for *susp_rate_black*, so this column is shaded gray.

As can be seen in columns 1 through 4 of Table 3, the rate of corporal punishment usage does not have significant effects on out-of-school suspensions in the 15 state sample. Note that if a column is shaded gray, instruments are either endogenous or irrelevant. In the case of the regression run on only black students, the instruments failed to pass the exogeneity test.

The same models as Table 3 are run using the corporal punishment dummy variable instead of corporal punishment rates, and coefficients from these regressions are presented in Table 4.

Table 4: Antisocial Behavior Dummy Variable Regressions (15 State Sample)

VARIABLES	(1) *susp_rate_ all*	(2) *susp_rate_ male*	(3) *susp_rate _female*	(4) *susp_rate_ white*	(5) *susp_rate_ black*
(1) *corp_punish_dummy*	0.00467	0.0162	-0.0158	-0.0103	0.00579
	(0.0200)	(0.0245)	(0.0176)	(0.0249)	(0.0390)
(2) F test for relevance	38.13	38.13	38.13	38.13	38.13
	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(3) J test for exogeneity	3.33	2.28	5.76	.45	15.06
	(p=.19)	(p=.32)	(p=.06)	(p=.80)	(p=.0005)
(4) Observations	3,740	3,740	3,739	3,699	3,483
(5) R-squared	0.359	0.313	0.368	0.128	0.115

Note. The coefficients for *corp_punish_dummy* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. Instruments were endogenous for *susp_rate_black*, so no coefficient is displayed.

As with the coefficients from regressions using corporal punishment rates, the coefficients in row 1 of Table 4 for the regressions using the dummy variable are insignificant. Once again, instruments did not pass the exogeneity test for the regression run on black students. The lack of significant coefficients here could be due to the fact that the 15 state sample includes several states where corporal punishment is legal but not used in schools.

I rerun the same regressions from Tables 3 and 4 on the eight state sample, which includes Alabama, Arkansas, Kentucky, Mississippi, Tennessee, Louisiana, Oklahoma, and Texas. The coefficients for the regressions using corporal punishment rates are shown in Table 5.

Table 5: Antisocial Behavior Regressions (8 State Sample)

	VARIABLES	(1) *susp_rate_ all*	(2) *susp_rate_ male*	(3) *susp_rate_ female*	(4) *susp_rate_ white*	(5) *susp_rate_ black*
(1)	*corp_punish_rate_all*	-0.749**				
		(0.326)				
(2)	*corp_punish_rate_male*		-0.511*			
			(0.275)			
(3)	*corp_punish_rate_female*			-1.365***		
				(0.497)		
(4)	*corp_punish_rate_white*				-0.900	
					(0.692)	
(5)	*corp_punish_rate_black*					-0.903**
						(0.423)
(6)	F test for relevance	13.04	12.08	11.17	9.95	14.90
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(7)	J test for exogeneity	.51	.54	2.91	2.52	9.27
		(p=.77)	(p=.24)	(p=.23)	(p=.28)	(p=.01)
(8)	Total Effect	-3.77%	-4.08%	-2.74%		
(9)	Total Effect (standard dev.)	-.44	-.40	-.37		
(10)	Observations	2,067	2,067	2,066	2,038	1,885
(11)	R-squared	0.207	0.191	0.170		

Note. These regressions were run on the 8 state sample of AL, AR, KY, MS, TN, LA, OK, and TX. The coefficients for different versions of *corp_punish_rate_all* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments were endogenous for *susp_rate_black* and barely irrelevant for *susp_rate_white*, so these columns are shaded gray. The Total Effect shows the percentage point difference in out-of-school suspension rate that occurs when a school goes from using no corporal punishment to using the average amount for a school that utilizes the practice. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

As can be seen in Column 4 of Table 5, the instruments fail relevance testing for the regression including only white students. Additionally, the instruments fail exogeneity testing for the regression run on black students in Column 5. In rows 1 through 3 of Table 5, coefficients for different versions of *corp_punish_rate_all* are negative and statistically significant to varying degrees, which suggests that the use of corporal punishment in high schools actually *decreases* the rate of out-of-school suspensions. If a school goes from using no corporal punishment to using the average level of corporal punishment for a school that utilizes the practice, the rate of out-of-school suspensions for all students goes down by 3.77 percentage points. This is a very surprising result that runs directly counter to my hypothesis. It should also be noted that the use of corporal punishment decreases the out-of-school suspension rate more for males (-4.08 percentage points) than females (-2.74 percentage points), but these decreases are similar in magnitude when measured in standard deviations (-.40 standard deviations for males versus -.37 standard deviations for females).

Regressions were also run on the same eight state sample using the corporal punishment dummy variable instead of using corporal punishment rates, and these regressions serve as a robustness check. In addition, using the dummy variable for corporal punishment helps show if the coefficients in the regressions using corporal punishment rates come from differences between schools that use corporal punishment and schools that do not or if these coefficients come mostly from differences within the group of schools that use corporal punishment. Column 5 of Table 6 shows that the instrumented variable model for black students fails the test for exogeneity, while the models shown in columns 1 through 4 are valid. In columns 1 through 4, coefficients are negative, and they are statistically significant in columns 1 through 3.

Table 6: Antisocial Behavior Dummy Variable Regressions (8 State Sample)

	VARIABLES	(1) *susp_rate_ all*	(2) *susp_rate_ male*	(3) *susp_rate_ female*	(4) *susp_rate_ white*	(5) *susp_rate_ black*
(1)	*corp_punish_dummy*	-0.0489**	-0.0487**	-0.0572***	-0.0384	-0.0981**
		(0.0201)	(0.0246)	(0.0180)	(0.0251)	(0.0401)
(2)	F test for relevance	30.12	30.12	30.12	30.12	30.12
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(3)	J test for exogeneity	.96	.54	2.79	2.28	10.26
		(p=.62)	(p=.76)	(p=.25)	(p=.32)	(p=.006)
(4)	Total Effect	-4.89%	-4.87%	-5.72%	-3.84%	
(5)	Total Effect (standard dev.)	-.57	-.47	-.76	-.43	
(6)	Observations	2,067	2,067	2,066	2,038	1,885
(7)	R-squared	0.324	0.299	0.292	0.107	0.065

Note. These regressions were run on the 8 state sample of AL, AR, KY, MS, TN, LA, OK, and TX. The coefficients for *corp_punish_dummy* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments were endogenous for *susp_rate_black*, so this column is shaded gray. The Total Effect shows the percentage point difference in out-of-school suspension rate that occurs when a school goes from not using corporal punishment to using corporal punishment. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Holding all else equal, if a school goes from goes from not using corporal punishment to using the practice, the out-of-school suspension rate for all students goes down by 4.89 percentage points. This difference is even larger than that calculated from the regressions using corporal punishment rates, where the out-of-school suspension rate for all students goes down by 3.77 percentage points when a school goes from using no corporal punishment to the average level for a school that does use physical punishment. This means that the effect of corporal punishment usage on antisocial behavior in high schools likely comes from differences between schools that use the practice and those that do not (and not from differences within the body of schools that use corporal punishment). Unfortunately, this is not a notion that I was able to test

by running regressions solely on schools that use corporal punishment, because instruments were irrelevant across the board for this sample.

When regressions were run on the nine state sample (the eight state sample plus Georgia), the negative, significant effects of corporal punishment on out-of-school suspensions held. These regressions are shown in Appendix E.

The strong negative effect of corporal punishment usage on out-of-school suspension rates is surprising, and it runs directly against my hypothesis. This negative effect could be for several reasons. It could be that administrators use corporal punishment instead of suspending students, though I doubt this explains all of the aforementioned effects. It also could be that students do not misbehave as much for fear of physical punishment, and in the process they are suspended at lower rates. Most likely, a combination of these two factors outweighs any increase in antisocial behavior that the use of corporal punishment causes. However, to get a broader picture of corporal punishment's effects on students, analysis of its possible academic ramifications is of paramount importance.

Academic Performance Analysis

To analyze the effect of corporal punishment usage on SAT participation rates, I use two-stage least squares instrumental variable regression with the SAT participation rate as the dependent variable and the rate of corporal punishment usage as the independent variable of interest. I then run the same model on male, female, white, and black subsets of the student population. The results of these regressions when they were run on the 15 state sample are shown below in Table 7. Note that *SAT_partipation_all* and all race/gender-based variants of the

variable have *participation* shortened to *part* in all tables below (e.g. *SAT_participation_female* becomes *SAT_part_female*). All instrumental variable models in Table 7 pass tests for instrumental relevance and exogeneity.

Table 7: Academic Performance Regressions (15 State Sample)

	VARIABLES	(1) *SAT_part_ all*	(2) *SAT_part_ male*	(3) *SAT_part _female*	(4) *SAT_part_ white*	(5) *SAT_part_ black*
(1)	*corp_punish_rate_all*	-3.428***				
		(0.840)				
(2)	*corp_punish_rate_male*		-2.537***			
			(0.614)			
(3)	*corp_punish_rate_female*			-6.065***		
				(1.525)		
(4)	*corp_punish_rate_white*				-5.580***	
					(1.862)	
(5)	*corp_punish_rate_black*					-2.033***
						(0.690)
(6)	F test for relevance	17.16	16.49	13.44	16.02	18.97
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(7)	J test for exogeneity	1.08	1.02	.42	.24	.69
		(p=.58)	(p=.60)	(p=.81)	(p=.89)	(p=.71)
(8)	Total Effect	-17.45%	-20.30%	-12.86%	-24.44%	-13.97%
(9)	Total Effect (standard dev.)	-1.36	-1.66	-.92	-1.63	-.78
(10)	Observations	3,740	3,740	3,739	3,698	3,482

Note. These regressions were run on the 15 state sample of states where corporal punishment in schools is allowed. The coefficients of different versions of *corp_punish_rate_all* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. The Total Effect shows the percentage point difference in SAT participation rate that occurs when a school goes from using no corporal punishment to using the average amount for a school that utilizes the practice. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

For the whole student population (column 1) and for all subsets of it (columns 2 through 5), the effect of corporal punishment on SAT participation rates is significantly negative. For the

student body as a whole, the SAT participation rate goes down by 17.45 percentage points if a school goes from using no corporal punishment to using the average level for a school that uses the practice. This is a change of -1.36 standard deviations, and it is significant at a level of 1%. To put this further into perspective, the average SAT participation rate in all three samples is around 20%, because only seniors and some juniors take the SAT. So, to have a change of 17.45 percentage points is dramatic. It is also worth noting that males are more adversely impacted by the use of corporal punishment than females (a change of -20.30 percentage points for males versus -12.86 percentage points for females), as one would expect. Additionally, if a school goes from using no corporal punishment to using the average level for a school that uses the practice, the decrease in SAT participation for white students (a change of -24.44 percentage points) is much stronger than the decrease for black students (-13.97 percentage points).

The same models from Table 7 are run using the corporal punishment dummy variable, and the coefficients for these regressions are shown in Table 8. As in Table 7, all instrumental variables pass tests for exogeneity and relevance in Table 8.

Table 8: Academic Performance Dummy Variable Regressions (15 State Sample)

	VARIABLES	(1) *SAT_part_all*	(2) *SAT_part_male*	(3) *SAT_part_female*	(4) *SAT_part_white*	(5) *SAT_part_black*
(1)	*corp_punish_dummy*	-0.209***	-0.217***	-0.208***	-0.192***	-0.195***
		(0.0437)	(0.0428)	(0.0474)	(0.0520)	(0.0598)
(2)	F test for relevance	38.13	38.13	38.13	38.13	38.13
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(3)	J test for exogeneity	2.79	4.35	2.73	3.27	1.02
		(p=.25)	(p=.11)	(p=.25)	(p=.19)	(p=.60)
(4)	Total Effect	-20.90%	-21.70%	-20.80%	-19.20%	-19.50%
(5)	Total Effect (standard dev.)	-1.63	-1.77	-1.48	-1.28	-1.10
(6)	Observations	3,740	3,740	3,739	3,698	3,482

Note. These regressions were run on the 15 state sample of states where corporal punishment in schools is allowed. The coefficients of *corp_punish_dummy* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. The Total Effect shows the percentage point difference in SAT participation rate that occurs when a school goes from not using corporal punishment to using corporal punishment. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Here, the difference in SAT participation rate between a school that uses corporal punishment and one that does not use it is even larger, at -20.90 percentage points. This means (as it did for the antisocial behavior regressions) that the coefficients in these regressions are likely picking up the differences between schools that use corporal punishment and those that do not (instead of differences within schools that use the practice). In Table 8, the effect of corporal punishment on SAT participation hovers around -20 percentage points for males, females, white students, and black students (columns 2 through 5).

I also ran instrumental variable regressions examining the effect of corporal punishment usage rates on SAT participation for the eight state sample, and the results of these regressions are shown in Table 9. Instruments are relevant and exogenous for all models in Table 9.

Table 9: Academic Performance Regressions (8 State Sample)

	VARIABLES	(1) *SAT_part_all*	(2) *SAT_part_male*	(3) *SAT_part_female*	(4) *SAT_part_white*	(5) *SAT_part_black*
(1)	*corp_punish_rate_all*	-3.050***				
		(0.787)				
(2)	*corp_punish_rate_male*		-2.196***			
			(0.568)			
(3)	*corp_punish_rate_female*			-5.195***		
				(1.417)		
(4)	*corp_punish_rate_white*				-5.195***	
					(1.895)	
(5)	*corp_punish_rate_black*					-2.281***
						(0.717)
(6)	F test for relevance	13.04	12.08	11.17	9.95	14.90
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(7)	J test for exogeneity	1.11	1.68	.90	2.85	.21
		(p=.57)	(p=.43)	(p=.64)	(p=.24)	(p=.90)
(8)	Total Effect	-15.34%	-17.52%	-10.44%	-22.81%	-15.12%
(9)	Total Effect (standard dev.)	-1.10	-1.32	-.69	-1.40	-.77
(10)	Observations	2,067	2,067	2,066	2,037	1,884

Note. These regressions were run on the 8 state sample (AL, AR, KY, MS, TN, LA, OK, TX). The coefficients of different versions of *corp_punish_rate_all* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. The Total Effect shows the percentage point difference in SAT participation rate that occurs when a school goes from using no corporal punishment to using the average amount for a school that utilizes the practice. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Strong, statistically significant negative coefficients persist when the sample is narrowed to eight states, though the coefficients are noticeably smaller than those found in the regressions

run on the 15 state sample. In the eight state sample, if a school goes from using no corporal punishment to using the average level for a school that physically punishes students, its SAT participation rate goes down by 15.34 percentage points. As with the corporal punishment rate regressions which were run on the full 15 state sample (Table 7), males are more adversely affected in SAT participation rate than females (a change of -17.52 percentage points for males versus -10.44 percentage points for females). Additionally, white students are once again more adversely impacted than black students (-22.81 percentage points for white students versus -15.12 percentage points for black students). The figures listed for "Total Effect" in Table 9 for the eight state sample are likely more accurate than those listed in Table 7 for the 15 state sample, because I am able to control for more cultural factors by narrowing the sample size down to what is typically considered the "Deep South."

Regressions using the corporal punishment dummy variable were also run on the eight state sample, and the coefficients for these regressions are shown in Table 10. Instruments do not pass exogeneity tests for the models run on male students and white students. However, models for all students, female students, and black students are valid. Like the dummy variable regressions for the 15 state sample in Table 8, the coefficients on the corporal punishment dummy variable are significantly negative in the valid models (columns 1, 3, and 5 of Table 10).

Table 10: Academic Performance Dummy Variable Regressions (8 State Sample)

	VARIABLES	(1) *SAT_part_all*	(2) *SAT_part_male*	(3) *SAT_part_female*	(4) *SAT_part_white*	(5) *SAT_part_black*
(1)	*corp_punish_dummy*	-0.192*** (0.0449)	-0.194*** (0.0435)	-0.194*** (0.0489)	-0.183*** (0.0540)	-0.227*** (0.0635)
(2)	F test for relevance	30.12 (p=.0000)	30.12 (p=.0000)	30.12 (p=.0000)	30.12 (p=.0000)	30.12 (p=.0000)
(3)	J test for exogeneity	4.41 (p=.11)	6.51 (p=.04)	3.03 (p=.22)	7.05 (p=.03)	1.14 (p=.56)
(4)	Total Effect	-19.20%		-19.40%		-22.70%
(5)	Total Effect (standard dev.)	-1.38		-1.27		-1.16
(6)	Observations	2,067	2,067	2,066	2,037	1,884

Note. These regressions were run on the 8 state sample (AL, AR, KY, MS, TN, LA, OK, TX). The coefficients of *corp_punish_dummy* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments are endogenous for *SAT_part_male* and *SAT_part_white*, so columns are shaded gray. The Total Effect shows the percentage point difference in SAT participation rate that occurs when a school goes from not using corporal punishment to using corporal punishment. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Here, if a school goes from using no corporal punishment to using the practice, the SAT participation rate for all students goes down by 19.20 percentage points. This effect is of virtually the same magnitude for the female section of the student body as it is for the whole student body. And while instruments were endogenous in the regression for males in this case, it must follow that the SAT participation rate of males goes down by a similar amount as it does for the female student body, because the coefficients of the corporal punishment dummy are very similar for regressions run on the whole student body and on just females. Black students are slightly more adversely affected than other students in Table 10, as the SAT participation rate of black students goes down by 22.70 percentage points if a school goes from not using corporal punishment to utilizing the practice.

Finally, I added Georgia to the eight state sample and ran regressions on this nine state sample, because Georgia is the only state omitted from the eight state sample where corporal punishment is still used in high schools frequently. Results from the eight state sample persisted in this nine state sample, and regression tables for this analysis are shown in Appendix E.

Across all three samples and using corporal punishment rates and a dummy variable, I found very large, negative, and statistically significant coefficients for corporal punishment's effects on SAT participation. This is in line with my hypothesis that the use of corporal punishment decreases SAT participation, but the magnitude and significance of these coefficients is far beyond what I expected going into this analysis. Additionally, it can be inferred from the results of my antisocial behavior analysis that this decrease in SAT participation does not come primarily from the channel I originally hypothesized. I thought that an increase in corporal punishment would lead to an increase in antisocial behavior, which would then adversely affect academic achievement. In fact, I find that the use of corporal punishment decreases antisocial behavior at a student body level, but it also heavily adversely impacts academic achievement. So, the second channel through which I hypothesized corporal punishment could cause lower academic achievement—the idea that corporal punishment creates a toxic learning environment and disengages students—is seemingly validated by my findings.

Conclusion:

If a school is looking into whether or not to use corporal punishment in disciplining its students, there is surely a cost-benefit analysis which it uses to weigh the benefit of decreased antisocial behavior against the loss in academic performance such a policy could impose. Using regressions from the eight state sample (found in Tables 5 and 9), this tradeoff can be described as follows: if a school goes from using no corporal punishment to using the average level for a school that utilizes the practice, the rate of out-of-school suspensions goes down by 3.77 percentage points, but the SAT participation rate of the student body goes down by 15.34 percentage points. If one is to interpret these variables as good measures of antisocial behavior and academic performance (which this paper holds to be the case), these changes can be interpreted as a decrease in antisocial behavior of .44 standard deviations along with a decrease of 1.10 standard deviations in overall academic performance. It is hard to believe that, given this tradeoff, any sensible school administrator would choose to allow the usage of corporal punishment to discipline students. Its effect on academic performance is simply too profound to ignore, despite the evidence that it might also decrease levels of antisocial behavior in the student body.

Though I am confident in the results presented in this paper, there is future research that could be done to solidify the academic and social ramifications of corporal punishment. The second wave of this research can be conducted in roughly two years, when the Office of Civil Rights releases the Civil Rights Data Collection dataset for the 2015-2016 school year. Once this dataset comes out, any researcher could create the same variables that I did by using the CRDC dataset along with the Common Core dataset for the corresponding school year. With two waves of data, this researcher could use entity fixed effects to find out what happens to different

markers of academic achievement and social well-being when schools increase or decrease their usage of corporal punishment.

If any state in the nine state sample used in this paper outlaws the practice of in-school corporal punishment, another possible offshoot of this paper emerges. It would be easy to trace the changes in different measures of academic and social well-being after the passing of this law, so this type of analysis would be ideal to answer the question of in-school corporal punishment's effects. However, until this type of legislation is enacted in the right place, this paper's results support the conclusion that the usage of corporal punishment should not be allowed in high schools.

While my research indicates that corporal punishment can decrease the occurrence of antisocial behavior within a student body, the price to be paid from an academic achievement perspective is simply too severe. Thus, to readers who came into this paper incredulous that in-school corporal punishment is still utilized on a wide scale, this paper will not change any minds. However, if there is anybody reading for whom in-school corporal punishment is the status quo, be wary that this practice comes at a price that is not worth paying under any circumstances.

References

Carr, S. (2014, April 8). Why Are Black Students Facing Corporal Punishment in Public Schools? The Nation. Retrieved November 6, 2016, from https://www.thenation.com/article/why-are-black-students-facing-corporal-punishment-publicschools/

Fuller, J. M. (2009). The Science and Statistics Behind Spanking Suggest that Laws Allowing Corporal Punishment Are in the Best Interests of the Child. Akron Law Review, 42(1), 243-317. Retrieved October 23, 2016, from http://ideaexchange.uakron.edu/cgi/viewcontent.cgi?article=1192&context=akronlawreview

Grogan-Kaylor, A. (2004, September). The effect of corporal punishment on antisocial behavior in children. Social Work Research, 28(3), 153-162. Retrieved from https://www.researchgate.net/profile/Andrew_Grogan-Kaylor/publication/261705604_The_effect_of_corporal_punishment_on_antisocial_behavior_in_children/links/5502fa020cf231de076fcae8.pdf.

Larzelere, R. E., & Kuhn, B. R. (2005, March). Comparing Child Outcomes of Physical Punishment and Alternative Disciplinary Tactics: A Meta-Analysis. Clinical Child and Family Psychology Review, 8(1), 1-37. doi:10.1007/s10567-005-2340-z

Maguin, E., & Loeber, R. (1996). Academic Performance and Delinquency. Crime and Justice, 20, 145-264. Retrieved from http://www.jstor.org/stable/1147645

McEvoy, A., & Welker, R. (2000, Fall). Antisocial Behavior, Academic Failure, and School Climate: A Critical Review. Journal of Emotional and Behavioral Disorders, 8(3), 130-141. doi:10.1177/106342660000800301

Mendez, L., & Knoff, H. (2003). Who Gets Suspended from School and Why: A Demographic Analysis of Schools and Disciplinary Infractions in a Large School District. Education and Treatment of Children, 26(1), 30-51. Retrieved from http://www.jstor.org/stable/42900535

Sparks, S. D., & Harwin, A. (2016, August 23). Corporal Punishment Use Found in Schools in 21 States. Education Week. Retrieved November 6, 2016, from http://www.edweek.org/ew/articles/2016/08/23/corporal-punishment-use-found-in-schoolsin.html

Straus, M. A., Sugarman, D. B., & Giles-Sims, J. (1997, August). Spanking by Parents and Subsequent Antisocial Behavior of Children. Archives of Pediatric and Adolescent Medicine, 151, 761-767. Retrieved from http://healingreligion.com/2490/html/CP24.pdf

Straus, M. A., & Stewart, J. H. (1999). Corporal punishment by american parents: National data on prevalence, chronicity, severity, and duration, in relation to child and family characteristics. Clinical Child and Family Psychology Review, 2(2), 55-70. doi:http://dx.doi.org/10.1023/A:1021891529770

Data Sources

Office for Civil Rights. (2016). 2013-14 Civil Rights Data Collection Public Use Data File. OCR_2016001. Date accessed: 2016-10-16.

National Center for Education Statistics. (2016). NCES Common Core of Data Public Elementary/Secondary School Universe Survey: School Year 2013-14. NCES 2016-150rev. Data accessed: 2016-12-14.

The University of North Carolina, Wilmington. (2017). Historical American Lynching Data Collection Project. Data accessed: 2017-02-01.

Appendix A: Variable List.

Variable Name	**Description**
state	state
schoolnumber	National Center of Education Statistics School ID
latitude	latitude of school
longitude	longitude of school
corp_punish_dummy	Does the school use corporal punishment?
midwest	school located in the midwest
south	school located in the south
west	school located in the west
northeast	school located in the northeast
district	name of school district
magnet	Does the school have a magnet program?
schoolname	name of school
fundingperstudent	salary expenditures (state+local) on all personnel per student
student_teacher_ratio	number of students per full time equivalency of teachers
percentwhite	percent of student body that is white
percentblack	percent of student body that is black
percentasian	percent of student body that is asian
percenthispanic	percent of student body that is hispanic
percentother	percent of student body that is of other heritage
enrollwhite	number of white students enrolled
enrollblack	number of black students enrolled
enrollasian	number of asian students enrolled
enrollhispanic	number of hispanic students enrolled
enrollother	number of other heritage students enrolled
enrollall	number of total students enrolled
SAT_participation_male	percent of all males taking the SAT
SAT_participation_female	percent of all females taking the SAT
SAT_participation_white	percent of white students taking the SAT
SAT_participation_black	percent of black students taking the SAT
SAT_participation_all	percent of all students taking the SAT
corp_punish_rate_male	percent of all males subjected to corporal punishment
corp_punish_rate_female	percent of all females subjected to corporal punishment
corp_punish_rate_white	percent of white students subjected to corporal punishment
corp_punish_rate_black	percent of black students subjected to corporal punishment
corp_punish_rate_all	percent of all students subjected to corporal punishment
issallm	percent of all males with at least one in school suspension
issallf	percent of all females with at least one in school suspension
isswhite	percent of white students with at least one in school suspension

issblack	percent of black students with at least one in school suspension
issall	percent of all students with at least one in school suspension
susp_rate_male	percent of all males with at least one out of school suspension
susp_rate_female	percent of all females with at least one out of school suspension
susp_rate_white	percent of white students with at least one out of school suspension
susp_rate_black	percent of black students with at least one out of school suspension
susp_rate_all	percent of all students with at least one out of school suspension
zipcode	zip code of school
freelunchall	the percentage of students who receive free lunches
reducedlunchall	the percentage of students who receive reduced-price lunches
SESall	the percentage of students who receive free or reduced-price lunches
lnfundingperstudent	ln of funding per student
lnenrollall	ln of overall student enrollment
city	school is located in a city
suburb	school is located in a suburb
town	school is located in a town
rural	school is located in a rural area
lynching	number of lynchings recorded that occurred in the same zipcode as the school

Appendix B: Descriptive Statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
latitude	3834.00	33.53	3.21	24.55	44.87
longitude	3834.00	-90.51	8.53	-114.76	-75.67
corp_punish_dummy	3832.00	0.15	0.36	0.00	1.00
midwest	3834.00	0.06	0.24	0.00	1.00
south	3834.00	0.88	0.32	0.00	1.00
west	3834.00	0.06	0.23	0.00	1.00
northeast	3834.00	0.00	0.00	0.00	0.00
magnet	3834.00	0.10	0.30	0.00	1.00
fundingperstudent	3823.00	4631.93	3531.70	0.00	50000.00
student_teacher_ratio	3750.00	19.36	38.68	0.00	916.67
percentwhite	3834.00	0.52	0.30	0.00	1.00
percentblack	3834.00	0.22	0.26	0.00	1.00
percentasian	3834.00	0.02	0.04	0.00	0.59
percenthispanic	3834.00	0.19	0.24	0.00	1.00
percentother	3834.00	0.05	0.10	0.00	1.00
enrollwhite	3834.00	499.66	464.69	0.00	3112.00
enrollblack	3834.00	228.10	306.96	0.00	2197.00
enrollasian	3834.00	30.57	73.19	0.00	1207.00
enrollhispanic	3834.00	259.31	467.18	0.00	4048.00
enrollother	3834.00	40.00	78.63	0.00	1826.00
enrollall	3834.00	1057.65	803.59	2.00	5352.00
SAT_participation_male	3833.00	0.19	0.12	0.00	1.00
SAT_participation_female	3833.00	0.23	0.14	0.00	1.00
SAT_participation_white	3786.00	0.22	0.15	0.00	1.00
SAT_participation_black	3561.00	0.22	0.18	0.00	1.00
SAT_participation_all	3833.00	0.21	0.13	0.00	1.00
corp_punish_rate_male	3833.00	0.01	0.05	0.00	0.81
corp_punish_rate_female	3834.00	0.00	0.02	0.00	0.40
corp_punish_rate_white	3826.00	0.01	0.03	0.00	0.40
corp_punish_rate_black	3757.00	0.01	0.05	0.00	1.00
corp_punish_rate_all	3833.00	0.01	0.03	0.00	0.44
issallm	3833.00	0.14	0.12	0.00	0.78
issallf	3833.00	0.09	0.09	0.00	0.62
isswhite	3787.00	0.09	0.09	0.00	1.00
issblack	3562.00	0.17	0.17	0.00	1.00
issall	3833.00	0.12	0.10	0.00	0.67
susp_rate_male	3833.00	0.10	0.10	0.00	0.99
susp_rate_female	3833.00	0.06	0.07	0.00	0.83

susp_rate_white	3787.00	0.06	0.08	0.00	1.00
susp_rate_black	3562.00	0.12	0.13	0.00	1.00
susp_rate_all	3833.00	0.08	0.08	0.00	0.90
freelunchall	3826.00	0.45	0.22	0.00	1.00
reducedlunchall	3826.00	0.08	0.06	0.00	1.00
lunchall	3826.00	0.53	0.22	0.00	1.00
lnfundingperstudent	3823.00	8.28	0.69	0.00	10.82
lnenrollall	3834.00	6.59	0.98	0.69	8.59
city	3834.00	0.21	0.41	0.00	1.00
suburb	3834.00	0.23	0.42	0.00	1.00
town	3834.00	0.18	0.39	0.00	1.00
rural	3834.00	0.38	0.49	0.00	1.00
lynching	3834.00	143.95	164.71	0.00	538.00

Appendix C: Relevance Tests for Instrumental Variables.

Normal variable names are used in this table as opposed to the lengthened versions present in the main body of this paper. *corporalpunishment* is the corporal punishment dummy variable, whereas *corp_punish_rate_all* is shortened to *corpall*. *corp_punish_rate_male* is shortened to *corpallm*.

Relevance Test Results				
Sample	Dependent Variable	Instruments Tested	Relevance Test F Value	Instruments Relevant?
15 State Sample	*corpall*	*city, suburb, town, rural*	17.16	yes
	corpallm	*city, suburb, town, rural*	16.49	yes
	corpallf	*city, suburb, town, rural*	13.44	yes
	corpwhite	*city, suburb, town, rural*	16.02	yes
	corpblack	*city, suburb, town, rural*	18.97	yes
	corporalpunishment	*city, suburb, town, rural*	38.13	yes
8 State Sample	*corpall*	*city, suburb, town, rural*	13.04	yes
	corpallm	*city, suburb, town, rural*	12.08	yes
	corpallf	*city, suburb, town, rural*	11.17	yes
	corpwhite	*city, suburb, town, rural*	9.95	no
	corpblack	*city, suburb, town, rural*	14.9	yes
	corporalpunishment	*city, suburb, town, rural*	30.12	yes
9 State Sample	*corpall*	*city, suburb, town, rural*	14.99	yes
	corpallm	*city, suburb, town, rural*	14.14	yes
	corpallf	*city, suburb, town, rural*	12.49	yes
	corpwhite	*city, suburb, town, rural*	12.8	yes
	corpblack	*city, suburb, town, rural*	16.39	yes
	corporalpunishment	*city, suburb, town, rural*	33.45	yes
Schools that use corporal punishment	*corpall*	*city, suburb, town, rural*	4.74	no
	corpallm	*city, suburb, town, rural*	3.7	no
	corpallf	*city, suburb, town, rural*	6.25	no
	corpwhite	*city, suburb, town, rural*	2.83	no
	corpblack	*city, suburb, town, rural*	5.18	no
For Relevance F Tests, Null Hypothesis: $F < 10$; Alternative Hypothesis: $F \geq 10$.				
Controls included in these tests are racial composition of student body, student-teacher ratio, funding per student, overall enrollment, and percentage of students receiving free or assisted lunches.				

Appendix D: Exogeneity Tests for Instrumental Variables.

All exogeneity J-Test results are for the four geographic instruments city, suburb, town, and rural. J has a chi-2 distribution with degrees of freedom equal to the degree of overidentification (2 in all cases). For Exogeneity J Tests, J = mF, where m is the number of instruments (3 w/ 1 omitted). Null Hypothesis: J ≥ threshold chi-2 value; Alternative Hypothesis: J < threshold chi-2 value. If Null hypothesis is rejected at 5% significance, the instrument is exogenous. The threshold J value for all tests is 5.99.

Normal variable names are used in this table as opposed to the lengthened versions present in the main body of this paper. *corporalpunishment* is the corporal punishment dummy variable, whereas *corp_punish_rate_all* is shortened to *corpall*. *corp_punish_rate_male* is shortened to *corpallm*.

Exogeneity Test Results (15 state sample)					
Dependent Variable	Independent Variable of interest	Exogeneity Test F value	Exogeneity Test J Value	Exogeneity Test chi-2 threshold	Exogenous?
suspall	*corpall*	1.16	3.48	5.99	yes
SATall	*corpall*	0.36	1.08	5.99	yes
suspallm	*corpallm*	0.9	2.7	5.99	yes
SATallm	*corpallm*	0.34	1.02	5.99	yes
suspallf	*corpallf*	1.93	5.79	5.99	yes
SATallf	*corpallf*	0.14	0.42	5.99	yes
suspwhite	*corpwhite*	0.23	0.69	5.99	yes
SATwhite	*corpwhite*	0.08	0.24	5.99	yes
suspblack	*corpblack*	4.89	14.67	5.99	no
SATblack	*corpblack*	0.23	0.69	5.99	yes
suspall	*corporalpunishment*	1.11	3.33	5.99	yes
SATall	*corporalpunishment*	0.93	2.79	5.99	yes
suspallm	*corporalpunishment*	0.76	2.28	5.99	yes
SATallm	*corporalpunishment*	1.45	4.35	5.99	yes
suspallf	*corporalpunishment*	1.92	5.76	5.99	yes
SATallf	*corporalpunishment*	0.91	2.73	5.99	yes
suspwhite	*corporalpunishment*	0.15	0.45	5.99	yes
SATwhite	*corporalpunishment*	1.09	3.27	5.99	yes
suspblack	*corporalpunishment*	5.02	15.06	5.99	no
SATblack	*corporalpunishment*	0.34	1.02	5.99	yes

Exogeneity Test Results (8 state sample)					
Dependent Variable	Independent Variable of interest	Exogeneity Test F value	Exogeneity Test J Value	Exogeneity Test chi-2 threshold	Exogenous?
suspall	*corpall*	0.17	0.51	5.99	yes
SATall	*corpall*	0.37	1.11	5.99	yes
suspallm	*corpallm*	0.18	0.54	5.99	yes
SATallm	*corpallm*	0.56	1.68	5.99	yes
suspallf	*corpallf*	0.97	2.91	5.99	yes
SATallf	*corpallf*	0.3	0.9	5.99	yes
suspwhite	*corpwhite*	0.84	2.52	5.99	yes
SATwhite	*corpwhite*	0.95	2.85	5.99	yes
suspblack	*corpblack*	3.09	9.27	5.99	no
SATblack	*corpblack*	0.07	0.21	5.99	yes
suspall	*corporalpunishment*	0.32	0.96	5.99	yes
SATall	*corporalpunishment*	1.47	4.41	5.99	yes
suspallm	*corporalpunishment*	0.18	0.54	5.99	yes
SATallm	*corporalpunishment*	2.17	6.51	5.99	no
suspallf	*corporalpunishment*	0.93	2.79	5.99	yes
SATallf	*corporalpunishment*	1.01	3.03	5.99	yes
suspwhite	*corporalpunishment*	0.76	2.28	5.99	yes
SATwhite	*corporalpunishment*	2.35	7.05	5.99	no
suspblack	*corporalpunishment*	3.42	10.26	5.99	no
SATblack	*corporalpunishment*	0.38	1.14	5.99	yes

Exogeneity Test Results (9 state sample)					
Dependent Variable	Independent Variable of interest	Exogeneity Test F value	Exogeneity Test J Value	Exogeneity Test chi-2 threshold	Exogenous?
suspall	*corpall*	0.2	0.6	5.99	yes
SATall	*corpall*	0.76	2.28	5.99	yes
suspallm	*corpallm*	0.13	0.39	5.99	yes
SATallm	*corpallm*	1.14	3.42	5.99	yes
suspallf	*corpallf*	0.99	2.97	5.99	yes
SATallf	*corpallf*	0.53	1.59	5.99	yes
suspwhite	*corpwhite*	1.25	3.75	5.99	yes
SATwhite	*corpwhite*	1.28	3.84	5.99	yes
suspblack	*corpblack*	2.79	8.37	5.99	no
SATblack	*corpblack*	0.26	0.78	5.99	yes
suspall	*corporalpunishment*	0.73	2.19	5.99	yes
SATall	*corporalpunishment*	2.34	7.02	5.99	no
suspallm	*corporalpunishment*	0.46	1.38	5.99	yes
SATallm	*corporalpunishment*	3.73	11.19	5.99	no
suspallf	*corporalpunishment*	1.51	4.53	5.99	yes
SATallf	*corporalpunishment*	1.31	3.93	5.99	yes
suspwhite	*corporalpunishment*	0.71	2.13	5.99	yes
SATwhite	*corporalpunishment*	3.64	10.92	5.99	no
suspblack	*corporalpunishment*	3.93	11.79	5.99	no
SATblack	*corporalpunishment*	0.73	2.19	5.99	yes

Appendix E: Regressions Run on the Nine state sample.

Antisocial Behavior

	Antisocial Behavior Regressions (9 State Sample)					
	VARIABLES	(1) *susp_rate_ all*	(2) *susp_rate_ male*	(3) *susp_rate_ female*	(4) *susp_rate_ white*	(5) *susp_rate_ black*
(1)	*corp_punish_rate_all*	-0.841*** (0.323)				
(2)	*corp_punish_rate_male*		-0.593** (0.271)			
(3)	*corp_punish_rate_female*			-1.526*** (0.504)		
(4)	*corp_punish_rate_white*				-0.936 (0.697)	
(5)	*corp_punish_rate_black*					-0.887** (0.401)
(6)	F test for relevance	14.99 (p=.0000)	14.14 (p=.0000)	12.49 (p=.0000)	12.80 (p=.0000)	16.39 (p=.0000)
(7)	J test for exogeneity	.60 (p=.74)	.39 (p=.82)	2.97 (p=.23)	3.75 (p=.15)	8.37 (p=.01)
(8)	Total Effect	-4.23%	-4.71%	-3.14%	-4.07%	
(9)	Total Effect (standard dev.)	-.49	-.46	-.42	-.45	
(10)	Observations	2,395	2,395	2,394	2,356	2,212
(11)	R-squared	0.213	0.188	0.173		

Note. These regressions were run on the 9 state sample of AL, AR, KY, MS, TN, LA, OK, GA, and TX. The coefficients for different versions of *corp_punish_rate_all* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments were endogenous for *susp_rate_black*, so this column is shaded gray. The Total Effect shows the percentage point difference in out-of-school suspension rate that occurs when a school goes from using no corporal punishment to using the average amount for a school that utilizes the practice. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Antisocial Behavior Dummy Variable Regressions (9 State Sample)

	VARIABLES	(1) *susp_rate_ all*	(2) *susp_rate_ male*	(3) *susp_rate_ female*	(4) *susp_rate_ white*	(5) *susp_rate_ black*
(1)	*corp_punish_dummy*	-0.0493***	-0.0500**	-0.0557***	-0.0469*	-0.0757**
		(0.0188)	(0.0228)	(0.0169)	(0.0247)	(0.0361)
(2)	F test for relevance	33.45	33.45	33.45	33.45	33.45
		(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(3)	J test for exogeneity	2.19	1.38	4.53	2.13	11.79
		(p=.33)	(p=.50)	(p=.10)	(p=.34)	(p=.003)
(4)	Total Effect	-4.93%	-5.00%	-5.57%	-4.69%	
(5)	Total Effect (standard dev.)	-.57	-.49	-.74	-.52	
(6)	Observations	2,395	2,395	2,394	2,356	2,212
(7)	R-squared	0.354	0.324	0.333	0.100	0.103

Note. These regressions were run on the 9 state sample of AL, AR, KY, MS, TN, LA, OK, GA, and TX. The coefficients for *corp_punish_dummy* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments were endogenous for *susp_rate_black*, so this column is shaded gray. The Total Effect shows the percentage point difference in out-of-school suspension rate that occurs when a school goes from not using corporal punishment to using corporal punishment. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Academic Performance

Academic Performance Regressions (9 State Sample)						
	VARIABLES	(1) *SAT_part_ all*	(2) *SAT_part_ male*	(3) *SAT_part _female*	(4) *SAT_part_ white*	(5) *SAT_part_ black*
(1)	*corp_punish_rate_all*	-2.795*** (0.732)				
(2)	*corp_punish_rate_male*		-2.096*** (0.532)			
(3)	*corp_punish_rate_female*			-4.591*** (1.322)		
(4)	*corp_punish_rate_white*				-4.939*** (1.776)	
(5)	*corp_punish_rate_black*					-2.065*** (0.661)
(6)	F test for relevance	14.99 (p=.0000)	14.14 (p=.0000)	12.49 (p=.0000)	12.80 (p=.0000)	16.39 (p=.0000)
(7)	J test for exogeneity	2.28 (p=.32)	3.42 (p=.18)	1.59 (p=.45)	3.84 (p=.15)	.78 (p=.68)
(8)	Total Effect	-14.06%	-16.66%	-9.46%	-21.48%	-13.71%
(9)	Total Effect (standard dev.)	-1.05	-1.30	-.64	-1.34	-.73
(10)	Observations	2,395	2,395	2,394	2,355	2,211

Note. These regressions were run on the 9 state sample (AL, AR, KY, MS, TN, LA, OK, TX, GA). The coefficients of different versions of *corp_punish_rate_all* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. The Total Effect shows the percentage point difference in SAT participation rate that occurs when a school goes from using no corporal punishment to using the average amount for a school that utilizes the practice. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.

Academic Performance Dummy Variable Regressions (9 State Sample)					
VARIABLES	(1) *SAT_part_all*	(2) *SAT_part_male*	(3) *SAT_part_female*	(4) *SAT_part_white*	(5) *SAT_part_black*
(1) *corp_punish_dummy*	-0.162***	-0.168***	-0.162***	-0.145***	-0.195***
	(0.0402)	(0.0390)	(0.0439)	(0.0491)	(0.0571)
(2) F test for relevance	33.45	33.45	33.45	33.45	33.45
	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)	(p=.0000)
(3) J test for exogeneity	7.02	11.19	3.93	10.92	2.19
	(p=.03)	(p=.004)	(p=.14)	(p=.004)	(p=.33)
(4) Total Effect			-16.20%		-19.50%
(5) Total Effect (standard dev.)			-1.10		-1.04
(6) Observations	2,395	2,395	2,394	2,355	2,211

Note. These regressions were run on the 9 state sample (AL, AR, KY, MS, TN, LA, OK, TX, GA). The coefficients of *corp_punish_dummy* are displayed with standard errors in parentheses below them. Regressions control for student-teacher ratio, funding per student, overall enrollment, student body racial composition, and the portion of students receiving free or assisted lunches. The threshold for the relevance test F statistic is 10 in all cases, and the J statistic must have a p value of more than .05 for instrumental variables to be considered exogenous. Instruments were endogenous for *SAT_part_all*, *SAT_part_male*, and *SAT_part_white*, so columns are shaded gray. The Total Effect shows the percentage point difference in SAT participation rate that occurs when a school goes from not using corporal punishment to using corporal punishment. Total Effect (standard dev.) converts this Total Effect figure into standard deviations of the dependent variable.