Table of Contents

Introduction: Punitive Zero Tolerance Policies . . . and negative outcomes associated with them

School Suspensions Generally

The use of exclusionary discipline policies in American schools has become increasingly prevalent over the last three decades as school disciplinary policies have shifted over time. A major turn towards punitive school discipline occurred in the late 1980's to early 1990's, the shift to the punitive zero tolerance policies prevalent today took hold.[1] The move towards harsh zero tolerance policies was initially in response to school administrator fears over increasing violence, drug-related problems, and gang activity.[2] Eventually, it was inscribed into federal law with the Gun-Free Schools Act of 1994 and its later versions which conditioned federal funding on mandatory one-year expulsion of students committing specified firearm offenses.[3] Even though the federal call for mandatory expulsion was limited to specific firearms and allowed for some discretion in application of the policy,[4] however, many states have expanded the use of zero tolerance widely beyond this reach.[5] Today, the potential applications for zero tolerance policies are numerous but vary significantly across states.[6] Stories of nonsensical applications of these policies are numerous,[7] as the policies are in wide-spread use around the country.[8]

The Data

One particular alarming outcome of the growing use of zero tolerance policies is that more and more children are being excluded from school, with out-of-school suspension "among the most widely used techniques."[9]

---

[1]   Russell J. Skiba & Kimberly Knesting, *Zero Tolerance, Zero Evidence: An Analysis of School Disciplinary Practice*, 92 NEW DIRECTIONS FOR YOUTH DEV. 17, (2001). Zero Tolerance disciplinary policies in schools have been broadly defined to mean a "philosophy or policy that mandates the application of predetermined consequences, most often severe and punitive in nature, that are intended to be applied regardless of the gravity of behavior, mitigating circumstances, or situational context." American Psychological Association, Zero Tolerance Task Force, *Are Zero Tolerance Policies Effective in the Schools? An Evidentiary Review and Recommendations*, 63 AM. PSYCHOLOGIST 852, 852 (2008) [hereinafter: APA Task Force].

[2]   Skiba & Knesting, *supra* note 5, at 19.

[3]   Hanson, *supra* note 3, at 303-04.

[4]   *Id*. at 305, 309.

[5]   *See, e.g.*, *id.* at 307; Skiba & Knesting, *supra* note 5, at 19; Christina L. Anderson, *Double Jeopardy: The Modern Dilemma for Juvenile Justice*, 152 U. PA. L. REV. 1181, 1185 (2004).

[6]   *See, e.g.*, ADVANCEMENT PROJECT & THE CIVIL RIGHTS PROJECT, HARVARD UNIV., OPPORTUNITIES SUSPENDED: THE DEVASTATING CONSEQUENCES OF ZERO TOLERANCE SCHOOL DISCIPLINE POLICIES (2000), *available at* http://civilrightsproject.ucla.edu/research/k-12-education/school-discipline/opportunities-suspended-the-devastating-consequences-of-zero tolerance-and-school-discipline-policies/crp-opportunities-suspended-zero tolerance-2000.pdf [hereinafter: OPPORTUNITIES SUSPENDED].

[7]   *See, e.g.*, *id.* at 4-7; JUDITH A. BROWNE, ADVANCEMENT PROJECT, DERAILED: THE SCHOOLHOUSE TO JAILHOUSE TRACK 11 (2003), *available at* http://www.advancementproject.org/digital-library/publications/derailed-the-schoolhouse-to-jailhouse-track [hereinafter DERAILED].

[8]   Skiba & Knesting, *supra* note 5, at 20 (citing a U.S. Department of Education report which stated that already in 1996-97 "94 percent of all schools ha[d] zero tolerance policies for weapons or firearms, 87 percent for alcohol, and 79 percent for violence or tobacco").

[9]   Skiba & Knesting, *supra* note 5, at 28*. See also* Russell J. Skiba et al., *Office Referrals and Suspension: Disciplinary Intervention in Middle Schools*, 20 EDUCATION AND TREATMENT OF CHILDREN 295, page not available in format I downloaded (1997) (finding Suspension to be by far the most common administrative disciplinary

2

For example, the number of primary and secondary public school students that were suspended from school at least once during a particular school year has risen from an estimated 1.7 million in 1974[10] to approximately 3.1 million in 1998[11] and 3.3 million in 2006,[12] a number which represents seven percent of all those students.[13] Somewhat counterintuitively, however, these high numbers of suspensions that exclude children from educational instruction are not predominantly being used to punish the most dangerous behaviors but instead often extend to trivial student behavior.[14] One author has termed this development "net widening," i.e. the sweeping into a punitive system of low-level offenders for which there is no credible deterrence effect from increasingly harsh punishment.[15]

Academic Costs

Given that the amount of instructional time a student receives is an important predictor of achievement outcomes,[16] the loss of instructional time that occurs when millions of students are excluded from school through suspensions hampers the academic development of our youth and diminishes their prospects for becoming productive and successful members of society.[17] In addition, these exclusionary policies inflict psychological costs on students as excessive punishment not only impedes learning and general childhood development but also subverts the relationship of students with, and their trust and confidence in, authority figures, which intensifies conflicts

---

decision—with one third of all disciplinary actions being suspensions—in a study of a large, urban Midwestern public school district).

[10] Johanna Wald & Daniel J. Losen, *Defining and Redirecting a School-to-prison Pipeline*, 99 New Directions for Youth Development 9, 10 (2001).

[11] *Id. See also* Anderson, *supra* note 11, at 1189 (citing OFFICE FOR CIVIL RIGHTS, U.S. DEP'T OF EDUC., FALL 1998 ELEMENTARY AND SECONDARY SCHOOL CIVIL RIGHTS COMPLIANCE REPORT: NATIONAL AND STATE PROJECTIONS (2000)).

[12] PLANTY ET AL., NATIONAL CENTER FOR EDUCATION STATISTICS, THE CONDITION OF EDUCATION 2009 70 (2009).

[13] *Id.*

[14] *See, e.g.*, Donald H. Stone & Linda S. Stone, *Dangerous & Disruptive or Simply Cutting Class; When Should Schools Kick Kids to the Curb?: An Empirical Study of School Suspension and Due Process Rights*, 13 J. L. & FAM. STUD. 1, 12 (2011) (reporting that in Maryland in 2006-07 "only 6.7% of suspensions were issued for dangerous substances , weapons, arsons and sex offenses combined"—with a similarly low 7.1% in 2007-08—while suspensions for disrespect, insubordination, and disruption made up 37.4% of all suspensions); *see also*, ADVANCEMENT PROJECT, TEST, PUNISH, AND PUSH-OUT: HOW "ZERO TOLERANCE" AND HIGH STAKES TESTING FUNNEL YOUTH INTO THE SCHOOL-TO-PRISON PIPELINE 4 (2010), *available at* http://www.advancementproject.org/digital-library/publications/test-punish-and-push-out-how-zero tolerance-and-high-stakes-testing-fu [hereinafter: TEST, PUNISH, AND PUSH-OUT] ("The[] punitive measures [applied as a result of zero tolerance policies] extend far beyond serious infractions; instead, the vast majority of punitive disciplinary consequences tend to result from relatively minor misbehavior or trivial student actions. In fact, the problem in most cases is not the student, but, rather, the adults who react inappropriately to youthful behavior.").

[15] *See* Anderson, *supra* note 11, at 1191-93.

[16] DANIEL J. LOSEN & RUSSELL SKIBA, SUSPENDED EDUCATION: URBAN MIDDLE SCHOOLS IN CRISIS 8 (2010) (citing other studies) [hereinafter: SUSPENDED EDUCATION].

[17] For example, even when suspension numbers were still much lower in 1975, the Children's Defense Fund reported that in the 53% of school districts which it analyzed in a study, over 4 million school days were lost by 1 million suspended students. CHILDREN'S DEFENSE FUND, SCHOOL SUSPENSIONS: ARE THEY HELPING CHILDREN? 56 (1975), *available at http://diglib.lib.utk.edu/cdf/main.php?catid=6.*

3

rather than mediating them.[18] Furthermore, often children either have no access to, or access to gravely deficient alternative education once they are suspended or expelled for a significant amount of time.[19]

Not surprisingly, then, research studies suggest that zero tolerance policies are not effective in either improving school climate by removing disruptive students or in reducing rates of misbehavior through a deterrent function.[20] In fact, these punitive policies might actually be counterproductive. There is evidence to suggest that schools with higher rates of suspensions have "less satisfactory ratings of school climate [and] less satisfactory school governance structures" and that "school suspension … appears to predict higher future rates of misbehavior and suspension among those students who are suspended. In the long term, school suspension and expulsion are moderately associated with a higher likelihood of school dropout and failure to graduate on time."[21] Furthermore, "emerging data indicate that schools with higher rates of school suspension and expulsion have poorer outcomes on standardized achievement tests"[22] independent of socio-demographic factors.[23]

Criminalization of Youth and Push Out

Even worse, research suggests that there is a positive relationship between the use of school suspension and youth incarceration.[24] Schools in a number of states now routinely refer their students, in some instances mandatorily, to juvenile justice, and in some cases adult criminal justice, authorities for school-related misbehavior.[25] Accompanying the shift towards zero tolerance discipline policies has been an increasing presence of law enforcement personnel on public school campuses around the country,[26] in effect "turn[ing] public schools

---

[18]   See Alicia C. Insley, *Suspending and Expelling Children from Educational Opportunity: Time to Reevaluate Zero Tolerance Policies*, 50 AM. U. L. REV. 1039, 1069-70 (2001)

[19]   See, e.g., Augustina Reyes, *The Criminalization of Student Discipline Programs and Adolescent Behavior*, 21 ST. JOHN'S J. LEGAL COMMENT. 73, 80-87 (2006) (explaining also the disproportionately negative effect on minorities imposed by current flaws in alternative education).

[20]   APA Task Force, *supra* note 6, at 854.

[21]   *Id.*

[22]   SUSPENDED EDUCATION, *supra* note 24, at 10.

[23]   See M. Karega Rausch & Russell J. Skiba, *The Academic Cost of Discipline: The Relationship Between Suspension/Expulsion and School Achievement* 19 (2006), *available at* http://www.agi.harvard.edu/Search/SearchAllPapers.php (finding based on two regression models "that school use of out-of-school suspension and expulsion is negatively related to achievement, even when socio-demographic variables are held constant").

[24]   Russell J. Skiba et al., *Consistent removal: Contributions of School Discipline to the School-prison Pipeline*, Paper presented at the Harvard Civil Rights Conference School-to-Prison Pipeline Conference, Cambridge, MA 17 (2003) (finding that across 37 states with available data "there is evidence of a clear relationship between school suspension and juvenile incarceration").

[25]   See, e.g., Thalia Gonzalez - *Keeping Kids In Schools: Restorative Justice, Punitive Discipline, and the School to Prison Pipeline* 7 (forthcoming 2012), *available at* http://works.bepress.com/thalia_gonzalez/3/ (explaining that "[c]urrently, forty-one states require schools to report students to law enforcement for various misbehaviors on campus"); *see also* OPPORTUNITIES SUSPENDED, SUPRA NOTE 12, at Appendix IV (listing in considerable detail reporting requirements of student misconduct to law enforcement for all fifty states).

[26]   See ADVANCEMENT PROJECT, EDUCATION ON LOCKDOWN: THE SCHOOLHOUSE TO JAILHOUSE TRACK 16-18, 23-43 (2005), *available at* http://www.advancementproject.org/digital-library/publications/education-on-lockdown-the-schoolhouse-to-jailhouse-track [hereinafter: EDUCATION ON LOCKDOWN] (explaining how more and more often "Law Enforcement Goes to School" and examining the extent and impact of increased law enforcement presence across four case studies); Aaron Sussman, Comment, *Learning in Lockdown: School Police, Race, and the Limits of Law*, 59 UCLA L. REV. 788 (2012). *See also* U.S. DEP'T OF EDUC., INDICATORS OF SCHOOL CRIME AND

into well-policed fortresses,"[27] and causing increasingly large numbers of school-based arrests.[28] Apart from the harms to all students who have to attempt to study in such difficult and distracting environments, racial minorities suffer disproportionately from increased school criminalization.[29]

Furthermore, zero tolerance policies, interacting with the "high stakes testing" regime that was implemented with the passage of the No Child Left Behind Act of 2001 (NCLB)[30] and its progeny, have been identified as a primary culprit for what has come to be known as the "push out" phenomenon.[31] Because the receipt of federal education funds is contingent on complying with the stringent, and numbers-driven, requirements of NCLB,[32] there is now a perverse incentive to exclude students from the educational environment through punitive school discipline. "[I]f a student acts up in class, it is no longer in educators' self-interest to address it by assessing the student's unmet needs or treating the incident as a teachable moment. It is much easier and more efficient to simply remove the child from class through punitive disciplinary measures . . . ."[33] Because being suspended can

---

SAFETY: 2009, at 69-70 (2009), *available at* http://nces.ed.gov/pubs2010/2010012.pdf [hereinafter: INDICATORS] (describing the increased use of various law enforcement and security tools at schools across the United States).

27   TEST, PUNISH, AND PUSH-OUT, *supra* note 21, at 15.

28   Gonzalez, *supra* note 35, at 8 (reporting very large increases in school-based arrests during the 2000s); TEST, PUNISH, AND PUSH-OUT, *supra* note 21, at 18-19.

29   *See* Sussman, Comment, *supra* note 33, at 793-4 & n.26 (2012) (explaining that "draconian measures used to target minor rule violations . . . are expanding, making the education system and the criminal justice system increasingly difficult to distinguish in low-income, non-white communities").

30   No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002) (codified as amended in scattered sections of 20 U.S.C.). The Act requires states to define and track so-called "Adequate Yearly Progress" requirements that measure and track the academic achievement of all public schools in a particular state. *See id.* at § 1111(b)(2)(B). The "measurable objectives" which each state sets for meeting such progress requirements are then measured by statewide "academic assessments." *See id.* at § 1111(b)(2)(G) (describing "Measurable Objectives" requirement); § 1111(b)(3) (describing "Academic Assessments" requirement). Taken together, these and other sections of the Act implement the high stakes testing regime I reference in this subsection.

31   *See* TEST, PUNISH, AND PUSH-OUT, *supra* note 21, at 3 (explaining that zero tolerance policies have interacted with high stakes testing practices to turn "schools in to hostile and alienating environments for many of our youth, effectively treating them as drop-outs-in-waiting" with the school-to-prison pipeline as the "devastating end result of these intertwined punitive policies")

32   *See* No Child Left Behind Act of 2001, *supra* note 38, at § 1111(a)(1) (conditioning the receipt of federal funds on the submission of a plan describing how a state will meet the academic standards, assessments, and accountability requirements of the Act).

33   TEST, PUNISH, AND PUSH-OUT, *supra* note 21, at 6 (internal quotation marks omitted). Deborah Gordon Klehr pointedly illustrates how students, especially those who are low-achievers academically and are prone to exhibit behavioral problems, are hit hard by this "perfect storm":

> [T]he unintended consequences of combining NCLB with zero tolerance can be a toxic mix for students . . . . Pressure to meet Adequate Yearly Progress combined with pressure to maintain safe schools and enforce zero tolerance policies means Maurice, the seventh grader who acts out in English class as a way to show his frustration with being unable to read, gets expelled for acting out, and his below proficient test scores are not counted against the school. Maybe this seventh grader swore at his teacher. He may be arrested for making a terroristic threat. Thus, in addition to facing an expulsion from school, Maurice may now face criminal charges. He has entered the School-to-Prison Pipeline and may never recover from the consequences. Let's assume Maurice is arrested, adjudicated delinquent, and sent to placement. If and when Maurice returns to his school district, he will likely face barriers to re-enrollment. Even if his expulsion was not permanent, Maurice's credits from juvenile delinquency placement may not transfer to his home district. The home district may attempt to enroll him in an alternative education program for disruptive youth despite the recommendation of his juvenile probation officer and his reformed behavior in

predict eventual school drop-out,[34] many students who get caught in this web of mutually reinforcing pressures do not graduate from high school, have fewer and worse job prospects, and are almost all but set up for future encounters with law enforcement and the criminal justice system.[35]

Economic Costs

Apart from these tragic social costs, there is also an economic cost to such a punitive regime. This cost stems from both the greater cost of running a crowded criminal justice system as well as from the lost productivity of youth that gets caught in the wheels without receiving the necessary skills to become productive members of society. In 2007, for example, "[s]tates spent about $5.7 billion . . . to imprison 64,558 youth committed to residential facilities. . . . [O]n average, it costs states $240.99 per day—around $88,000 a year—for every youth in a juvenile facility."[36] Economists Mark Cohen and Alex Piquero estimate that the monetary present value of saving a "high-risk youth" from a life of crime range from $2.6 to $5.3 million if the youth can be saved by age eighteen, from $3.2 to $5.5 million if the youth can be saved by age fourteen, and from $3.2 and $5.5 million if the youth can be saved by age ten.[37] These costs are constituted by a staggering amount of individual cost items relating to the three categories of costs related to future crime, drug involvement, and high school dropout.[38]

Clearly, we should be extremely critical of school discipline policies that impede our youth's academic potential, criminalize our children, and waste enormous amounts of social and economic resources. As it turns out, however, the general costs of punitive school discipline policies also implicate deep notions of equality and inequality because they do not impose their harsh effects on all societal groups equally.

---

placement. If he is allowed to return to a regular school program, he is likely to be watched closely for any misbehavior while, at the same time, unlikely to receive support services. Finally, the psychological effects of the criminalization of his earlier behavior may lead Maurice to drop-out of school, return to the criminal justice system, or both.

Deborah Gordon Klehr, *Addressing the Unintended Consequences of No Child Left Behind and Zero Tolerance: Better Strategies for Safe Schools and Successful Students*, 16 GEO. J. ON POVERTY L. & POL'Y 585, 602-03 (2009).

[34] *See* Gonzalez, *supra* note 35, at 12 (explaining that "[o]nce removed from schools, students experience decreased academic achievement, further fueling negative attitudes and leading to increased dropout rates"); Russell W. Rumberger, *Why Students Drop Out of School*, in DROPOUTS IN AMERICA: CONFRONTING THE GRADUATION RATE CRISIS 131, 143-44(Gary Orfield ed., 2004).

[35] Gary Orfield, *Losing our Future: Minority Youth Left Out*, in DROPOUTS IN AMERICA, *supra* note 43, at 1, 1 ("Dropping out often leads to economic and social tragedy. High school dropouts are far more likely than graduates to be unemployed, in prison, unmarried or divorced, and living in poverty.").

[36] JUSTICE POLICY INSTITUTE, THE COSTS OF CONFINEMENT: WHY GOOD JUVENILE JUSTICE POLICIES MAKE GOOD FISCAL SENSE 4 (2009).

[37] Mark A. Cohen & Alex R. Piquero, *New Evidence on the Monetary Value of Saving a High Risk Youth*, 25 J. QUANTITATIVE CRIMINOLOGY 25, 46 (2009).

[38] *Id.* at 27-46 (investigating the "Cost of Individual Crimes," "Victim Costs," "Criminal Justice-Related Costs Due to Career Criminals," "Opportunity Cost of Career Criminal's Time While Incarcerated," "Opportunity Cost of Resources Associated with the Manufacture and Sale of Drugs," "Drug Rehabilitation Expenses," "Reduced Productivity Due to Decreased Work Ability," "Medical Costs Associated with Overdose or Other Drug-Related Illness," "Premature Death due to Drug Abuse," "Additional Crime Committed by Drug Users," "Criminal Justice Costs Associated with Drug Use," "Third-Party Costs Associated with Drug Crime," and "Lost Wages And Productivity" associated with high school dropout).

Racial Disproportionality in School Discipline

Equally alarming as the general societal costs associated with punitive school discipline as outline above is the fact that minority youth, and especially African American students, have consistently been found to be disproportionately affected by school disciplinary actions, including out-of-school suspensions, for more than thirty years.[39] And not only has racial disproportionality in school discipline been a consistent finding, it also seems to have increased in severity over time.[40]

According to data from the Office of Civil Rights of the U.S. Department of Education, for example, African American students were suspended twice as "heavily"—as measured by the percentage of black enrollment that had been suspended one day or more—as compared to White students in 1972-73, but were disciplined three times as heavily in 2006-07.[41] Similarly, in an unprecedentedly detailed investigation into the subject using a nationally representative sample of schools, Professor Russell Skiba and his colleagues found African American students to be more than twice as likely to receive a disciplinary office referral as compared to White students at the elementary school level, and more than 3.7 times as likely to receive such referral in middle school.[42] African American students were significantly more likely to receive an office referral *for all offense categories* under investigation.[43] Furthermore, once referred to the office, African American students are significantly more likely to receive out-of-school suspensions as their punishment for a particular disciplinary offense in both elementary and middle school.[44] While this disparity in punishment holds true for all infraction types at the elementary school level, it is reduced to only certain infraction types at the middle school level.[45]

What is significant is that this racial disproportionality in school discipline cannot be credibly explained entirely by differences in socioeconomic status or differential rates of misbehavior.[46] Therefore, while low socioeconomic status is a risk factor for school suspension, race has been found to make a contribution to differences in suspension rates that are independent of socioeconomic status.[47] Similarly, "investigations of student behavior, race, and discipline have yielded no evidence that African American over-representation in school suspension is due to higher rates of misbehavior, regardless of whether the data are self-reported, or based on

---

[39]   *See* RUSSELL J. SKIBA ET AL, INDIANA EDUCATION POLICY CENTER, THE COLOR OF DISCIPLINE: SOURCES OF RACIAL AND GENDER DISPROPORTIONALITY IN SCHOOL PUNISHMENT 3-5 (2000), *available at* http://www.iub.edu/~safeschl/cod.pdf (listing and describing prior studies that found racial disproportionality in school discipline before confirming such findings with own study) [hereinafter: THE COLOR OF DISCIPLINE]. For findings of such racial disproportionality in specific locales, see for example OPPORTUNITIES SUSPENDED, *supra* note 12, and DERAILED, *supra* note 13.

[40]   *See* DANIEL J. LOSEN, NATIONAL EDUCATION POLICY CENTER, DISCIPLINE POLICIES, SUCCESSFUL SCHOOLS, AND RACIAL JUSTICE (2011), *available at* http://nepc.colorado.edu/publication/discipline-policies.

[41]   *Id.* at 4, Figure1.

[42]   *Race is not neutral*, *supra* note 40, at 93-94.

[43]   *Id.*

[44]   *Id.* at 95.

[45]   *Id.* These types, in the nomenclature of the author, were disruption, moderate infractions, and tardy/truancy.

[46]   Russell J. Skiba et al., *African American Disproportionality in School Discipline: The Divide between Best Evidence and Legal Remedy*, 54 N.Y.L. SCH. L. REV. 1071, 1088 (2010)

[47]   *Id.* at 1088 (citing studies).

analysis of disciplinary records. If anything, studies have shown that African American students are punished more severely for less serious or more subjective infractions."[48]

Furthermore, numerous studies have documented that African American students are disciplined more frequently for offenses which are ambiguous and vague while White students are more likely to be disciplined more heavily in clearly delineated offense categories.[49] It seems more likely that racial stereotyping (conscious or unconscious) as well as cultural mismatch between teachers and students are at work and can explain at least some part of existing racial disproportionality in school discipline.[50]

The available data suggest that the foregoing racial disparities in school discipline translate to similar disparities in who is referred to the juvenile justice system[51] as well as the criminal justice system.[52]

This paper is interested in evaluating one particular approach to school discipline which has been offered as an effective means to counter what has been described above: Restorative Justice. In the following, I will shortly review what Restorative Justice is, and what the existing literature has found about its effectiveness in reducing the negative effects of punitive school discipline.

Literature Review

Restorative Justice as a Potential Solution

What is Restorative Justice?

Restorative Justice can be described as an alternative method of dealing with misbehavior distinguished from the two currently dominant models of retribution (justice model) and rehabilitation (welfare model).[53] Its core values focus on "healing rather than hurting, moral learning, community participation and community caring, respectful dialogue, forgiveness, responsibility, apology, and making amends" in an attempt to restore victims, offenders, as well as broader affected communities to a more positive place after something bad has happened.[54] Maybe the most familiar form of disciplinary decisionmaking associated with Restorative Justice is victim-offender mediation (VOM), a practice that has been utilized periodically in the US criminal justice system.[55] "Victim-offender mediation is [a practice available to] victims who want to have a mediation meeting with the offender to

---

[48]   *Id.*

[49]   *See, e.g.*, OPPORTUNITIES SUSPENDED, *supra* note 12; DERAILED, *supra* note 13.

[50]   *See Race Is Not Neutral*, *supra* note 40, at 87.

[51]   *See, e.g.*, Sean Nicholson-Crotty et al., *Exploring the Impact of School Discipline on Racial Disproportion in the Juvenile Justice System*, 90 SOC. SCI. Q. 1003 (2009); *see also* TEST, PUNISH, AND PUSH-OUT, *supra* note 21, at 19; EDUCATION ON LOCKDOWN, *supra* note 37, at 23-43 (describing racial disparities in four case studies).

[52]   *See, e.g.*, Sharon Dolovich, *Creating the Permanent Prisoner*, *in* LIFE WITHOUT PAROLE: AMERICA'S NEW DEATH PENALTY? (Charles J. Ogletree, Jr. & Austin Sarat, eds., forthcoming)

[53]   John Braithwaite, *Restorative Justice: Assessing Optimistic and Pessimistic Accounts*, 25 Crime and Justice 1, 4 (1999).

[54]   *Id.* at 6.

[55]   *See, e.g.*, Mark S. Umbreit, *Restorative Justice Through Victim-Offender Mediation: A Multi-Site Assessment*, 1 WESTERN CRIMINOLOGY REV. 1, 7-9 (1998) http://wcr.sonoma.edu/v1n1/umbreit.html.

discuss how the crime affected them and how the offender can repair the harm. Victim-offender mediations are conducted by trained mediators who are sensitive to the needs of victims and their families."[56]

The goal of VOM, and restorative justice more broadly, is to first identify more specifically what harm has occurred, and to then develop—through dialogue as opposed to top-down punishment—a mutually agreeable solution, of how to repair the harm but also reintegrate the perpetrator into the broader community, for example through a "sentence" of community service.[57] Restorative Justice attempts to counteract problems within the current punitive frame of dealing with misbehavior, which include a lack of deterrence from excessive punishment, excessive imprisonment, and victim alienation, by implementing a collaborative process that focuses on repair of harm through reconciliation and dialogue and greater inclusion of the stakeholders in any particular incident.[58]

Some authors have proposed Restorative Justice models as a potentially effective alternative way of administering school discipline[59]—especially as compared to zero tolerance policies—because they take into account the needs of the multiple actors involved in a disciplinary preceding as well as the multiple levels of harm that a school experiences when it has to deal with violence in any of its many forms.[60] Scholars stress that Restorative Justice's greater focus on accountability, reintegration and inclusion (instead of exclusion and exiling), community building, and the building of problem-solving skills is particularly beneficial for schools because it allows for the development of a safe, collaborative, and positive environment in which students are more likely to strive.[61] Restorative Justice models incorporate different kinds of practices at varying levels of sophistication, ranging from simple circle discussions[62] and affective statements[63] to "continuum models" and whole-school implementation across a variety of school activities beyond disciplinary practices.[64]

Because Restorative Justice focuses on reintegration, dialogue, collaboration and mutual respect, it seems to be a useful approach to countering the problems created by current regimes of punitive school discipline as described in Part I. But are Restorative Justice methods *actually* effective on the ground in doing so?

Restorative Justice – A Potential Solution?

Available studies suggest that Restorative Justice can indeed be beneficial in addressing issues associated with punitive school discipline and in reducing the exorbitant numbers of school suspensions that prevail in U.S. schools today.

For example, a study conducted by the Thelton E. Henderson Center for Social Justice at the University of California Berkeley, School of Law reported very positive results surrounding the implementation of a restorative

---

[56]   *Id.* at 7; *see also id.* at 12-13 (describing VOM process).
[57]   *See id.* at 5-7, 12-13.
[58]   *See id.* at 2-3, 5-7.
[59]   *See, e.g.*, Cara Suvall, Essay, *Restorative Justice in schools: Learning from Jena High School*, 44 HARV. C.R.-C.L. L. REV. 547 (2009); Hanson, *supra* note 3; Anderson, *supra* note 11; Thalia Gonzalez - *Keeping Kids In Schools: Restorative Justice, Punitive Discipline, and the School to Prison Pipeline* (forthcoming 2012), *available at* http://works.bepress.com/thalia_gonzalez/3/.
[60]   *See* Suvall, *supra* note 91, at 559-62.
[61]   Gonzalez, *supra* note 100, at 17.
[62]   I will provide a definition for this for the final version of the paper.
[63]   I will provide a definition for this for the final version of the paper.
[64]   *Id.* at 18-19.

justice program at Cole Middle School in West Oakland, California.[65] Although the school had to be closed due to declining enrollment only two years after the implementation of the program,[66] the report found that the average suspension rate at the school dropped from fifty suspensions per one hundred students to only six suspensions per one hundred students on average for the two years after the program was implemented.[67] Students also reported strong positive feedback on how the program helped to reduce problematic behavior such as fighting and helped build relationships with other students.[68]

Similarly, the International Institute for Restorative Practices (IIRP) found positive results occurring in schools both in the U.S. (specifically six schools in Pennsylvania) and internationally after those schools implemented Restorative Justice programs.[69] With regards to the U.S. schools, the IIRP reported reductions in categories ranging from general suspension rates to the frequency of individual "Behavioral Incidents" categories over different time periods. Across the board, suspension rates were reduced after Restorative Justice programs were instituted in these schools (as well as at the international schools that were examined).[70]

A pilot project Restorative Justice disciplinary practices in four Minnesota school districts also reported encouraging success.[71] Although a number of implementation challenges, including the gathering of useable baseline data, were reported,[72] out-of-school suspensions went down in the aftermath of implementation of a Restorative Justice program at every school with accurate/useable baseline data.[73]

Furthermore, Professor Thalia Gonzalez, in a very comprehensive description of schools that have implemented Restorative Justice programs in the U.S. so far, generally reported decreases in suspension rates after implementation of Restorative Justice programs.[74] At North High School in the Denver Public Schools District— Gonzalez's "case study" school—for example, out-of-school suspensions had declined by 34% four years after implementation of the Restorative Justice program.[75]

Finally, and most pertinent to this project, the Denver Public Schools District (DPS) as a whole has also reported significant reductions in out-of-school suspensions after Restorative Justice practices were implemented in the 2006-07 school year through the Denver Public Schools Restorative Justice Project.[76] Because of early

---

[65] MICHAEL D. SUMNER ET AL., SCHOOL BASED RESTORATIVE JUSTICE AS AN ALTERNATIVE TO ZERO TOLERANCE POLICIES: LESSONS FROM WEST OAKLAND (2010)

[66] *Id.* at 10.

[67] *Id.* at 31.

[68] *Id.* at 20.

[69] INTERNATIONAL INSTITUTE FOR RESTORATIVE PRACTICES, IMPROVING SCHOOL CLIMATE: FINDINGS FROM SCHOOLS IMPLEMENTING RESTORATIVE PRACTICES (2009), *available at* http://www.safersanerschools.org/uploads/news_pdfs/91666_IIRP-Improving-School-Climate.pdf.

[70] *Id.*

[71] Nancy Riestenberg, Presentation at the National conference of the Hamilton Fish Institute on School and Community Violence, *Restorative Measures in Schools: Alternatives to Suspensions – In-School Behavior Intervention Grants* (2004), *available at* gwired.gwu.edu/hamfish/merlin-cgi/p/.../d/.../04Riestenbergpdf/

[72] *Id.* at 42.

[73] *Id.* at 43-49.

[74] Gonzalez, *supra* note 100.

[75] *Id.* at 45. For a description of Restorative Justice use and results throughout the Denver Public School District, see *supra* notes 165-68 and accompanying text.

[76] *See* MYRIAM L. BAKER, DPS RESTORATIVE JUSTICE PROJECT: YEAR THREE, YEAR END REPORT: 2008-2009 3-4 (2009), *available at*

successes with the program, it has expanded continuously and during the 2008-2009 school year served 1235 students.[77] Schools not only experienced increases in timeliness and student social skills after the implementation of Restorative Justice,[78] but were also able to reduce the rate of suspensions per 100 students by almost ten percent over three years of utilizing Restorative Justice.[79]

Limitations of Current Literature, Theory of Current Study

What unites the current literature on the potential of Restorative Justice to reduce the harsh impact of punitive school discipline, then, is that most inquiries into the issue have shown the potential promise of Restorative Justice but have been limited to either data from individual schools or points in time, do not take into account control variables or an appropriate control group, or both. This study addresses both limitations comprehensively by expanding both the number of schools investigated, using a variety of control variables to better isolate the actual effect of Restorative Justice implementation on a reduction in punitive school discipline, and comparing discipline outcomes in Restorative Justice schools to a large control group of schools over time.

In doing so, I focus on one main outcome variable of interest: out-of-school suspension numbers. Because the prior literature summarized above consistently showed a reduction in out-of-school suspensions at the individual schools in which Restorative Justice had been implemented, I assumed that Restorative Justice had an actual independent, and negative, effect on the number of out-of-school suspensions which a particular school will impose on its students. Accordingly, for the first part of my study, I hypothesized that such an effect should also be found if one had the chance to look at a larger group of Restorative Justice schools in the aggregate, if one controlled for other variables that might drive the use of out-of-school suspensions as a response to misbehavior at school, and if one compared the suspension numbers of Restorative Justice schools with those of schools that did not officially change their approach to school discipline to Restorative Justice. This first part of my study is simply an extension of the current literature, imposing greater empirical comprehensiveness on already existing anecdotal research findings. It builds on the same general theory underlying these earlier investigations that Restorative Justice as a more collaborative, dialogue-based approach to student misbehavior that focuses on mutual agreement and integration should lead to lesser use of exclusionary disciplinary sanctions.

In the second part of my study, I attempt to extend the current literature on racial disproportionality in school discipline. As summarized above, such literature consistently shows that racial disproportionality in school discipline exists, and has existed for decades. In this study, I investigate whether the implementation of Restorative Justice actually reduces such disproportionality. Based on my training in the field of Critical Race Theory (CRT), I believe that the ultimate source of racial disproportionality lies in a long historical process of negative racialization of non-white groups in the United States as racial "others" who do not "truly" belong to the U.S. social fabric and

---

www.restorativejusticecolorado.org/_literature_55812/Denver_Public_Schools_Restorative_Justice_Program_Final _Report_2008-2009.
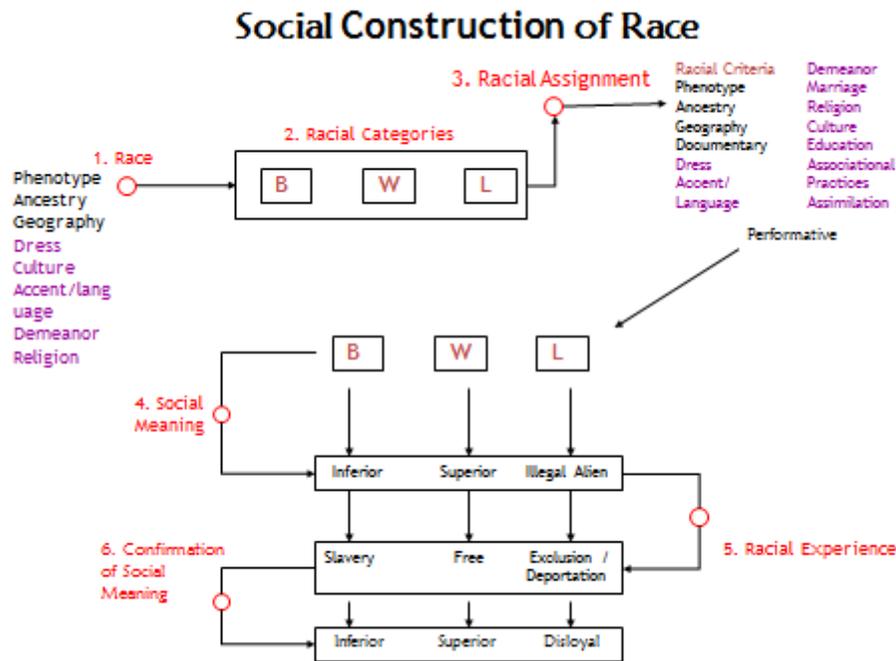
[77]   *Id.*
[78]   *Id.* at 10-14.
[79]   *Id.* at 15-16.

can be treated accordingly.[80] While scholars have developed different specifications of the foundational tenet of CRT that race is a social construction, Devon Carbado has developed the following schematic as a heuristic for describing the phenomenon.

Figure 1: The Social Construction of Race[81]

Social Construction of Race

Step 1 attempts to give content to "this thing we call *race*."[82] In particular, it acknowledges the fact that people in the United States think of "race" along different dimensions which include the phenotype, ancestry, dress, culture, accent, religion and more of a person whose race is being determined.

Steps 2 and 3 describe the process by which individuals then use their impression of another person's race based on such dimensions to sort that person into one of a number of different racial categories through the process of racial assignment. During the process of racial assignment, people engage any number of "racial criteria," some of which overlap with factors used in determining race itself, to determine into which racial category a particular individual should be put. Choices of dress, one's accent or language, general demeanor, religious and cultural practices, one's education and associational practices, and one's perceived level of assimilation into the predominant community culture might all play a role in how people are assigned to particular racial categories in different

---

[80]   David Simson, Exclusion, Punishment, Racism, and Our Schools: A Critical Race Theory Perspective on School Discipline (manuscript on file with author).

[81]   Devon W. Carbado, Discrimination on the Basis of Racial Orientation (manuscript on file with author). I was introduced to both this schematic, and the theory and arguments underlying the following discussion of the schematic, by Professor Carbado in his lectures in my law school course on CRT.

[82]   *Race: The Power of an Illusion*, FACING HISTORY AND OURSELVES, http://www.facinghistory.org/resources/library/race-power-illusion (last visited Feb. 13, 2012) (previewing the content of the 3-part documentary "Race: The Power of an Illusion").

contexts. Significantly, "even when a person does not intend to [actively] manage her identity, . . . the racial meanings others ascribe to her . . . will turn at least in part on her performative identity."[83]

Therefore, the performative aspect of race and racial assignment is intricately tied to step 4, the social meanings associated with particular racial categories. In the broadest sense, the racial hierarchy in the United States has developed a framework which associates whiteness with superiority and dominance,[84] blackness with inferiority and lack of worth,[85] and Latino identity with a presumption, among other things, of illegal immigration status.[86]

As society develops and acts upon such social meanings, members of specific racial categories are bound to have particular racial experiences that reflect the existing racial hierarchy: step 5 in the schematic. Once a society has deeply invested itself in ascribing the meaning of inferiority to blackness, for example, slavery becomes not only a possibility but indeed seen as "desirable," even for its victims.[87] On the other hand, whites, marked and perceived as the superior race, are then seen as entitled to all the benefits of freedom and liberty that were thought to form the foundation of American society.[88] Finally, the racial meaning of presumptive illegality leads to efforts to push

---

[83]   Carbado, *supra* note 52, at 1609.

[84]   *See, e.g.*, Plessy v. Ferguson, 163 U.S. 537, 559 (1896) (Harlan, J., dissenting) ("The white race deems itself to be the dominant race in this country. An so it is, in prestige, in achievements, in education, in wealth and in power. So, I doubt not, it will continue to be for all time, if it remains true to its great heritage and holds fast to the principles of constitutional liberty.").

[85]   One particularly infamous rendition of this idea was delivered by then United States Supreme Court Chief Justice Roger Taney in Dred Scott v. Sandford:

> It is difficult at this day to realize the state of public opinion in relation to that unfortunate race [blacks], which prevailed in the civilized and enlightened portions of the world at the time of the Declaration of Independence, and when the Constitution of the United States was framed an adopted. But the public history of every European nation displays it in a manner too plain to be mistaken. They had for more than a century before been regarded as beings of an inferior order, and altogether unfit to associate with the white race, either in social or political relations; and so far inferior, that they had no rights which the white man was bound to respect; . . . This opinion was at that time fixed and universal in the civilized portion of the white race. It was regarded as an axiom in morals as well as in politics, which no one thought of disputing, or supposed to be open to dispute. . . .

Dred Scott v. Sandford, 60 U.S. (19 How.) 393, 407-08 (1856). *Cf.* Randall L. Kennedy, *Racial Critiques of Legal Academia*, 102 HARV. L. REV. 1745, 1751 (1989) ("[O]f all the many racially derogatory comments about people of color, particularly Negroes, none has been more hurtful, corrosive, and influential than the charge that they are intellectually inferior to whites.")

[86]   The most prominent example of such attitudes towards Latinos/as as presumptively illegal, especially where they comprise a large portion of the local population, can be found in recent anti-immigration bills such as Arizona's S.B. 1070 or Alabama's H.B. 56, which empower law enforcement and school officials to inquire into the immigration status at a routine traffic stop or at school. There has been considerable controversy over such measures because of their potential to encourage racial profiling. Such profiling, especially in Southern states like Arizona, is widespread. *See, e.g.*, U. S. DEPARTMENT OF JUSTICE, CIVIL RIGHTS DIVISION, LETTER TO BILL MONTGOMERY, COUNTY ATTORNEY OF MARICOPA COUNTY (Dec. 15, 2011) (notifying the county attorney of Maricopa County of the rampant racial discrimination practiced by the Maricopa County police department, especially as it relates to racial profiling and mistreatment of Latinos/as).

[87]   *See* Dredd Scott, *supra* note 88, at 407 (arguing that the history of Western societies clearly shows a perception of blacks as being "so far inferior, that they had no rights which the white man was bound to respect; and that the Negro might justly and lawfully be reduced to slavery *for his benefit*") (emphasis added).

[88]   Before the Civil War, for example, citizenship via naturalization was only available to "free white person[s]." *See* Ozawa v. United States, 260 U.S. 178, 194-95 (1922). Only, after the Civil War in 1870 was this citizenship statute expanded to "aliens of African nativity and to persons of African descent." *Id.* at 195.

Latinos/as out of the United States via all means that are legally available or make their lives in the United States miserable.[89]

Completing the vicious circle of the social construction of racial hierarchy is step 6, where the racial experiences to which members of particular racial categories are being exposed serve to confirm existing racial meanings. That is to say, society comes to think that if you are categorized as black, and you have been legally forced into slavery, it must be that this is because you are inferior to those who have always been free and not so subjugated. If you are white, and enjoying liberty and freedom, this must be because you are superior. If you are categorized as Latino/a and have been the subject of legislation encouraging your deportation from the United States, this must be because your presence in the country was illegal and inappropriate in the first place.

Applying this general framework to this study, we can hypothesize a similar dynamic which looks something like Figure 2.

Figure 2. The Social Construction of Race in the Context of School Discipline
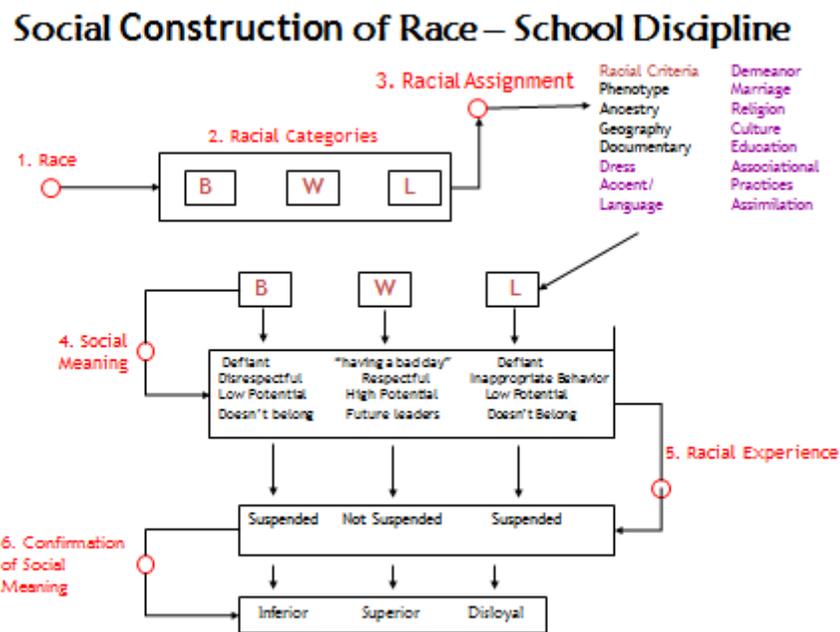
Taking step 1 for granted for the purposes of this discussion (i.e. recognizing that race is a fluid category whose definition depends on different dimensions), students are assigned by their teachers and school administrators—those who are involved in making a disciplinary decision—into the racial categories "white," "black," and "Latino." Different racial criteria might be salient depending on the particular context of the interaction, but given the face-to-

---

[89]   Arizona's and Alabama's anti-immigration laws, for example, are based explicitly on the policy of "self-deportation" by attrition and attempt to make life so difficult for immigrants without proper immigration status that they will "voluntarily" exit the country. Similarly, some proponents of Arizona's recent law against ethnic studies courses in public high schools have not hidden the fact that this law was intended only to apply to courses in Chicano studies.

face nature of many incidents leading to a disciplinary decision,[90] it is particularly likely that phenotype and certain performative racial assignment criteria, in particular dress, accent, general demeanor, and associational practices of the student will be highly influential in the racial assignment process of the decision-maker.

Through a complex and interlocking process influenced by longstanding notions of racial stigma, societal stereotypes and implicit bias which derive in part from such stigma, differential perception and evaluation of the same event by members of the racial majority and minority, and normative baselines as to what constitutes "appropriate" behavior, the disciplinary decision-maker evaluates the behavior of the student within an existing framework of social meanings associated with the student's racial category. In situations in which there is at least some ambiguity as to whether a disciplinary violation has occurred,[91] these meanings might be the decisive factor in evaluating whether a student was "defiant" or "having a bad day," respectful or disrespectful/behaving "inappropriately," dangerous and threatening or harmless. I argue that this evaluation of student behavior will generally and predictably be one which affects minority youth negatively.

This evaluation, then, determines which disciplinary action will be taken, and in turn which racial experience the student is subjected to. I hypothesize that, at a minimum, the portion of racial disproportionality in school discipline which cannot be explained by socioeconomic factors and rates of actual misbehavior can be attributed to this process.[92] Completing the vicious cycle, the experiences of this nation's youth serve to confirm and rigidify broader social meanings that associate inferiority and a lack of true societal belonging with blackness, presumptive illegality to Latino/a identity, and superiority and societal leadership with whiteness.

If one accepts the argument that racial disproportionality in school discipline which is based on a long history of both explicit and subtle racial discrimination against non-white groups is inappropriate, the need to find alternative approaches that reduce such disproportionality becomes clear. I hypothesized that Restorative Justice can provide such an alternative, and that it would be effective in reducing racial disproportionality in school discipline for the following reasons.

---

[90]   At the stage of the initial office-referral, for example, a teacher or school safety officer will be present at the particular incident and initiate the disciplinary process. Similarly, because students who are suspended, even for short amounts of time, have a due process right to be granted at least an informal hearing at which they are presented the evidence against them and at which they can provide their side of the story, the principal or other administrator making the disciplinary decision will likely have at least cursory face-to-face interaction with the student. *See* Goss v. Lopez, 419 U.S. 565, 581 (1975) ("Students facing temporary suspension have interests qualifying for protection of the Due Process Clause, and due process requires, in connection with a suspension of 10 days or less, that the student be given oral or written notice of the charges against him and, if he denies them, an explanation of the evidence the authorities have and an opportunity to present his side of the story.").

[91]   This ambiguity can occur at two different points: offense definition; and behavior evaluation. Particular offense categories might be defined so vaguely that there will always be ambiguity as to whether a violation has occurred or not. Candidates falling into this category are offenses such as "defiance of authority." *See* Opportunities Suspended, *supra* note , at 4-5 (describing the many different behaviors for which students could be suspended for "defiance of authority" using the experiences of a ten-year old African American girl). Similarly, ambiguities could arise as to whether an offense category has been met, even if it is fairly clearly delineated. For example, it might not always be clear whether there was an assault on a teacher if a teacher is attempting to break up a fight between students and gets hurt in the scuffle. Whether a student will be charged with, and suspended or even expelled for, assault on a teacher or not will depend on the teacher's subjective evaluation of the behavior of the student in the particular situation.

[92]   *See supra* notes 59-61 and accompanying text

For one thing, Restorative Justice might be an attractive alternative to zero tolerance policies because it is a practice that can provide a forum for the voice of minorities to show that systematic inequalities are still a reality in society. CRT scholars, for example Professor Russell Robinson, have pointed out that that there is a wide gulf between perceptions of whites and African Americans on the issue of racial discrimination in the United States.[93] This phenomenon is called perceptual segregation and posits that "[b]lack and white people tend to perceive allegations of racial discrimination through substantially different perceptual frameworks."[94] In particular, whites generally subscribe to a "colorblindness perspective" which "views discrimination as an aberration from a colorblind norm, and . . . regards most forms of race-consciousness as socially disruptive" while blacks more often subscribe to a "pervasive prejudice perspective" which "views discrimination as a commonplace event, rooted in daily social dynamics."[95] The existence of such perceptual segregation is thus highly significant in the context of school discipline because it affects a person's "very definition of racial discrimination."[96] A similar process is likely at play in the context of disparate treatment of Latinos/as.

Disciplinary decisionmakers who subscribe to the colorblindness perspective will be more likely to suspend African American and Latino/a students for two interacting reasons. First, they will be less likely to be aware of the fact that their decisionmaking in a particular case is influenced by notions of racial stigma and negative racial stereotypes towards African Americans and Latinos/as. After all, if you are convinced that colorblindness is society's norm and that thus racial discrimination is no longer a significant problem, you are more likely to believe that your decision to suspend a student for "dangerous," "threatening," or "defiant" behavior in an ambiguous situation was "on the merits" when in fact it might not have been. Second, they will be less likely to see the consistent findings of pervasive racial disproportionality in school discipline in the United States as the product of biased decisionmaking and thus as a cause for concern or a change in their own decisionmaking.

Restorative Justice has the potential to bridge the gap of understanding created in part by perceptual segregation. In particular, its focus on dialogue and on giving all parties to a situation that resulted in harm an opportunity to interact and share their perspectives might lead to the discovery of misunderstandings and different perceptions of the interactions between members of a school community in the first place. After all, perceptual differences are maintained by the "failure . . . to interact meaningfully with [others] and to engage those of other races on racial issues."[97] Restorative Justice, especially when it is implemented in schools on a broad and daily basis, encourages not only interaction per se, but *meaningful* interaction and gives a voice to those who otherwise might not have a safe place to express their thoughts and emotions.

---

[93]   Russell Robinson, *Perceptual Segregation*, 108 Colum. L. Rev. 1093, 1106-13 (marshaling empirical evidence showing that African Americans more often perceive that African Americans as a group are subject to more frequent and blatant discrimination, and that measures taken to remedy such discrimination are often inadequate, as compared to whites).

[94]   *Id.* at 1106.

[95]   *Id.* at 1117.

[96]   *Id.* It should be noted that the theory of perceptual segregation deals with differing perceptions between racial groups on average. African Americans are not immune from subscribing to the colorblindness perspective, with Supreme Court Justice Clarence Thomas probably being the most prominent example. [find good source for this point].

[97]   *Id.* at 1133.

Restorative Justice circles, affective statements, conferences, and the like can bring to the surface the causes of instances of "behavioral leakage"[98] by encouraging both students and disciplinary decisionmakers to be honest about how they perceived the other person's behavior in the interaction that might lead to a suspension for, say, "defiance." A teacher may find out, for example, that the student felt neglected by the teacher, not sufficiently respected, or that issues outside of the classroom weighed on their behavior in class. The teacher might then understand both how his or her own behavior might have caused the situation to escalate into a suspension-worthy event, or that what the student might need is not being kicked from school for a number of days but a showing of respect and care. Even more importantly, students have a greater ability to share their belief that a teacher's behavior towards them was inappropriate and that they believe that their race played a role in this process. In the best case scenario, teachers will reconsider their own potential biases and adjust their behavior appropriately. In any case, they will be reminded that a different interpretation of their behavior exists, and that it might impose racial harm on the students they are committed to serve. In this way, Restorative Justice programs represent a real-life implementation of a very important maxim in Critical Race Theory: the need to "look to the bottom."[99]

In this study, I attempt to measure empirically whether Restorative Justice indeed has a positive effect on reducing racial disproportionality in school discipline and thus comports with the theory laid out above. I do so utilizing information from the following dataset.

Data Collection

Data Sources

The dataset for the current project stems from two different surveys conducted by different sub-departments of the Department of Education (Dept. of Ed) of the federal government. Suspension data was collected from the Civil Rights Data Collection Survey (CRDC), which is administered in two to four year intervals by the Office of Civil Rights (OCR) within the Dept. of Ed. It collects information about a variety of civil rights related programs and issues in a broad sample of U.S. public primary and secondary schools, including disciplinary outcomes disaggregated by race. The CRDC is administered in two parts: The first collects suspension data at the end of a particular school year. In a second-wave part of the survey, the CRDC collects enrollment numbers from the following fall semester. Data from the CRDC is freely available on the internet at the website ocrdata.ed.gov. The website allows visitors to download suspension data in comma separated value (csv) files.

The suspension data made available to the public via the CRDC suffers from two complications, however. First, in an effort to anonymize the data and avoid identification of individual students, OCR staff rounds reported

---

[98]   By behavioral leakage I mean the process through which implicit biases towards racial minorities have been found to correlate with more negative interactions with, for example, blacks. Negative non-verbal behavior, in turn, is very likely to create positive feedback loops of retaliatory responses in social interaction, escalating conflict in the process. *See* Jerry Kang, *Trojan Horses of Race*, 118 HARV. L. REV. 1489, 1523-25 (2005).

[99]   *See* Mari Matsuda, *Looking to the Bottom: Critical Legal Studies and Reparations*, 22 HARV. C.R.-C.L. L. REV. 323 (1987) Also think about using Kimberlé Williams Crenshaw, *Foreword: Toward a Race-Conscious Pedagogy in Legal Education*, 11 Nat. Black L.J. 1, 8 (1989) (proposing a shift in the "frame" of discussion of racism in society including a "depriveleging [of] dominant perspectives" and providing proper space for minority perspectives in legal pedagogy).

suspension numbers to counts of 5. In an effort to attempt imputation of values, I inquired into the way in which OCR rounds such numbers, but the office refused to answer my question. Therefore, suspension counts reported by the CRDC might over-, or underreport the actual number of suspensions handed out by a particular school. This is a problem which is particularly likely to influence summary statistics such as suspension percentages when suspension counts are very low, and thus introduces additional error into any analysis of CRDC suspension data.

Second, data from the CRDC is not sufficient to answer empirical questions going beyond counting the total number of suspensions handed out by a particular school. Because of the temporal disconnect between the point in time at which suspension numbers and enrollment numbers are being solicited from school respondents, as mentioned above, any calculation of statistics that rely on enrollment numbers, such as suspension percentages and the like, are bound to be inaccurate. To alleviate this problem, I collected data from a second source: the Common Core of Data (CCD) survey administered and collected by the National Center for Education Statistics. The CCD is a survey which is administered every year and solicits basic information such as enrollment, school level, pupil teacher ratio, free lunch rates, and the like from every public school in the United States. Data from the CCD are also publicly available from the CCD website http://nces.ed.gov/ccd/pubschuniv.asp.

Because the CCD is administered every year, I was able to match its information with suspension data from the CRDC to produce a dataset that had both suspension and enrollment numbers from the same school year. CRDC and datasets were merged based on a unique school code with which schools are identified in both the CRDC and the CCD.

Study Sample

The sample of schools under investigation in this paper contains suspension numbers from the CRDC and general school information from the CCD on all schools in two school districts: the Denver Public Schools district (DPS), and the Santa Fe Public Schools district (SFPS). These two districts were chosen non-randomly. Instead, they were chosen because based on my research they contain the two largest concentrations of schools utilizing Restorative Justice methods in their school discipline procedures in the United States. I collected data from these two districts for two school years: 2005-06 and 2009-10. I chose these dates because they are the years in which the latest two iterations of the CRDC were collected, and because Restorative Justice was implemented in a large number of schools in those districts between these two CRDC iterations. Thus, suspension numbers from these two schools years should show the impact of the implementation of this alternative approach to school discipline. The two districts contain a combined number of 34 schools that began using Restorative Justice in their disciplinary practices after 2006 but before 2009.

These two districts are also attractive to study because they are large enough to contain significant control populations. In particular, there were 292 non-Restorative Justice entries in the dataset containing data from individual schools for either 2005-06 or 2009-10. Because not all of the non-Restorative Justice schools existed in both 2006 and 2009, and because certain schools had to be dropped from the analysis either because they contained information which suggested erroneous data entry, or because they were unusual types of schools—such as vocational schools or special education schools—the final dataset on which I performed my analyses contained a

grand total of 143 schools across the two districts. 113 of those schools were "non-Restorative Justice" schools, and 30 of them had implemented Restorative Justice into their disciplinary practices between the 2005-06 and 2009-10 school years.

Variables and Variable Coding

Dependent Variables – Overall Suspension Numbers

My outcome variables revolve around a single metric: out-of –school suspensions. I chose suspensions as the basis for all of my outcome variable for two reasons: First, because out-of-school suspensions (hereinafter: suspensions), in many places, are the most common method with which administrators respond punitively to misbehavior at school, suspension numbers are likely to be large enough to make meaningful statistical analysis possible. Furthermore, because suspensions are so commonly used, they are also more likely to be recorded, and thus reported, accurately. Second, as the most frequently used tool of exclusionary discipline, reducing the number of suspensions doled out by school administrators should be the primary goal of those who are interested in alleviating the negative effects of punitive school discipline described above.

Suspension Percentage – Measurement Error

The main input into all of my dependent variables in this study was the suspension percentage of individual schools. Suspension percentage simply refers to the total number of suspensions recorded by an individual school for a given school year divided by the total enrollment of that school for that school year. The total number of suspensions was arrived at by recoding some of the data downloaded from the CRDC internet database. Before the 2009-10 CRDC, schools were instructed to report suspensions as the number of unique individual students having received one (or more) suspensions during a particular school year. This measure represented the total number of suspensions reported for all schools in the 2005-06 CRDC iteration which, when divided by total enrollment, gave me suspension percentages at the school level. Total suspension numbers in 2009-10, however, had to be calculated differently by recoding CRDC data. This was so because the 2009-10 CRDC survey began to collect information on suspensions in a different format. Starting with the 2009-10 CRDC, suspension numbers are now collected for the total number of students who have been suspended once only, as well as for the total number of students who have been suspended more than once. Thus, to keep the data consistent, I added the totals of students suspended once and students suspended more than once to arrive at the same metric for total suspensions that the CRDC used in 2005-06.

As mentioned above, it was necessary to merge the CRDC and CCD datasets to be able to calculate suspension percentage because enrollment numbers in the CRDC are measured at a different point in time than CRDC suspension numbers. Accordingly, the enrollment figures that represent the denominator of my suspension percentage metric come from the CCD dataset for the school year for which the CRDC collected suspension numbers.

There are three potential sources of measurement error affecting my results for the suspension percentage metric. The first stems from the fact mentioned above that the OCR rounds suspension numbers to counts of five in

the CRDC. The second stems from the different way in which total suspensions were being recorded in the CRDC until the 2009-10 iteration of the survey (i.e. collecting suspensions as the number of unique individuals receiving one or more suspensions). It is possible that school administrators misunderstood or disregarded this instruction and simply recorded the total number of instances leading to a suspension, which may include the same individual student multiple times and skew suspension numbers upward. In fact, one school had to be dropped from my dataset because it recorded a suspension percentage of over 100 percent. Third, combining the data from two different datasets collected by two different government agencies has the potential to introduce measurement error. It might be, for example, that two different school administrators recorded the responses to the CRDC and CCD and made their decision on how to count fringe cases based on different criteria. The same process might occur when the OCR and the NCES clean the data for public use.

Suspension Percentage Difference

The main dependent variable in my analysis of whether the implementation of Restorative Justice programs in Denver and Santa Fe was successful in causing a decrease in suspensions overall was the difference in the suspension percentage of an individual school between 2005-06 and 2009-10. I used this metric as my main dependent variable because I wanted to investigate not only whether the implementation of Restorative Justice led to a reduction in suspensions overall (which the prior literature had done), but also whether a potential reduction was significantly different from a similar reduction other schools that did not implement Restorative Justice might also have experienced over the same time frame. Because the variable suspension percentage difference was normally distributed (see Appendix B), I did not perform any variable transformations on it.

Suspension Percentage Ratio – "No Zero" (Natural Log)

Because a measure of suspension percentage difference privileges schools with high suspension percentages to begin with and schools that implemented Restorative Justice between 2005-06 and 2009-10 had higher suspension percentages in 2005-06 (see Appendix A and Figure 1, *infra*), I created a second dependent variable in my attempt to measure differences in overall suspension percentages between schools with and without Restorative Justice programs: the ratio of a school's suspension percentage in 2009-10 divided by 2005-06. I created this particular ratio because I am interested in whether suspensions *declined* over time. Therefore, I calculated my ratio variable using suspension percentages of individual schools from the 2009-10 school year in the numerator and suspension percentages from 2005-06 as the comparative baseline in the denominator. Because this variable looks at a ratio of two suspension percentage numbers, it reduces the privilege received by schools with high initial suspension percentages.

This variable, however, suffers from the weaknesses that it does not include data from any schools that had an initial suspension percentage of zero in 2005-06 because division by zero is mathematically not possible. Thus, those cases fall out of the analysis. Furthermore, the variable in its initial form was highly skewed to the right. Therefore, I transformed the variable by taking the natural logarithm of the untransformed suspension percentage ratio described above. As with the raw suspension percentage ratio, all cases that had a value of zero in the

denominator of the ratio—i.e. schools that did not have any suspensions in 2005-06—fall out of the analysis. However, with this particular metric, all schools that do not have any suspensions in 2009-10 also fall out of the analysis because taking the natural logarithm of zero is impossible.

Still, I wanted to use this metric because I thought that it would be interesting to look only at schools that had some minimal level of punitive school discipline both in 2005-06 and in 2009-10 and to compare the influence of the implementation of Restorative Justice among those schools. This particular variable is very normally distributed and thus highly amenable to traditional statistical analysis (see Appendix B).

Dependent Variables – Racial Disproportionality Analysis

For my analysis of racial disproportionality in school discipline, I use the following dependent variable:[100]

Black Suspension Percentage Disparity

Black Suspension Percentage Difference was calculated in the exact same way as Hispanic Suspension Percentage Difference. However, because black enrollment in the two school districts under study was much smaller than Hispanic enrollment, only 58 schools had large enough ethnic diversity to fulfill my requirements of enrollment of more than 20 students in both ethnic categories. This variable was also approximately normally distributed (see Appendix C).

Treatment Variable – Restorative Justice

My treatment variable was a dummy variable that indicated whether or not a school had implemented Restorative Justice principles into their school disciplinary practices between 2005-06 and 2009-10. I collected such information with the help of Professor Thalia Gonzalez at Occidental College, who is currently engaged in in-depth research on the effectiveness of Restorative Justice programs across the country, Tim Turley, who is the Restorative Justice coordinator of DPS, as well as a Restorative Justice Initiative End of Year Report from 2008-2009 produced by SFPS.

It should be noted that the implementation of Restorative Justice was not identical across all Restorative Justice schools in my sample. Within DPS, some schools implemented Restorative Justice more in-depth than others, and within SFPS, some schools used Restorative Justice only for some disciplinary issues but not others. However, I made the choice to include all schools that had implemented Restorative Justice in an organized manner (for ex. by making a Restorative Justice coordinator available, having a formal peer court or peer panel program

---

[100]   In my initial conceptualization of this paper, I also wanted to investigate Hispanic racial disparity in school disciplinary outcomes in the two school districts under study. Therefore, my theory section includes much discussion of the racialization of Latinas/-os in the United States. However, initial analysis showed that at least for the time period under investigation in this study, there was no initial racial disparity in suspension numbers for Latina/-o and white students. Therefore, the implementation of Restorative Justice programs could not have alleviated this problem. As I increase the size of my dataset, I hope to be able to pick up on this point of inquiry again.

using Restorative Justice principles, or consistently training teachers in the use of Restorative Justice methods in their class room) deliberately and for two reasons. First, doing so allowed me to increase the sample of schools which could fairly be characterized as Restorative Justice schools, thus increasing the statistical power of my analysis. Second, because Restorative Justice is so different from punitive school discipline in its approach to handling misbehavior at school, I was willing to make the assumption that coordinated implementation in any organized manner—regardless of whether it was 100 percent complete or not—would exert a measurable effect on the disciplinary practices of those schools. Furthermore, because Restorative Justice is a complex approach to resolving disputes which features many different practices, it is almost impossible to determine whether any school has "completely" implemented Restorative Justice, and whether the form of implementation at one school would also be appropriate at another school.

That being said, my estimates are thus conservative and might underestimate the potential influence of implementing Restorative Justice fully at a particular school. The estimates are made even more conservative by the fact that some schools implemented Restorative Justice at later points within the time period between 2005-06 and 2009-10, and thus might not have had as much/enough time to implement the practices to a point where they would exert a potential influence on suspension numbers. Further, at least at DPS, language encouraging the use of Restorative Justice in resolving disciplinary disputes had found its way into the district's student discipline code.[101] Thus, some of the schools not designated as Restorative Justice schools in my dataset might have actually changed their approach to move away from punitive school discipline (Tim Turley indicated as much in an email to me) and might have depressed the contrast to the formally designated Restorative Justice schools in my dataset. Lastly, I kept one school in my dataset as a Restorative Justice school even though the school formally opted out of the program during the 2008-09 year. I did so because I assumed that this school, having used Restorative Justice for a couple of years, was more appropriately considered within the treatment group, rather than the control group.

As mentioned above, the total count of schools in my dataset was 113 non-Restorative Justice schools and 30 Restorative Justice schools.

Control Variables

In an effort to isolate the effect of Restorative Justice on suspension outcomes as much as possible, I also utilized a number of control variables collected from the CCD. I surveyed all variables collected by the CCD and determined that the following variables were those that were potentially related to suspension outcomes in my

---

[101] *See* Baker, *supra* note 84 (citing Policies and Procedures: Policy JK – Student Discipline, Denver Public Schools, http://www.dpsk12.org/policies/Policy.aspx?-db=policy.fp3&-format=detail.html&lay=policyview&-sortfield=File&-op=eq&Section=J&-recid=32883&-find= (last visited February 20, 2012) (explaining that "[t]here are three types of intervention strategies that are available: Administrative, Restorative, and Skill-based/Therapeutic" and noting the expectation that "[s]chools should minimize the use of out-of-school suspensions, recommendations for expulsion, and referrals to law enforcement, to the extent practicable" since "[i]t is a goal of the Denver Public Schools and the Board of Education that the juvenile and criminal justice systems be utilized less frequently to address school-based misconduct.").

dataset:[102] school enrollment (total); enrollment by the five major ethnic categories Hispanic, black, white, Asian American, and American Indian; pupil teacher ratio; level of school; Title 1 eligibility; and the number of students eligible for free school lunch. Furthermore, because Restorative Justice schools had such a high level of suspension percentage in 05-06, I also controlled for initial suspension percentage.

Suspension Percentage

Suspension percentage as a control variable simply represented the suspension percentage of a given school for the 2005-06 school year. However, because suspension percentage proved to be highly skewed to the right, I transformed this variable by taking the natural logarithm of the 2005-06 suspension percentage. I added 1 to the suspension percentage number of which I took the natural logarithm in order to preserve the schools that had a suspension percentage of zero in 2005-06.[103] For the distribution of this control variable, see Appendix C.

Total Enrollment

The total enrollment of each school was reported as such in the CCD survey. However, the distribution of the data was skewed to the right. Therefore, I conducted a log transformation to make the enrollment data more normally distributed and useable in my statistical analysis (see Appendix C).

Ethnic Enrollment

I measured ethnic enrollment numbers in percentages of total enrollment rather than actual counts to avoid high correlation with total enrollment. Like total enrollment, and even more pronouncedly so, ethnic enrollment percentage numbers, with the exception of Hispanic enrollment, were highly skewed to the right. I conducted natural log transformations to make the data more normally distributed. Hispanic suspension percentage turned out to be so high across both districts that the distribution of Hispanic students could not be improved by transforming the data. Thus, Hispanic percentage is the only non-transformed ethnic enrollment category in the dataset (see Appendix C).

Pupil Teacher Ratio

Pupil Teacher Ratio as collected by the CCD measures the number of students per full-time equivalent teacher at a particular school. While there was one outlier in the dataset, and transforming the data by taking the square root of the ratio slightly improved the visible shape of the data (see Appendix C), the untransformed ratio proved to be the most useful metric in my analysis and thus was used instead of the transformed data.

Level of School

Level of School is a categorical variable that indicates whether the school is a primary school, middle school, or high school. The CCD includes the following four school levels:

---

[102] There were some other variables in the CCD survey, such as geographic location of the school in relation to an urban center or whether or not the school was a charter school. Because they did not vary in my dataset, however, I did not use them in my analysis.

[103] I used the same log transformation method for the rest of my controls as well. I only refrained from adding 1 to the natural logarithm of my suspension percentage ratio variable because I thought that doing so might skew the results inappropriately.

1 = Primary (low grade = PK through 03; high grade = PK through 08)
2 = Middle (low grade = 04 through 07; high grade = 04 through 09)
3 = High (low grade = 07 through 12; high grade = 12 only)
4 = Other (any other configuration not falling within the above three
categories, including ungraded)

However, in order to make the dataset consistent, I recoded al schools that had level 4 assigned to them into the categories into which they would most naturally fall. I recoded ten schools from level 4 into level 3 based on the grades they offered. Four of those schools were not in level 3 initially because they also offered 6th grade education, but because they offered all grades up to 12th grade, I coded them as high schools. All other schools with a level 4 designation in my dataset should have already been classified as high schools, but for some reason the CCD seemed to have a faulty designation. Ultimately, I had 105 level 1 schools, 20 level 2 schools, and 18 level 3 schools in my final dataset.

Title 1 Eligibility of the School

Title 1 Eligibility of a school is a binary variable that indicates whether a school is eligible to receive grants under Title I, Part A (Title I) of the Elementary and Secondary Education Act providing financial assistance to schools with high percentages of low income students to support them in their academic progress. In my dataset, 103 schools were eligible for Title 1 funding while 40 were not.

Free Lunch Students

The CCD free lunch variable indicates how many students are eligible to participate in the Free Lunch Program under the National School Lunch Act. I recoded this variable as a percentage by dividing the number of free lunch eligible students by total enrollment. Free lunch rate, like Hispanic enrollment percentage, proved to be most useful in an untransformed state. For its distribution, see Appendix C.

Methods and Results

General Method of Analysis

I began my statistical analysis of the suspension data of the 143 schools in my dataset by identifying the most appropriate statistical model that would predict my dependent variables of interest using my control variables but not my treatment variable. In other words, I attempted to identify which set of my control variables would predict my outcome variable of interest independent of my treatment variable. I did so by first identifying the pairwise spearman correlations between each control variable and my particular dependent variable, testing the significance of such correlation, and confirming the correlation by visual inspection of pairwise scatterplots. I then ran OLS regressions (since all of my dependent variables are continuous) to see whether highly correlated variables were actually useful in predicting my dependent variable of interest. I then specified the model that used the least amount of control variables to achieve the highest predictive power (as measured by the R^2 value of individual regressions). Finally, I then matched the schools in my control group with schools in my treatment group along the

control variables so identified, using the coarsened exact matching (CEM) method, and then compared dependent variable means between matched treatment and control groups using a simple t-test.

Results

Overall Suspension Percentage Difference

In terms of baseline data for suspension percentages in my dataset, schools that would implement Restorative Justice programs between 2005-06 and 2009-10 interestingly had much higher initial suspension percentages than schools that would not implement Restorative Justice (see Figure 1).

Figure 1. Initial Suspension Percentage Rates of Schools That Would and Would Not Implement Restorative Justice
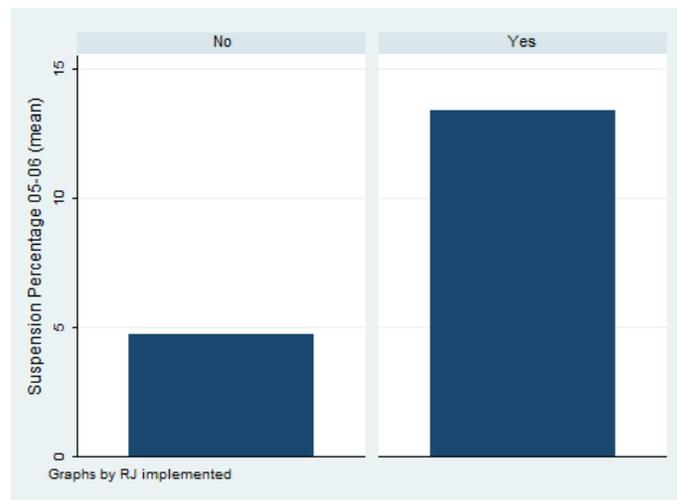
Pairwise correlation analysis identified only (logged) suspension percentage and the level of school as significantly correlated with overall Suspension Percentage Difference. Visual inspection of pairwise scatterplots and regression analysis, however, identified enrollment, free lunch rate, and black enrollment percentage as also useful predictors of Suspension Percentage Difference. A very strong relationship between suspension percentage and suspension difference was particularly noteworthy. This is not surprising and confirmed my initial belief that using raw suspension percentage difference would privilege schools with high initial suspension percentages, and thus implicitly RJ schools as well. A second model that does not use enrollment as a control and achieved similar predictive power was also specified (see Appendix D for regression output of the best predictive models).

As the first step in my analysis of whether suspension difference was significantly different between Restorative Justice and non-Restorative Justice schools, I ran a t-test comparing the means of suspension percentage difference between the two groups without controlling for any other variables. For purposes of this study, I was interested in one-tail differences (left-tail) in means rather than in the more commonly tested two-tail difference. This is so because prior literature had identified reductions in suspension numbers at schools that had implemented Restorative Justice in their disciplinary practices. No research that I am aware of had found any increases. Thus, I hypothesized that the average suspension percentage difference should be higher in Restorative Justice schools as

25

compared to non-Restorative Justice schools. Because the Restorative Justice variable was coded as 0 for non-Restorative Justice schools and as 1 for Restorative Justice schools, the difference in means in the t-test would be negative if the suspension percentage difference was higher in Restorative Justice schools. A left-tailed significance test would tell me whether such difference was in fact significantly smaller than zero (rather than just different from zero in either direction, as a two-tailed test would measure). In fact, without controlling for any other variables, Restorative Justice schools had a significantly higher suspension percentage difference between 2005-06 and 2009-10 as compared to non-Restorative Justice Schools. Suspension percentages declined, on average, by 3.3 percent in RJ schools, and by only .87 percent in non-RJ schools (see Table 1). This result was highly significant.

Table 1: T-Test Results for Difference in Suspension Percentage Difference Means Between RJ and non-RJ schools Without Matching

	N	Mean Suspension Percentage Difference	Standard Error
non-RJ schools	113	.8686	.455
RJ schools	30	3.328	1.089
Difference in Means		-2.459*	1.045
		P (T < t) = .01*	

Left-Tail Significance: ** p<0.01, * p<0.05

Next, I identified the imbalance in the control variables between schools in the treated group—i.e. schools that implemented Restorative Justice between 2005-06 and 2009-10—and schools in the control group—non-Restorative Justice schools—using the imb command that is part of the coarsened exact matching package in Stata. I found a Multivariate L1 distance[104] of 0.931 for the model as a whole and substantial imbalances for individual control variables, especially for suspension percentage and the level of school. I then identified potential cutoff points for coarsened exact matching[105] and performed coarsened exact matching. First, I matched on individual levels of school, as well as above and below the median for all other control variables. 22 RJ schools could be matched to 40 non-RJ schools. The multivariate L1 distance declined to 0.655, a significant improvement in model balance, and the following improvement in sample balance was achieved (see Table 2).

---

[104] Multivariate L1 distance is a value between 0 and 1 which refers to the "difference between the multidimensional histogram of all pretreatment covariates in the treated group and that in the control group." Matthew Blackwell et al., *cem: Coarsened Exact Matching in Stata*, Stata Journal (Feb. 22, 2010). The L1 value by itself is not very useful analytically, as it only gives a measure of overall imbalance in the sample among the control variables. However, it is useful as a comparison value for L1 scores that can be achieved after having matched the cases in the sample using coarsened exact matching.
[105] Coarsensed exact matching temporarily coarsens the data into a specified number of bins (or strata) that contain particular ranges of each control variable (as specified by the researcher) and then matches schools exactly based on what bin they are in for each control variable. Specifying different numbers of bins determines how many schools can be successfully matched.

Table 2: Difference in the Means of Control Variables Before and After Coarsened Exact Matching (Model 1)

	Enrollment (log)	Suspension Percentage 05 (log)	Level of School	Free Lunch Rate	Black Enrollment Percentage (log)
Differences in Means before Matching					
non-RJ schools	5.96	1.25	1.23	57.43	2.18
RJ schools	6.50	2.19	2	56.42	1.69
Differences in Means After Matching					
non-RJ schools	6.37	1.59	1.85	40.78	1.81
RJ schools	6.52	1.86	1.85	45.34	1.71

I then reran the t-test to see if the significance in the differences in suspension percentage difference had disappeared. Indeed, it had (see Table 3).

Table 3: T-Test Results for Difference in Suspension Percentage Difference Means between RJ and non-RJ schools After Coarsened Exact Matching (Model 1)

	N	Mean Suspension Percentage Difference	Standard Error
non-RJ schools	40	. 1.366	.739
RJ schools	22	2.619	.973
Difference in Means		-1.253	1.265
		P (T < t) = .16	

Left-Tail Significance: ** p<0.01, * p<0.05

Nevertheless, the mean Suspension Percentage Difference among RJ schools was still twice as large as that of non-RJ schools. Suspension percentages at RJ schools declined, on average, by 2.6 percent between 2005-06 and 2009-10 while suspension percentages at non-RJ schools declined by only about 1.4 percent.

I then ran my second model (without enrollment as a control variable), repeating the steps from above. Again, a significant improvement from an initial L1 imbalance of 0.836 to 0.767 could be achieved. 25 RJ schools were matched with 50 non-RJ schools. The following improvement in sample imbalance resulted from matching (see Table 4).

Table 4. Difference in the Means of Control Variables Before and After Coarsened Exact Matching (Model 2)

	Enrollment (log) – not matched in this Model	Suspension Percentage 05 (log)	Level of School	Free Lunch Rate	Black Enrollment Percentage (log)
Differences in Means before Matching					
non-RJ schools	5.96	1.25	1.23	57.43	2.18
RJ schools	6.50	2.19	2	56.42	1.69
Differences in Means After Matching					
non-RJ schools	5.89	1.92	2	52.70	1.95
RJ schools	6.51	2.01	2	51.62	1.75

The second model resulted in a sample that was much more imbalanced on enrollment (which was not matched in this model). However, a much better match was achieved on initial suspension percentage. Because suspension percentage seems to be the primary driver of Suspension Percentage Difference, this model is thus preferable to the first. The t-test on the balanced sample in Model 2 produced the following results (see Table 5).

Table 5: T-Test Results for Difference in Suspension Percentage Difference Means between RJ and non-RJ schools After Coarsened Exact Matching (Model 2)

	N	Mean Suspension Percentage Difference	Standard Error
non-RJ schools	50	1.585	.723
RJ schools	25	3.419	1.181
Difference in Means		-1.834	1.319
		$P(T < t) = .08$	

Left-Tail Significance: ** $p<0.01$, * $p<0.05$

The left tail significance of the difference in the means of Suspension Percentage Difference improved to a point of near-significance in this model. Again, the mean Suspension Percentage Difference at RJ schools, amounting to negative 3.4 percent, was more than twice as large as the reduction of 1.6 percent in non-RJ schools.

Finally, I retested this second model by matching even more closely on suspension percentage—the strongest predictor of suspension difference. I matched in all four quartiles of initial logged suspension percentage and left the other matching instructions equal. 22 RJ schools could be matched to 50 non-RJ schools. An L1 reduction from 0.836 to .75 was achieved. The following improvement in sample imbalance resulted from matching (see Table 5).

Table 5. Difference in the Means of Control Variables Before and After Coarsened Exact Matching (Model 3)

Differences in Means before Matching					
	Enrollment (log) – not matched in this Model	Suspension Percentage 05 (log)	Level of School	Free Lunch Rate	Black Enrollment Percentage (log)
non-RJ schools	5.96	1.25	1.23	57.43	2.18
RJ schools	6.50	2.19	2	56.42	1.69
Differences in Means After Matching					
non-RJ schools	5.96	1.95	1.91	52.36	2.10
RJ schools	6.46	1.97	1.91	50.04	1.83

The t-test on the balanced sample in Model 3 produced the following results (see Table 6).

Table 6: T-Test Results for Difference in Suspension Percentage Difference Means between RJ and non-RJ schools After Coarsened Exact Matching (Model 2)

	N	Mean Suspension Percentage Difference	Standard Error
non-RJ schools	50	1.585	.723
RJ schools	22	3.388	1.344
Difference in Means		-1.803	1.406
		P (T < t) = .10	

Left-Tail Significance: ** p<0.01, * p<0.05

The left tail significance of the difference in the means of Suspension Percentage Difference stayed at approximately the same level as with Model 2. Once again, the doubly as large reduction in suspension percentages at RJ schools compared to non-RJ schools appeared in the analysis.

Overall Suspension Percentage Ratio – "No Zero" (Natural Log)

With regards to my second outcome variable for overall suspension levels—the natural log of the suspension percentage ratio between 2009-10 and 2005-06—pairwise correlation analysis identified suspension percentage in 2005, level of school, and free lunch rate as potentially useful controls. Visual inspection of pairwise scatterplots, identified white enrollment percentage as a potentially useful control variable. Regression analysis led me to settle on a final model including suspension percentage, level of school, white enrollment percentage, either black or Hispanic enrollment percentage, and free lunch rate as control variables (see Appendix E).

Once again, I ran a t-test comparing the means of suspension percentage difference between the two groups without any control variables. For purposes of this variable I was interested in right-tail one-tail differences in means. A right-tailed significance test would tell me whether the difference in suspension percentage ratio means was significantly larger than zero. This would indicate that the ratio was higher for non-RJ schools, which in turn would mean that as compared to the baseline of 2005-06, suspensions would have declined more for RJ schools than for non-RJ schools. Without controlling for other variables, there was no significant difference in suspension percentage ratios (natural log) between non-RJ and RJ schools (see Table 7). The right-tail p-value was not significant at p = .26.

Table 7: T-Test Results for Difference in Suspension Percentage Difference Means between RJ and non-RJ schools Before Coarsened Exact Matching

	N	Mean Suspension Percentage Ratio (log)	Standard Error
non-RJ schools	70	-.255	.094
RJ schools	25	-.370	.121
Difference in Means		.115	.174
		P (T > t) = .26	

Right-Tail Significance: ** p<0.01, * p<0.05

Note: 48 cases were dropped because they either had a suspension percentage of 0 in 2005-06 and thus fell out of the analysis when the suspension percentage ratio was taken or had a suspension percentage of 0 in 2009-10 and thus fell out of the analysis when the natural logarithm was taken.

To avoid repetition of the steps described above with regards to Suspension Percentage Difference, I note only the results of the best model using this metric here. The best model (i.e. the model with both a high imbalance reduction and the greatest possible number of matched cases) was a model using suspension percentage, level of school, white enrollment percentage, Hispanic enrollment percentage, and free lunch rate as control variables and cutting the control variables (with the exception of school level) at their median. It matched 32 non-RJ schools with 20 RJ schools and achieved the following balance in the sample after matching (see Table 8).

Table 8. Table 5. Difference in the Means of Control Variables Before and After Coarsened Exact Matching (Best Model).

Differences in Means before Matching					
	Suspension Percentage 05 (log)	Level of School	White Enrollment Percentage (log)	Hispanic Enrollment Percentage	Free Lunch Rate
non-RJ schools	1.25	1.23	2.58	57.26	57.43
RJ schools	2.19	2	2.68	63.61	56.42
Differences in Means After Matching					

| non-RJ schools | 1.84 | 2 | 2.57 | 55.73 | 53.84 |
| RJ schools | 2.05 | 2 | 2.67 | 61.92 | 55.79 |

The t-test on the balanced sample in this Model produced the following results (see Table 9).[106]

Table 9: T-Test Results for Difference in Suspension Percentage Difference Means between RJ and non-RJ schools After Coarsened Exact Matching (Model 2)

	N	Mean Suspension Percentage Ratio (log)	Standard Error
non-RJ schools	32	-.071	.145
RJ schools	20	-.372	.144
Difference in Means		.301	.216
		P (T > t) = .09	

Right-Tail Significance: ** p<0.01, * p<0.05

These results mean that the suspension percentage in 2009-10 for RJ schools was only about 69 percent of what it had been in 2005-06, while the suspension percentage for non-RJ schools was about 93 percent of what it had been in 2005-06.[107] The fact that RJ schools had a lower ratio was almost statistically significant at p= .09.

Racial Disproportionality – Black Suspension Percentage Disparity

For this variable, it is important to first look at initial levels of Black Suspension Percentage Disparity at time 0 (2005-06). Consistent with the wealth of prior research on the persistence of black racial disproportionality in school discipline, there is significant such disparity in my dataset as well, a disparity which is greater in schools that would implement Restorative Justice between 2005-06 and 2009-10 than in those schools that would not. Black suspension percentage, on average, was more than 10 percent higher than white suspension percentage in RJ schools, whereas average black suspension percentage was "only" slightly less than 5 percent higher than average white suspension percentage in non-RJ schools (see Figure 1).

---

[106] A similar model which matched even more closely on suspension percentage and matched 17 RJ schools with 32 non-RJ schools closely mirrored this model but with a slightly higher right-tail p = .14.

[107] exp(-.372) = .68935424; exp(-.071) = .93146189.

Figure 1: Initial Black Suspension Percentage Disparity by Treatment in 2005-06

Graphs by RJ implemented
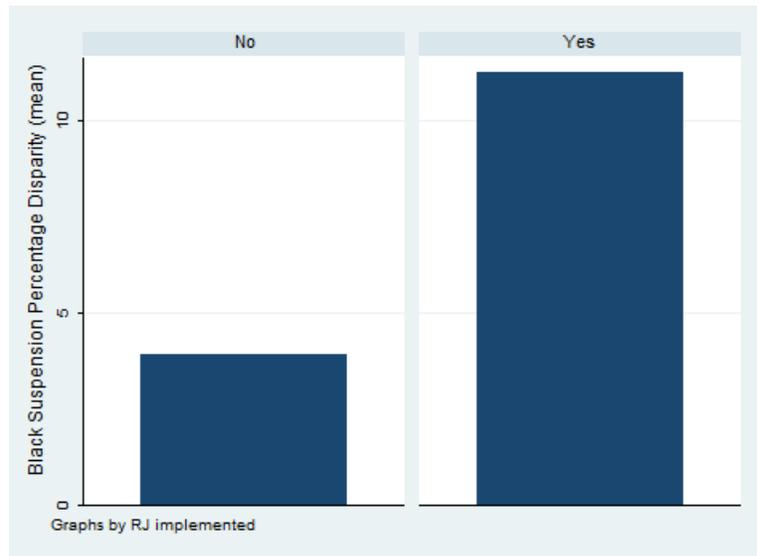
Therefore, I proceeded to investigate whether the implementation of Restorative Justice would significantly reduce such disparity—both absolutely and compared to non-RJ schools. As in my general suspension percentage analysis above, initial correlation analysis, visual inspection of correlation scatterplots, and regression analysis helped me identify the most appropriate control variables for a model comparing Black Suspension Percentage Disparity between non-RJ and RJ schools. While no particularly strong relationship with any control variables was detectable, level of school and initial suspension percentage emerged as potential control candidates. Because this racial disproportionality analysis again focuses on suspension percentage differences, my interest shifts back to a left-tailed test inquiring whether the reduction in Black Suspension Percentage Disparity was greater for RJ schools (producing a negative difference in means in the t-test) as compared to non-RJ schools).

As with all of the analysis above, I conducted an initial t-test without matching, which produced the following results (see Table 10).

Table 10: T-Test Results for Difference in Suspension Percentage Difference Means between RJ and non-RJ schools After Coarsened Exact Matching (Model 2)

	N	Mean Black Suspension Percentage Disparity Difference	Standard Error
non-RJ schools	45	-.862	1.01
RJ schools	13	4.57	3.01
Difference in Means		-5.436	2.47
		P (T < t) = .02*	

Left-Tail Significance: ** p<0.01, * p<0.05

Without controlling for other variables, schools that implemented Restorative Justice reduced their Black Suspension Percentage Disparity by significantly more than their non-RJ counterparts. In particular, RJ schools reduced their existing Black Suspension Percentage Disparity by about 4.6 percent, while such disparity actually increased in schools that had not implemented Restorative Justice. Matching schools on the level of school, above and below the median of initial suspension percentage, or in all 4 quartiles of initial suspension percentage only reduced the number of control schools with which Restorative Justice schools could be matched and increased the standard error around the estimate of the mean difference in Black Suspension Percentage Disparity for non-RJ schools (and thus increased the p-value of the model to a maximum left-tail p = .059).[108] The general results, however, remained the same across all matching models. That is, RJ schools significantly reduced their Black Suspension Percentage Disparity by a much greater amount than non-RJ schools. In fact, non-RJ schools across all models seemed to increase their Black Suspension Percentage Disparity slightly.

Discussion of Results

The current study produced a number of promising results. At the most general level, the results from this study of 143 primary, middle, and high schools in two large public school districts in the Mid-, Southwest suggest that the implementation of Restorative Justice programs can substantially reduce school reliance on punitive disciplinary measures. Schools that would implement Restorative Justice programs between time zero of this study, 2005-06, and time 1, 2009-10, seemed to be those schools that had larger initial problems with punitive school discipline to begin with. Restorative Justice schools showed a much larger initial suspension percentage than non-Restorative Justice schools. The first part of my study inquired into whether these schools were able to reduce their problems with high suspension rates, and if so, whether Restorative Justice seemed to have something to do with that. Results of a t-test that compared only the mean Suspension Percentage Difference between 2005-06 and 2009-10 indeed found that schools that had implemented a Restorative Justice program between these two years were able to reduce their initial suspension percentages by slightly more than 3 percent, a figure much greater than the approximately 0.9 percent reduction achieved by non-RJ schools. This difference in the reduction of suspension numbers was highly significant.

However, because an exploration of the relationship of Suspension Percentage Difference over time revealed that such differences are strongly driven by initial baseline suspension percentages—which is not surprising because schools with higher initial suspension percentages also have more to reduce in absolute terms—I needed to investigate whether it was actually the implementation of Restorative Justice that was driving this significant reduction, or the fact that RJ schools also happened to have larger initial suspension percentages. After controlling for initial suspension percentage, and a variety of other variables that are potentially related to Suspension Percentage Difference, however, the fact remained that RJ schools had reduced their suspension percentage by more than non-RJ schools. While the difference in the reduction of suspension percentages was no longer statistically significant at a p < .05 significance level, each model specification which matched RJ schools

---

[108] Matching schools on black enrollment percentage in 2005-06 also did not change the results.

with non-RJ schools on those variables found to predict Suspension Percentage Difference the most showed a reduction in suspension percentages for RJ schools that was twice as large as that found in non-RJ schools. The model that almost perfectly matched schools on the strongest predictor of Suspension Percentage Difference—initial suspension percentage—showed RJ schools outperforming non-RJ schools by about double (RJ school reduction in suspension percentage = 3.4 percent; non-RJ school reduction = 1.6 percent). There is only a 10 percent chance that this difference was not in part due to Restorative Justice. This is an encouraging finding given that results from this study likely undervalue the influence of RJ implementation. Even in an environment in which some schools in Santa Fe implemented Restorative Justice very close to the 2009-10 school year, and in which other schools in Denver also potentially used non-punitive approaches to solve their discipline problems, Restorative Justice implementation showed a strong effect in reducing the percentage of students suspended at Restorative Justice schools.

The second metric I used to measure whether Restorative Justice had been successful in reducing punitive school discipline—the logged ratio of a school's suspension percentage in 2009-10 divided by its suspension percentage in 2005-06—also showed promising results. This particular metric investigated only the results of those schools that had some suspensions in both years under investigation. This should not be a negative, however, because schools that had at least one year without any suspensions are not likely to experience the brunt of problems associated with punitive school discipline which prompted this study in the first place. Furthermore, this metric is useful because by using a ratio instead of a raw difference in suspension percentages, it does not privilege schools with a high initial suspension percentage as much.

A t-test that investigated whether the ratio of 2009-10 suspension percentage to 2005-06 suspension was significantly smaller for RJ schools—thus suggesting that suspension percentages had declined comparatively more for RJ schools than for non-RJ schools—showed that there was no significant difference between RJ and non-RJ schools on this metric—at least, without controlling for other factors that might drive the suspension percentage ratio. After controlling for those factors— including initial suspension percentage, level of school, white enrollment percentage, either black or Hispanic enrollment percentage, and free lunch rate—the difference in the suspension percentage ratio for RJ schools as compared to non-RJ schools moved in the correct direction: upward. In particular, in the best model which compared 32 non-RJ schools to 20 RJ schools, the mean logged suspension percentage ratio for RJ schools was -.372 and only -.071 for non-RJ schools. This means that the suspension percentage in 2009-10 for RJ schools was only about 69 percent of what it had been in 2005-06, while the suspension percentage for non-RJ schools was about 93 percent of what it had been in 2005-06.[109] Again, this difference was almost statistically significant (right-tail p = .09). Certainly, it is substantively significant in providing support for the proposition that RJ seems to work in reducing suspension percentages in schools that implement it.

Thus, this research provides a statistically more intricate confirmation of what anecdotal case studies had suggested in the prior literature: Restorative Justice helps in reducing school suspensions and allows schools to move away from being punitive in their resolution of disputes, keeping students in school instead.

---

[109]   exp(-.372) = .68935424; exp(-.071) = .93146189.

Finally, I investigated whether the implementation of Restorative Justice significantly reduced racial disproportionality in school discipline vis-à-vis African American students. In particular, I analyzed whether the disparity in black suspension percentage as compared to white suspension percentage—measured by the difference between black suspension percentage and white suspension percentage)—was reduced by a greater amount in schools that implemented Restorative Justice than in those that did not. I confined my analysis on this point to only those schools that had white as well as black enrollment of over 20 students. I did so because otherwise small fluctuations in total suspension numbers and/or enrollment numbers would have improperly skewed my results.

Once again, the results are positive. Without matching schools across a number of potential control variables, the 13 RJ schools that were part of this subsample reduced their black suspension percentage disparity by about 4.5 percentage points, while non-RJ schools actually increased their disparity by slightly less than 1 percent. Importantly, the difference in these two numbers is highly significant (left-tail $p = .02$). Even after controlling for those variables that might potentially influence these numbers, the results do not change. While the number of control schools fell when I attempted to match schools in this small subsample, thus increasing both the standard error around the non-RJ school estimate and increasing the p-value of the t-test as a result, in each specification of the model black suspension percentage disparity decreased substantially in RJ schools and increased slightly in non-RJ schools. Consequently, there is strong evidence that, at least in the two school districts under investigation in this study, Restorative Justice is helpful in addressing what has been a decade long problem of African American disproportionality in school discipline. It is possible that the process addressed in my theory section is at play. Schools implementing Restorative Justice give their students a chance to tell their story, and give others a chance to hear that story. To the extent that stereotyping and cultural prejudice drive African American racial disproportionality in school discipline, such disproportionality is likely to be reduced where disciplinary decisionmakers and stereotyped students are encouraged to face each other individually, as equal human beings, rather than in a punitive authority structure that is susceptible to the occurrence of "behavioral leakage." The statement of a teacher relating his experience with the positive power of Restorative Justice sums it up nicely:

"I was having a bad day. A student was being very disruptive and he wasn't going to back down. I attacked him with sarcasm and embarrassed him. He got really angry and came at me physically. We went to the dean of students, who asked me: ' How many days\ [meaning, 'How many days of suspension do you want to give him'? I knew I'd had a part in this; my sarcasm had set the boy off. I said to him: 'You can have three days of suspension or take part in a [Restorative Justice] circle.' He chose the circle. Administrators, counselors, students and teachers attended. He told everyone where he was coming from. Everyone said how they had been affected by the incident. We both apologized. It ended in hugs. The student was never disruptive or missed my class again."

Conclusion

I started my inquiry with the premise that punitive school discipline is a problem of massive proportions in the United States, a problem which affects millions of children every year. Consequently, the need for a potential remedy for such disparity is clear. Restorative Justice offers such a potential remedy, promoting dialogue and mutual understanding instead of punishment and distrust. Restorative Justice appears to have helped a significant

number of schools in two large public school districts tackle their large problems with high overall suspension levels as well as large African American disproportionality in school suspensions.

However, much research remains to be done. For one thing, the sample in this study suffered from a number of significant weaknesses and measurement errors introduced by the way in which the Department of Education collects its suspension numbers from public schools across the country. While rounding suspension numbers to whole counts of five might not affect the Department's research on nationwide patterns, it has the potential to seriously affect research within smaller school samples, especially where schools with low enrollment numbers are prevalent. This data problem is, of course, exacerbated in the realm of research into racial disproportionality in school discipline, because ethnic enrollment and suspension numbers are by definition smaller subsets of total suspensions and enrollment overall.

Furthermore, more in-depth research needs to be done on schools in which the level of Restorative Justice implementation is similar in intensity. While Restorative Justice is too complex to expect perfectly matched implementation across schools, implementation in this particular sample took place at different points in time and at different levels of intensity. Thus, much more granular research is possible and should be undertaken.

For the time being, however, this paper reports positive results for those who are passionate about responding to punitive school discipline with collaboration, dialogue, and mutual understanding. Turns out, it works!

Appendix A – Initial Suspension Percentages at Schools with and without Restorative Justice Programs

Figure 1: Mean Initial Suspension Percentages for Schools that Did and Did Not Implement Restorative Justice Programs between 2005-06 and 2009-10.
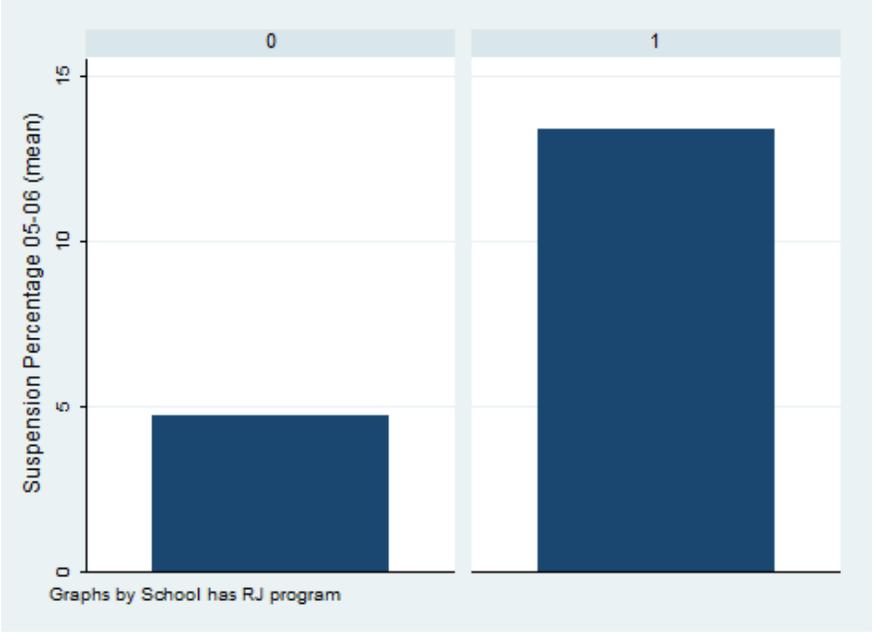
Appendix B – Distributions of Dependent Variables

Figure 2: Distribution of Suspension Percentage Difference
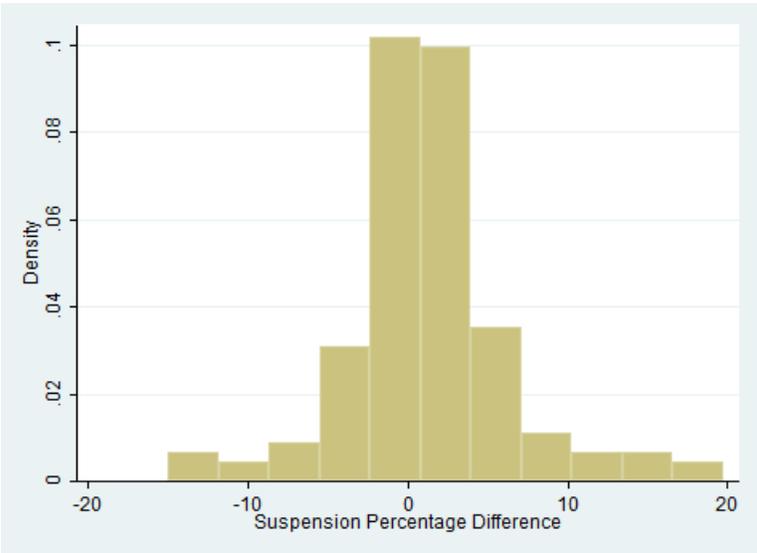
Figure 3: Distribution of Logged Suspension Percentage Ratio
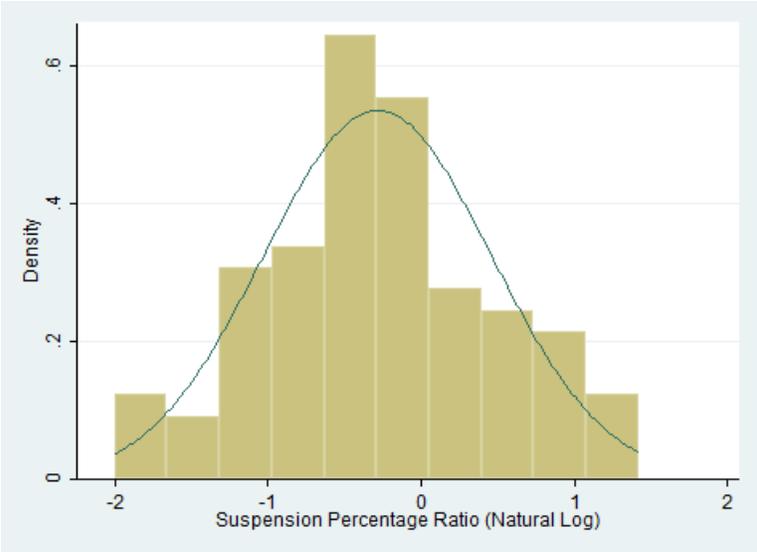
Figure 4: Distribution of Hispanic Suspension Percentage Disparity (Enrollment > 20)
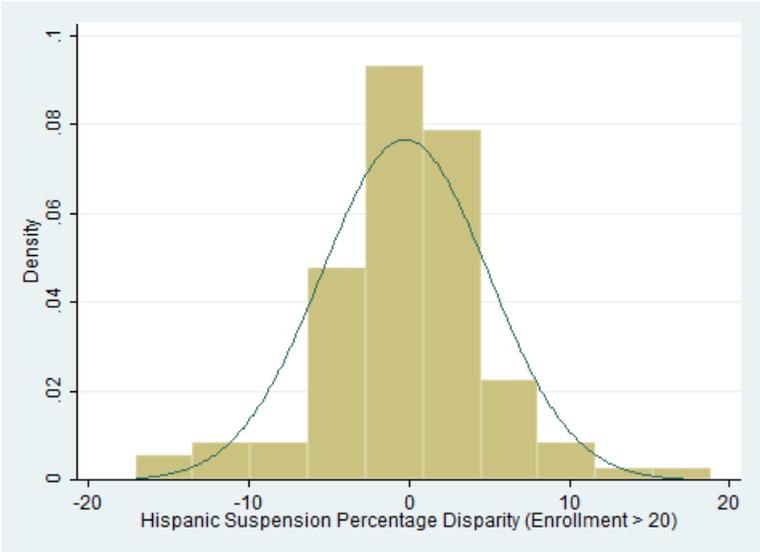
Figure 5: Distribution of Black Suspension Percentage Disparity (Enrollment > 20)
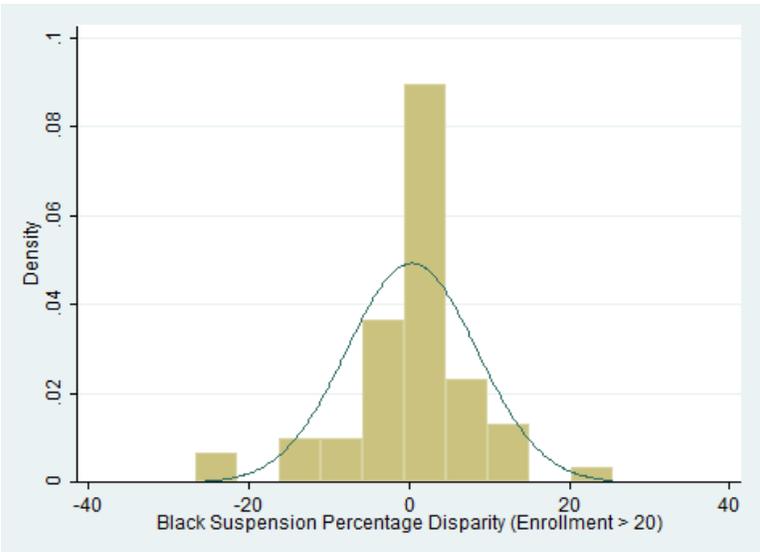
Appendix C: Distributions of Control Variables

Figure 6: Distribution of Suspension Percentage 05-06 and Logged (+1) Suspension Percentage in 05-06
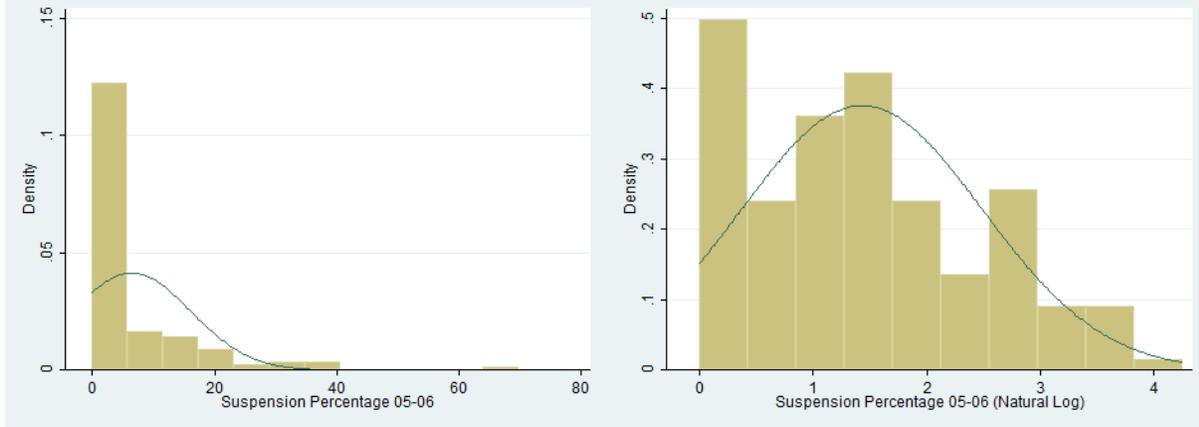
Figure 7: Distribution of Total Enrollment and Logged (+1) Total Enrollment
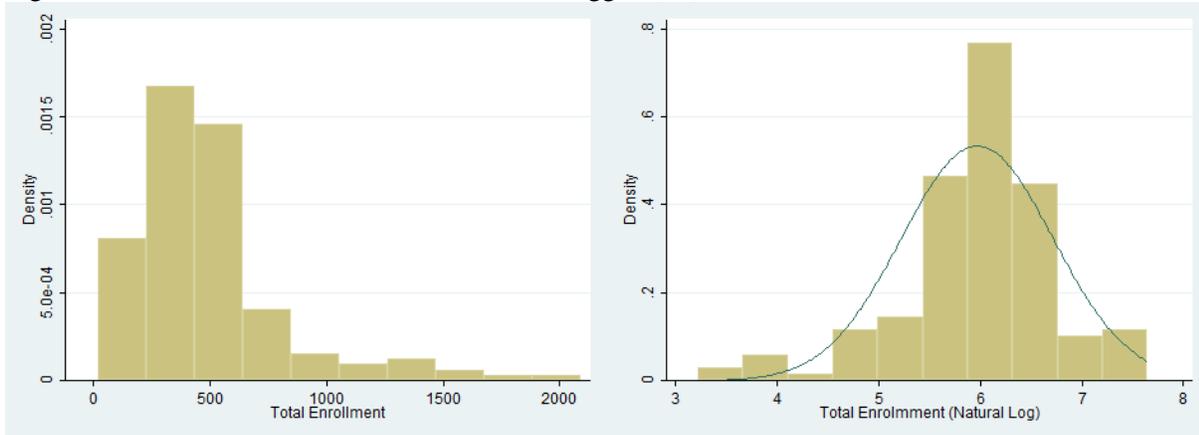
Figure 8: Distribution of Hispanic enrollment (untransformed)
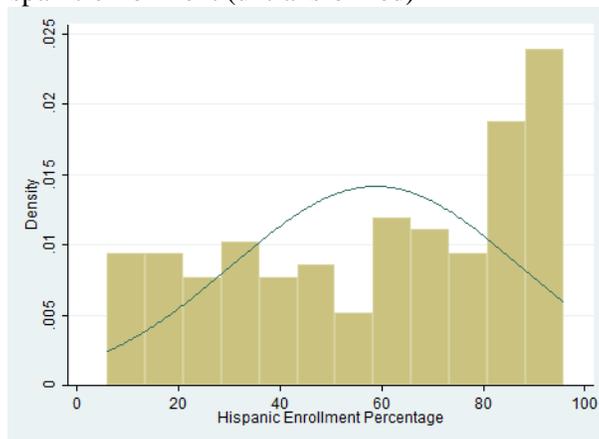
Figure 9: Distribution of Black Enrollment Percentage and Logged (+1) Black Enrollment Percentage
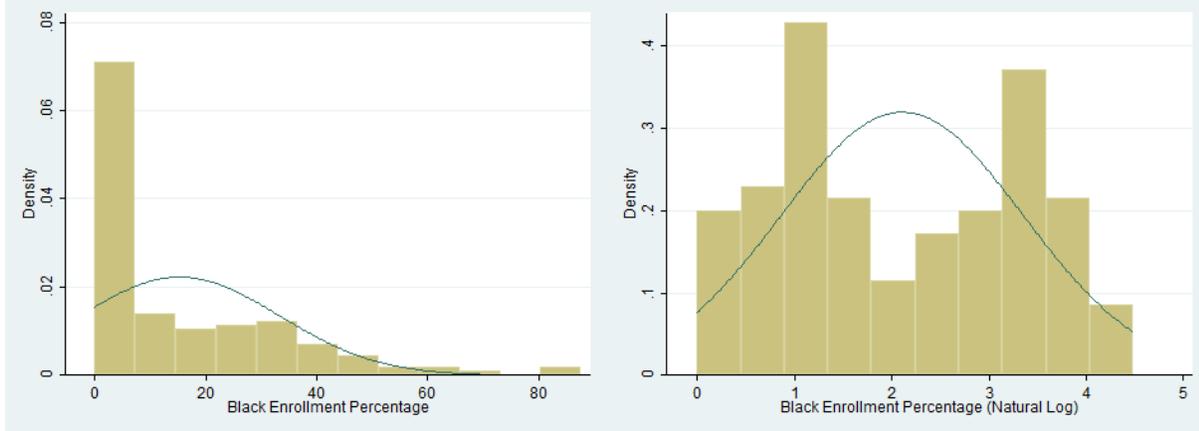
Figure 10: Distribution of Asian American Enrollment Percentage and Logged (+1) Asian American Enrollment Percentage
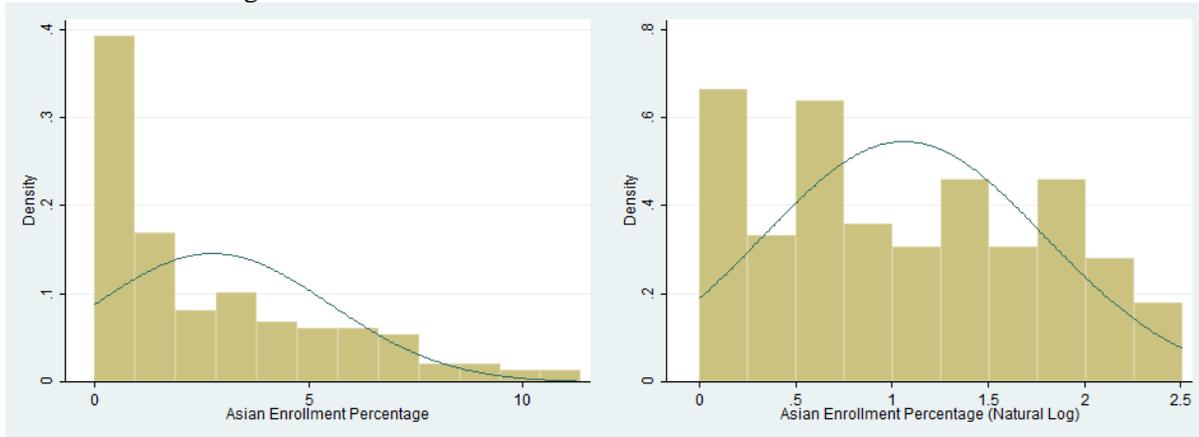
Figure 11: Distribution of American Indian Enrollment Percentage and Logged (+1) American Indian Enrollment Percentage
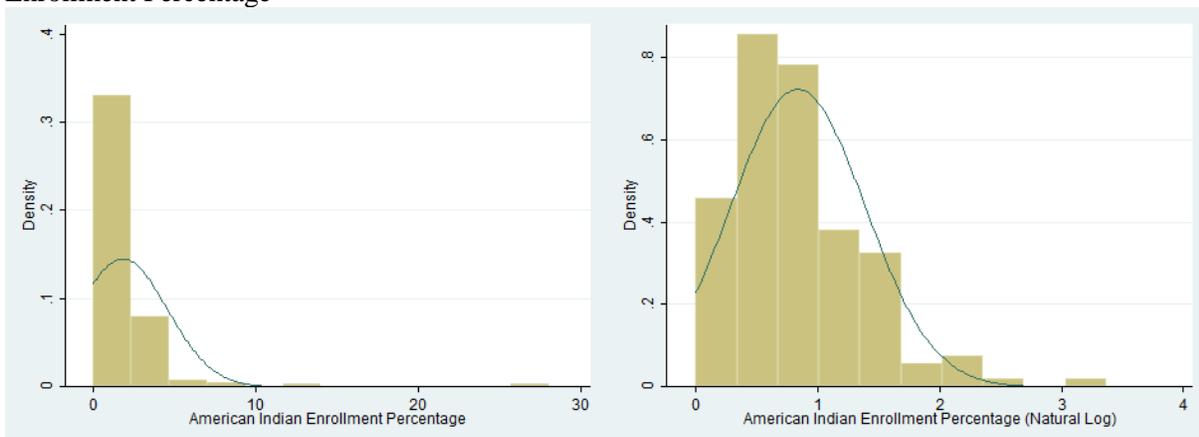
Figure 12: Distribution of White Enrollment Percentage and Logged (+1) White Enrollment Percentage
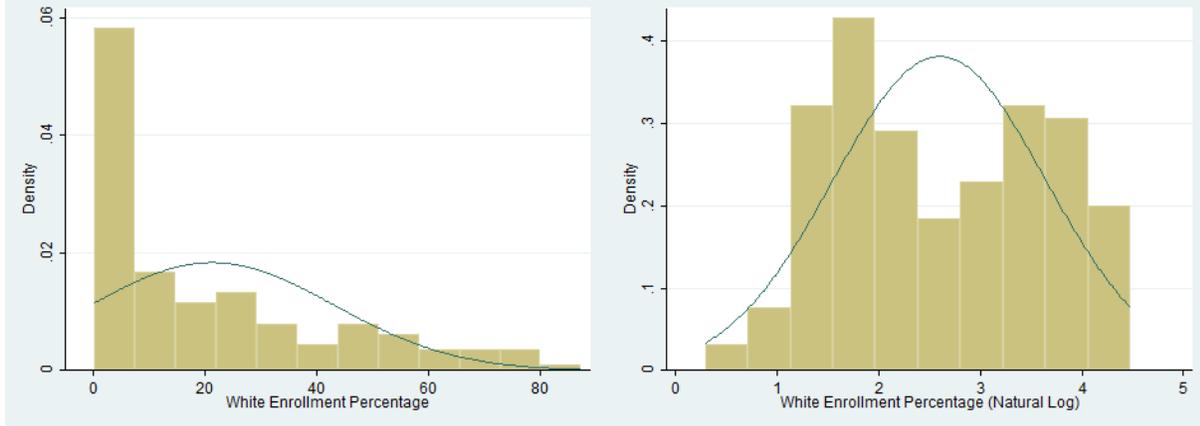
Figure 13: Distribution of Pupil Teacher Ratio and Square Root of Pupil Teacher Ratio
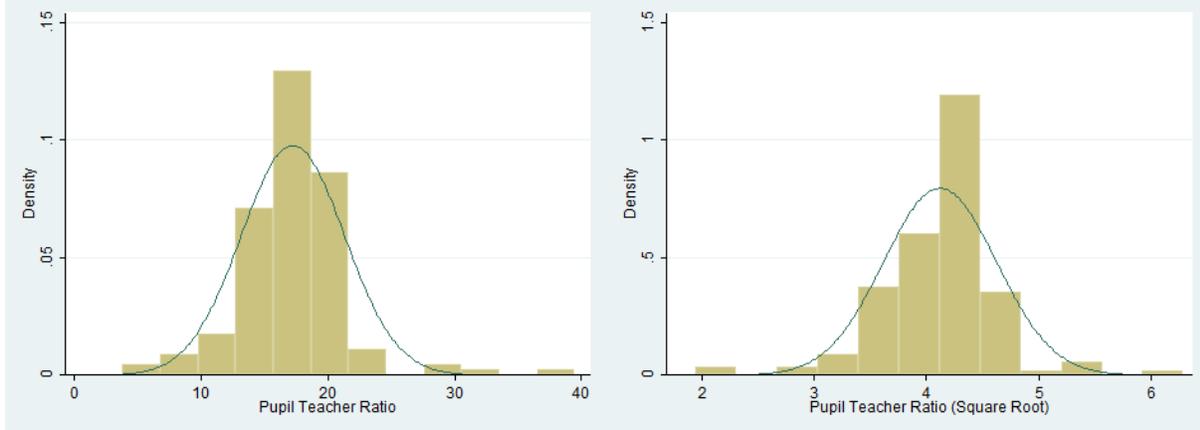
Figure 14: Distribution of Free Lunch Percentage
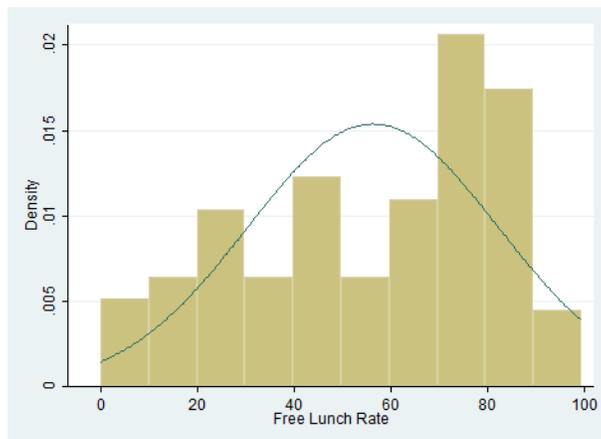
Appendix D: Regression Output of Models Predicting Suspension Percentage Difference

OLS Regression Analyses of the Predictors of
Overall Suspension Percentage Difference between 2005-06 and 2009-10

	Model 1	Model 2
Enrollment 05-06 (log)	-1.08 (0.63)	
Suspension Percentage 05-06 (log)	4.45** (0.54)	4.18** (0.54)
School Level 2	-1.46 (1.71)	-1.41 (1.74)
School Level 3	-3.23* (1.44)	-3.28* (1.43)
Free Lunch Rate 05-06	-0.04** (0.01)	-0.03** (0.01)
Black Enrollment Percentage 05-06 (log)	-0.42 (0.27)	-0.37 (0.27)
Constant	5.08 (3.64)	-1.28 (0.92)
Observations	143	143
R^2	0.48	0.47

Robust standard errors in parentheses

** $p<0.01$, * $p<0.05$

Note: Robust standard errors were used to control for heteroskedasticity in control variables.

Appendix E: Regression Output of Models Predicting Suspension Percentage Ratio

Table 1: OLS Regression Analyses of the Predictors of
the Square Root of Suspension Percentage Ratio between 2005-06 and 2009-10

Enrollment 05-06 (log)	0.23**
	(0.06)
Free Lunch Rate 05-06	0.01**
	(0.00)
Black Enrollment Percentage 05-06 (log)	0.11**
	(0.03)
Suspension Percentage 05-06 (log)	-0.20*
	(0.09)
School Level 2	0.25
	(0.13)
School Level 3	0.18
	(0.13)
Constant	-1.01*
	(0.39)
Observations	118
R^2	0.21

Robust standard errors in parentheses
** $p<0.01$, * $p<0.05$
Note: Robust standard errors were used to control for heteroskedasticity in control variables.

Appendix F: Scatter Graphs Plotting Dependent vs. Independent Variables

Suspension Percentage Difference as Dependent Variable
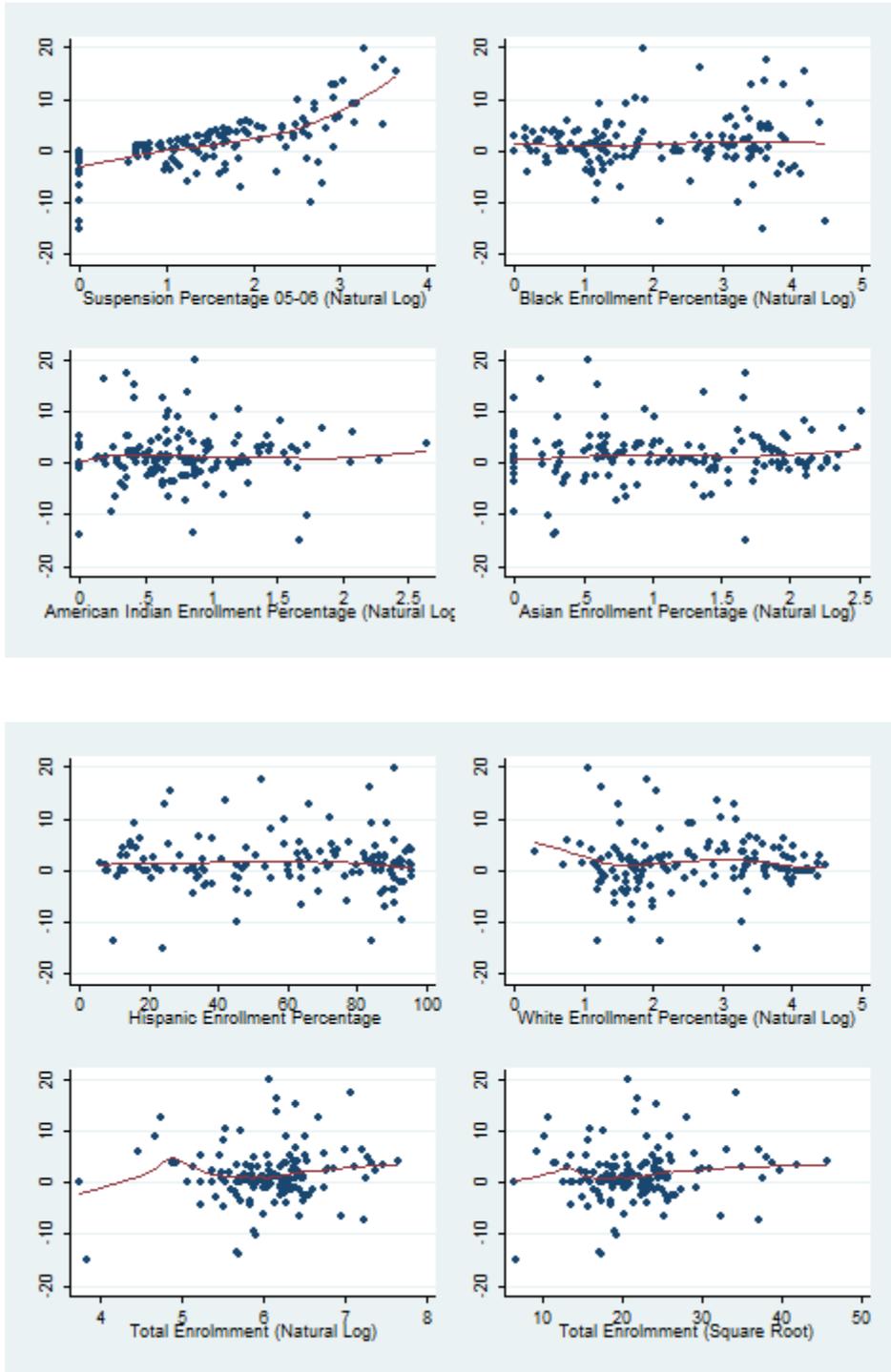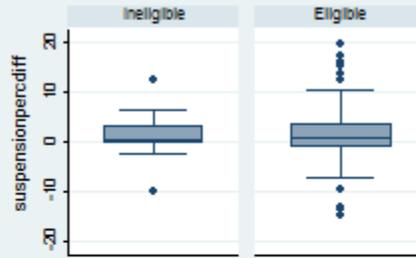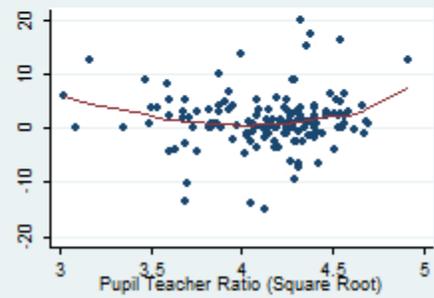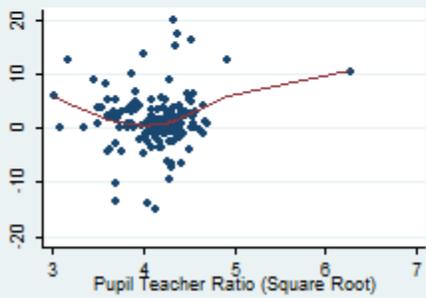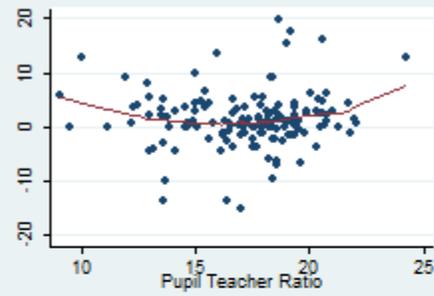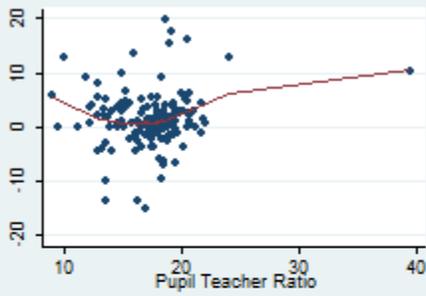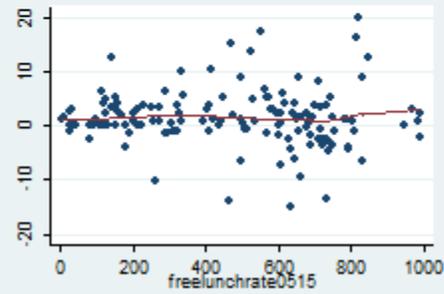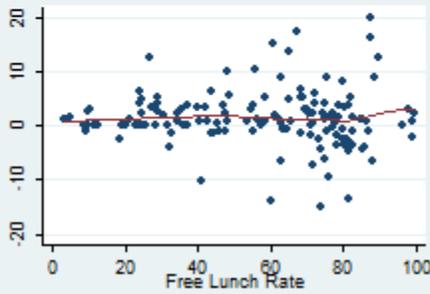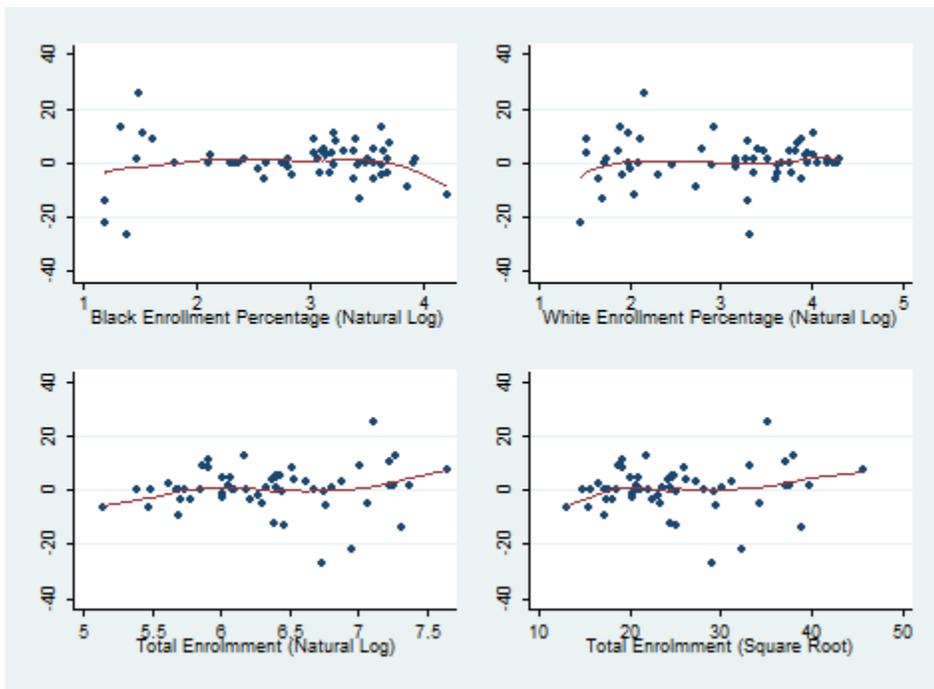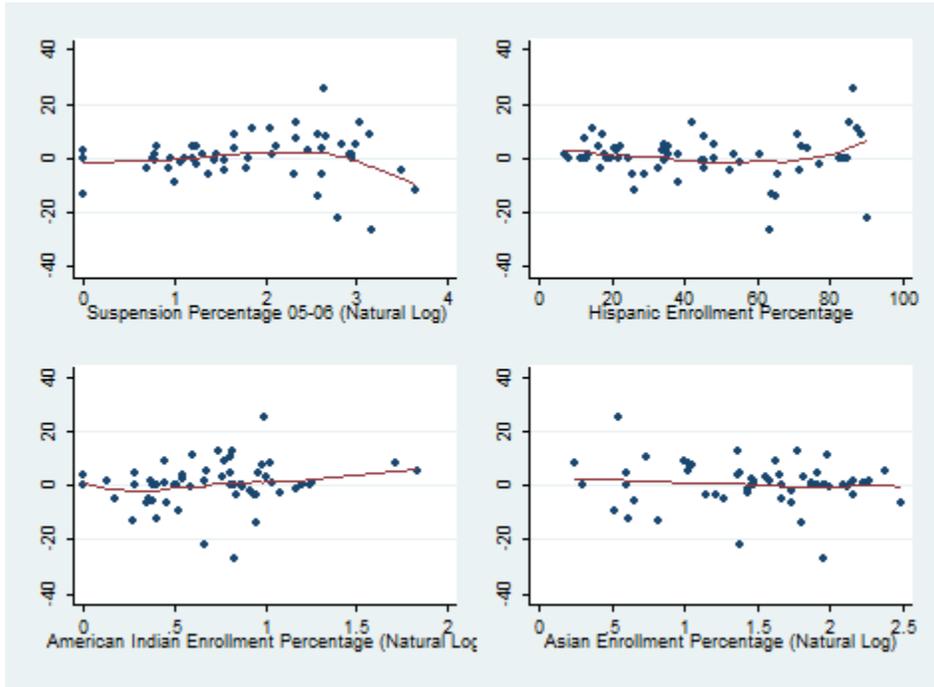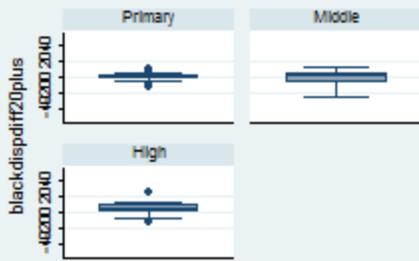
Suspension Percentage Ratio (log) as dependent variable
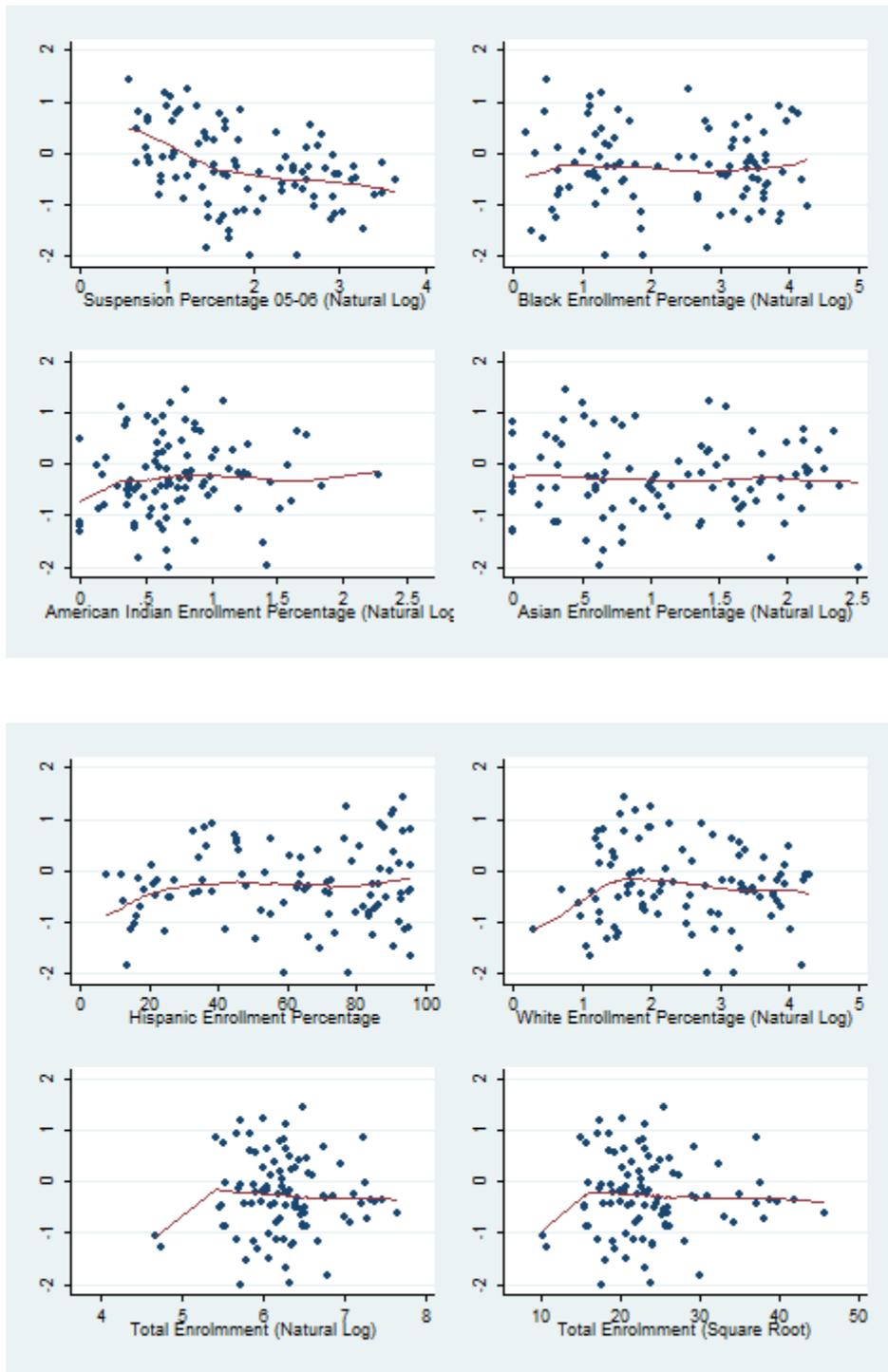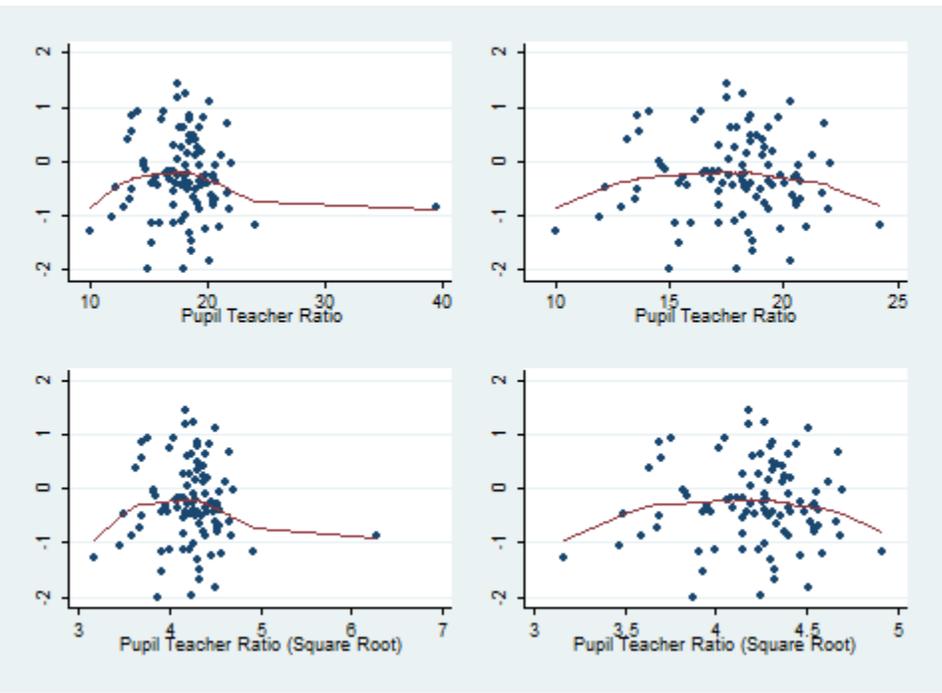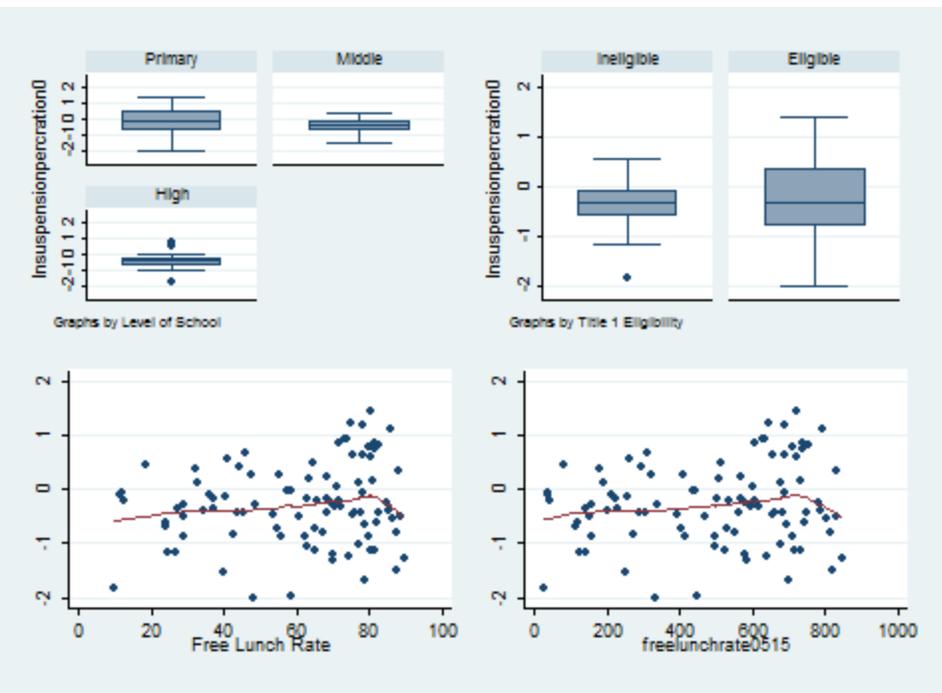
Black Suspension Percentage Disparity as Dependent Variable