Technical Report

IS THERE NO EXCUSE? THE EFFECTS OF THE NEW ORLEANS SCHOOL REFORMS ON SCHOOL DISCIPLINE

EDUCATION RESEARCH ALLIANCE FOR NEW ORLEANS

Mónica Hernández, Vanderbilt University and Tulane University

March 19, 2019

EducationResearchAllianceNOLA.org

Is There No Excuse? The Effects of the New Orleans School Reforms on School Discipline

Mónica Hernández[1]

Tennessee Education Research Alliance
Peabody College
Vanderbilt University

February 27th, 2019

Abstract

During the last decade, public debate about discipline practices that exclude students from schools has intensified. Some have also raised concerns that the increasing market-based and test-based school accountability could make exclusionary discipline worse. This paper studies the effect of the post-Katrina New Orleans school reforms on school discipline, as measured by out-of-school suspensions and expulsions, using a synthetic control method. The results indicate that the school reforms temporarily increased the rate of expulsions three years after the reforms initiated. This increase diminished immediately, returning expulsion rates in New Orleans back to its pre-Katrina levels. The results are qualitatively similar for out-of-school suspensions, though these results are less precise and robust.

[1] Research Associate. monica.hernandez@vanderbilt.edu. This article was mostly developed during my postdoctoral fellowship at the Education Research Alliance for New Orleans at Tulane University. Thanks to Douglas Harris, Nate Barrett, Brian Jacob, Christian Buerger, Briana Diaz, Sara Slaughter and Lindsay Weixler for helpful comments, as well as attendees at APPAM, the CIERS seminar at the University of Michigan and the economics seminar at EAFIT University and the Central Bank of Colombia. All errors are my own.

1. Introduction

The last decade has seen a shift in the national discourse on school discipline. The attention has moved from zero-tolerance policies to alternative discipline policies that keep schools safe and preserve learning opportunities for all students (APA, 2008).[2] Schools have been encouraged to adopt positive school climate strategies and to reduce the use of exclusionary discipline. However, the prevalence of out-of-school suspensions and expulsions is still high, especially among minority students. In 2014, 5.3% of all students in public schools were suspended out-of-school and 0.2% were expelled. These same rates were 2 to 3 times higher among black students, 13.5% of whom were suspended out-of-school and 0.4% of whom were expelled.[3]

This push for the reduction of exclusionary discipline practices has happened in the midst of several important and ambitious school reforms. From New York to New Orleans, school systems have implemented reforms with market-based accountability, through the expansion of the charter school system, and test-based accountability, through the establishment of sanctions for poor school performance. Several of these reforms have been linked to improvements in student academic achievement.[4] However, many critics have raised concerns about the reforms' potential to increase disciplinary incidents in schools. Market-based accountability may encourage the growth of No Excuses charter school models that display higher suspension and expulsion rates due to their strict discipline policies (Angrist, Pathak and Walters, 2013). Test-based accountability, in addition, can incentivize low-performing schools to game the testing system by suspending or expelling low-performing students around testing time (Figlio and Gletzler, 2006; Jacob, 2005; Cullen and Reback, 2006; Figlio, 2006).

---

[2] Zero-tolerance refers to 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 (APA, 2008).

[3] Author's calculations based on the U.S. Department of Education Office of Civil Rights data.

[4] For evidence regarding test-based accountability reforms see Ladd (1999), Jacob (2005), Figlio and Rouse (2006), and Dee and Jacob (2011), among others. For evidence regarding market-based accountability reforms see Abdulkadiro, Angrist, Dynarski, Kane & Pathak (2011), Dobbie and Fryer (2011) and Harris and Larsen (2015).

This paper studies the effect of the post-Katrina New Orleans school reforms on school discipline, as measured by student out-of-school suspensions (OSS) and expulsions. The New Orleans school reforms, which were part of the efforts to rebuild the city after the devastation caused by Katrina, included several elements of market-based and test-based accountability. The reforms established a charter school-based system with autonomous schools that make their own decisions, including discipline policies and practices, and compete with one another for student enrollment. They also gave the authorizers responsible for overseeing New Orleans schools the power to take-over low performing schools and significantly changed the teacher workforce.[5]

The mechanisms through which the New Orleans school reforms could have affected school discipline are varied. School choice, higher performance-based accountability, a transformed teacher workforce with different demographic characteristics, and decreased regulation of school policies and practices within the charter-based school system could have all affected students' behavior and/or school discipline policies and practices. Drawing upon the literature that studies school choice, school accountability, charter schools and racial discipline disparities,[6] this paper sheds light into how these mechanisms could have operated to affect school discipline.

New Orleans represents the most intense test-based and market-based accountability reform ever implemented in the United States (Harris and Larsen, 2016). Recent studies have shown that the reforms have had promising early effects on students' academic achievement (Harris and Larsen, 2016; Harris and Larsen, 2018; Ross, Harris and Liu, 2016), and other large urban school districts have started adopting similar charter-based portfolio strategies. According to the Center for Reinventing Public Education (CRPE), one of the main advocates of portfolio

[5] For a detailed description of the New Orleans school reforms, see Harris and Larsen (2015).
[6] For a thorough review of this literature, see Barrett, McEachin, Mills and Valant (2017).

strategies, currently there are 39 districts that belong to its portfolio-strategy network.[7]

In order to assess the effect of the New Orleans post-Katrina school reforms on discipline, this paper uses a synthetic control method. In this method, I compare the change in disciplinary incidents among New Orleans students, relative to the change in disciplinary incidents among a synthetic control group of students. I construct the synthetic control using two samples. The first is a pooled sample of students enrolled in Louisiana public schools, representative of all students in the state. The second is a panel of students who returned to their original school districts after the hurricane, which allows controlling for changes in student population that happened with Katrina.

One of the main threats to identification in empirical papers that study school discipline is measurement error, particularly from how incidents are reported. The number of *actual* discipline incidents are not observed, only those *reported* by schools. Similar to studies of crime, the difference is in whether students get caught and whether school personnel choose to make an official record. To address this threat, I perform a robustness check where I restrict the outcome to out-of-school suspensions for serious and specific offenses, which tend to be reported with more accuracy.

The results indicate that the school reforms temporarily increased the rate of expulsions in New Orleans schools. By 2009, three years after the reforms began, expulsions were 1 to 3 times higher than their control counterpart. This large increase was not driven by the establishment of new charter schools. Instead, it was driven by public schools directly run by the statewide Louisiana Recovery School District (RSD) that were in transition to become charter schools.[8] One year later, in 2010, the increase in expulsions started to reverse, coinciding with the increased political and legal pressures coming from a lawsuit challenging the discipline

[7] This list includes New York City, Los Angeles, Philadelphia, Chicago, Detroit, New Orleans, Memphis, Nashville, St. Louis, Cleveland, and Denver.
[8] RSD is a statewide government organization that was created before Katrina to take over and close underperforming schools.

practices applied to special education students.[9] By 2012, six years after the reforms started, the increase in expulsions had reversed completely, and school exclusionary discipline rates returned to pre-reform levels.

The results regarding out-of-school suspensions suggest sustained increases, although the findings depend on the severity and specificity of the offenses. The rate of suspensions for specific offenses (i.e. offenses that are well-defined, such as being habitually tardy or getting involved in fights) increases between 40% and 60%, but the result depends on the sample used. To check if inaccurate reporting is affecting the results, I focus on suspensions for serious offenses, which tend to be recorded with more accuracy. By serious offenses I mean offenses that are well defined and severe, such as getting into fights or having habits that injure others. In this case, the analysis reveals effects that are larger and more robust, with increases ranging between 60 and 90%, depending on the sample.

This paper is the first to assess the effects of school reforms that combine both market-based and test-based accountability on school discipline. Prior studies examining the effect of school reforms on discipline incidents have focused on test-based accountability. Holbein and Ladd (2017) provide evidence that North Carolina schools failing to make Adequate Yearly Progress (AYP) under NCLB experienced increases in reported students' misbehaviors. Chiang (2009), on the other hand, found no evidence of increases in disciplinary incidents in Florida schools that did not make AYP under NCLB. This paper adds to this nascent literature by studying the first reform that paired increased test-based accountability with a charter-school based system and a turnover of the teacher workforce.

The rest of this paper is organized as follows. Section 2 provides the context of the reforms and a conceptual framework to understand their effects on school discipline. Section 3 explains the empirical strategy that shapes the research design, while Section 4 describes the data. Section 5 presents the results of the synthetic

[9] The Southern Poverty Law Center (SPLC) filed a federal lawsuit in 2010 against the Louisiana Department of Education, related to New Orleans schools' disciplinary practices with special education students. The lawsuit documented how New Orleans schools violated federal laws stating that students cannot be disciplined for behaviors that are a manifestation of their disabilities.

control method. Section 6 presents robustness checks, while Section 7 discusses and concludes.

2. Context and Conceptual Framework

2.1. The New Orleans School Reforms

Prior to Hurricane Katrina, New Orleans was one of the lowest performing school districts in the United States. Seventy percent of eighth grade students were below proficiency level in math and 74 percent in English. The Orleans Parish School Board (OPSB) had earned a reputation for corruption and incompetence, since it could not account for $71 million in federal money. When Katrina hit in August 2005, 110 school buildings, out of 126, were destroyed. The entire school population, consisting of around 60,000 students and 4,300 teachers, was forced to relocate. Given the effort that rebuilding the education system required, education leaders decided to use this opportunity to turnaround the school system and implement one of the most aggressive and ambitious school reforms in the recent history of the United States (Harris, 2015; Sims and Rossmeier, 2015; Harris and Larsen, 2016).

With the school reforms, New Orleans changed from a traditional local school district that allocated resources and ran school operations, to a charter school-based system with few attendance zones and autonomous schools making their own decisions, including discipline policies and practices. This created market-based accountability, with families choosing amongst a larger set of schools and, therefore, schools competing to attract and keep their students. Test-based accountability for schools also increased. Before the storm, school closure decisions were based on student enrollment, finances and/or political decisions. After the reforms, school closure and takeover depended solely on schools' performance on standardized achievement tests and high school graduation (Bross, Harris, & Liu, 2016). This means that a charter school's students had to perform well on achievement tests in order for that school to remain open.

The state government moved oversight of almost all the city's public schools from the locally elected OPSB, to the statewide Louisiana Recovery School District (RSD). Most OPSB schools were quickly turned into charter schools, and over the first decade, so too were all RSD schools. While these changes happened, RSD was directly in charge of running many of the public schools that would become charter schools. Through the reform, the school system leaders (OPSB and RSD) switched their focus from operating to authorizing and overseeing charter schools (Harris and Larsen, 2016).

Finally, the teacher workforce changed significantly with the school reforms. In the aftermath of Katrina, the teacher union contract expired and was not renewed. All teachers were fired. Charter schools, which now make up more than 95% of the city public schools, were not required to hire certified teachers. This combination of policy shifts led to significant changes in the teacher workforce, with an influx of new teachers from Teach for America, The New Teacher Project, TeachNOLA, and other non-university-based alternative preparation programs. The percent of black teachers in New Orleans dropped from 71% to 49%, while the percent of students belonging to minority groups remained at around 95%. A similar decline was observed in the proportion of teachers with local roots and in teachers' years of experience (Barrett and Harris, 2015).[10]

2.2. The Anecdotal Evidence on the Effect of School Reforms on Discipline

New Orleans education leaders faced challenges unique to this setting and openly admitted to learning as they went (Gross, Tuchman and Yatsko, 2016). In creating an entirely new system of schooling, New Orleans leaders performed a balancing act between establishing new schools and developing an entirely new governance structure with new institutions to recruit and develop charter school operators (e.g., New Schools for New Orleans), recruit a new teacher workforce to the city (e.g., Teach for America and TeachNOLA), and provide information to families to help them choose schools (New Orleans Parents Guide). The state RSD

[10] The percentage of teachers who graduated from New Orleans-based colleges dropped from 60% in 2005 to 34% in 2014.

existed prior to Katrina but had only a handful of staff and had not been designed to carry out its new responsibilities.

Schools were given autonomy over discipline policies and practices (Charpentier, 2008). RSD did establish a Student Code of Conduct in 2008 that outlined the due process for expulsions and out-of-school suspensions in RSD direct-run schools (RSD, 2008); however, the due process was not enforced and, as reported by a chief judge of the Orleans Parish Juvenile Court, students could be "out of school for up to two months because of mistakes, abuses and misunderstandings over the process" (Charpentier, 2008). Schools might not have followed the Code of Conduct because of the transitional nature of the RSD and because enforcement would have been difficult, especially with RSD's small staff. RSD schools were aware that they were subject to potential closure or takeover, so they had strong incentives to create measurable academic success.[11]

Indeed, it is possible that some school leaders would have felt encouraged to use exclusionary discipline at high frequencies. Paul Vallas, the RSD superintendent, instituted some strict policies to handle school safety and discipline. Vallas hired a security firm and a former New Orleans Police Department commander to ensure safety at RSD schools (Simon, 2007a; Reckdahl, 2007). During the 2008 school year, RSD officials encouraged several RSD elementary schools to introduce parts of the KIPP model, whose charter schools are highly structured and script student movement throughout their buildings (Carr, 2008).

In the 2009 school year, community stakeholders started voicing concerns about the alarming rates of expulsions and out-of-school suspensions in some New Orleans schools. Child advocacy groups denounced the regular use of exclusionary discipline practices for minor misbehaviors, disproportionately targeting black students (Sullivan and Morgan, 2010). As Figure 1 shows, the percent of students expelled from schools reached a peak in 2009. This peak was driven by increased

[11] Although the reform legislation did not state explicitly that all RSD schools would be converted to charter schools, it was clear that the goal was having a majority-charter school district. Paul Vallas, the first RSD superintendent, declared that "We (RSD) want to have a system of charters and independent schools (…) We're shifting to reconstitute charter schools from failing schools" (Simon, 2007b).

discipline rates in schools that had not yet converted to charter schools and that, in the meantime, were run directly by RSD (Figure 2).

Partly because of the increase in exclusionary discipline, the Southern Poverty Law Center (SPLC) filed a federal civil rights lawsuit in 2010 against the Louisiana Department of Education (LDOE) and the Board of Elementary and Secondary Education (BESE). The lawsuit documented how New Orleans students with disabilities were excluded from schools and disciplined without the procedural safeguards required by federal law,[12] how RSD schools disciplined special education students at much higher rates than the state average, and how students were suspended or expelled because of manifestations of their disabilities (P.B. vs. Pastorek, 2010).

The SPLC lawsuit attracted national attention and exposed New Orleans' discipline disparities (Samuels, 2010; Mock, 2010). The lawsuit and the community complaints surfaced against a backdrop of growing national attention on out-of-school suspensions and expulsions and the "school-to-prison pipeline" (Gross, Tuchman and Yatsko, 2016). Under mounting political and legal pressure, the RSD fully centralized public school expulsions at the beginning of the 2012-2013 academic year. The new centralized expulsion system, run by RSD's Student Hearing Office, established common criteria for expelling a student and a hearing process that enforces the application of those criteria to expulsions.[13] Both RSD direct-run schools and RSD charter schools had to abide by the new regulations. OPSB schools soon agreed to comply with them, as well. However, schools maintained their autonomy over suspensions.

After the 2009 peak, New Orleans witnessed a sharp decrease in the expulsion rate, compared to the rest of the state (Figure 1), as well as a smaller decrease in the

[12] The lawsuit stated that LDOE failed to comply with its general supervisory responsibilities under the Individuals with Disabilities Education Improvement Act of 2004 (IDEA). The violations included the failure to abide by IDEA's procedural safeguards for discipline. For instance, the lawsuit documented the case of a student who had been suspended out of school and restrained and held against his will several times throughout the school year, as a consequence for manifestations of his disability (P.B. vs. Pastorek, 2010).

[13] Gross, Tuchman and Yatsko (2016) describe in great detail the process leading to the centralized expulsion system and how it works.

rate of out-of-school suspensions (Figure 3). This decrease, however, was not driven by the centralized expulsion system. Expulsions began to decrease in 2010, coinciding with the SPLC lawsuit and the national attention it brought to New Orleans discipline disparities. The centralized expulsion system came later in 2012, when the expulsion rate had already reached its lowest level in over two decades.

2.3. The Mechanisms Behind the Effect of School Reforms on Discipline

The literature on test-based accountability, market-based accountability and racial discipline disparities provides a conceptual framework to understand the mechanisms through which the reforms might have affected school discipline.

2.3.1. Market-based school accountability

Market-based school accountability implies school choice for families, which could affect school discipline in different ways. First, the heightened competition, might motivate schools to suspend and/or exclude students who are low performing or whose behavior disrupts learning in the classroom, thus making the school look better on state accountability metrics and making the school more attractive to potential attendees. Hoxby (1999) provides evidence in support of this argument, showing that school leaders in intense competition districts report responding to classroom misbehavior with stricter discipline.[14]

Second, families may have preferences about discipline practices, which could influence parents' choice of schools and therefore compel schools to move towards policies that are favored more by parents (Hoxby, 1999). In a study of Texas charter schools, Weiher and Tedin (2002) find that discipline is the second most important school attribute considered by Texas charter parents.[15] Steele, Vernez, Gottfried and Schwam-Baird (2011) also find that New Orleans charter school parents report discipline as one of the three most important school attributes. However, the

[14] Hoxby (1999) studies school choice based on the number of school districts (and thus choice among school districts) that each metropolitan area offers.
[15] It is not clear whether parents favored more strict or flexible discipline policies.

direction of the effect in this case is unclear as some parents may prefer strict and exclusionary discipline and others may prefer more positive approaches.[16]

Another aspect of market-based accountability that might impact school discipline is that it provides charter school leaders with autonomy over what happens in their schools, including discipline policies and practices. Some argue that having autonomy is important because discipline is part of the school culture and school leaders know their students best (Medler, 2016; Greene, 2016; Petrilli, 2016). According to this strand of the literature, autonomy would allow school leaders to implement discipline practices that are the best fit for their teachers and students.

In practice, there is mixed evidence on how charter schools affect discipline. On one hand, Imberman (2011) provides evidence that students in start-up charter schools[17] experience large reductions in disciplinary infractions, which are lost once students return to regular public schools. On the other hand, several authors have expressed concern over the discipline policies and practices of some charter schools, particularly No Excuses schools. The No Excuses label commonly refers to schools with high academic expectations and a college-going culture (Thernstrom and Thernstrom, 2003; Carter, 2000). These schools also often feature strict behavior codes, extended instructional time, and targeted instruction (e.g., tutoring) for low-performing students (Whitman, 2008).

In New Orleans, many perceived the No Excuses model to be deeply ingrained in some charter schools.[18] While the use of a No Excuses strategy has demonstrated improvements on student achievement (Angrist, Pathak, and Walters, 2013; Dobbie and Fryer, 2016; Cheng, Hitt, Kisida and Mills, 2017), it has also been controversial due to its strict disciplinary regime. In a systematic review of the evidence on No

[16] If parents think their children might benefit more from strict discipline practices and policies, then suspensions and expulsions might increase with school choice. If, on the other hand, parents prefer more positive approaches towards discipline, suspensions and expulsions might decrease.

[17] According to the author, start-up charters are schools with voluntary enrollment that begin as charters. In contrast, conversion schools are those that were previously traditional public schools and later converted to charter status (Imberman, 2011).

[18] "A growing number of schools, particularly charters, embrace a 'no excuses' or 'whatever it takes' attitude toward closing the achievement gap" (Carr, 2010). "The outcomes driven, no-excuses model is common among charter schools in New Orleans" (Williams, 2014).

Excuses schools, Angrist, Pathak, and Walters (2013) found that No Excuses schools have higher out-of-school suspensions rates than other types of urban charter schools. Some argue that the No Excuses schools are paternalistic and punitive (Boyd, Maranto and Rose, 2014; Horn and Wilburn, 2013; Lack, 2009) and that their strict rules may reduce students' aspirations (Goodman, 2013). According to these critics, the introduction of No Excuses charter schools could have led to a rise in expulsions and out-of-school suspensions, without a governing entity to ensure that sanctions were fair and consistent.

2.3.2. Test-based accountability

The second mechanism through which the reforms could affect school discipline is related to intense test-based accountability. The closure and takeover policy that came with the school reforms meant more pressure for schools, in addition to the accountability that already existed with NCLB. The higher accountability could have altered schools' incentives to change the composition of their student bodies, by removing distracting and low-performing students from the classroom and the school. Any of these strategic behaviors could have increased the rate of expulsions and out-of-school suspensions.[19]

The introduction of NCLB in 2002 provides evidence of how test-based accountability could alter the incentives of New Orleans schools. When NCLB was introduced in 2003, New Orleans schools displayed an abrupt and atypical rise in the expulsion rate (Figure 1, Panel A). A large part of this rise happened in March, when standardized achievement tests were administered (Figure A.4 in the Appendix). It is plausible that school administrators initially reacted to NCLB by expelling low-achieving students around testing times. Section 5.1 describes this point in further detail.

Besides altering schools' incentives, accountability pressures could have also changed how students behave when they are in school, increasing their level of

[19] Figlio (2006) finds that pressured schools in Florida assigned longer disciplinary suspensions to low-achieving students in tested grades around testing dates. This, however, is less likely to happen in New Orleans, because LDOE requires Louisiana students to take the tests, even if they are in the middle of a suspension.

anxiety or decreasing their motivation (Wheelock, Haney, and Bebell, 2000; Hoffman, Assaf, and Paris, 2001; Jones, 2007). Holbein and Ladd (2017) provide evidence that North Carolina schools failing to make AYP under NCLB saw increases in incidents that led to out-of-school suspensions, as well as increases in sexual offenses and offenses that are reportable to law enforcement agencies. Chiang (2017), on the other hand, found no evidence of increases in disciplinary incidents in Florida schools that faced the threat of closure or takeover.[20]

2.3.3. Changes in the teacher workforce

The third and final potential mechanism behind the change in discipline is the change in the teacher workforce. The workforce that emerged post-reform was younger, had less experience, fewer local roots and was less black in proportion. These changes raised concerns about the new teachers' ability to make instruction culturally relevant, manage a classroom with students who had experienced trauma due to Katrina and pre-Katrina impoverishment, and serve as role models for the majority black student population (Thompson, 2011; Buras, 2012).

Several studies have pointed out the ways in which a mismatch in the race of teachers and students affects the rate at which students are subject to exclusionary school discipline. Nationally, black students are rated as less disruptive and are suspended less often when they are rated by black teachers (Downey and Pribesh, 2004; Dee, 2005; Wright, 2015). In an evaluation of North Carolina teacher-student relationships, Lindsay and Hart (2017) found that a match to a same-race-gender teacher led to a two percentage point decrease, on average, in how often black male and female students experienced exclusionary discipline.

There are several factors that could explain the differential rating of black students by black and white teachers. Barrett, McEachin, Mills and Valant (2017) conceptualize race disciplinary disparities as coming from i) actual differences in the behaviors of black and white students, or ii) differences in the way schools

[20] Researchers have also found that accountability encourages teachers to focus attention on students at the margin of being proficient (Neal and Schanzenbach, 2010). Given this, other students who receive lower levels of attention may be more likely to act out and engage in misbehaviors, as documented by Holbein and Ladd (2017).

interpret behaviors as infractions and/or assign punishments. According to the first explanation, black students who have non-black teachers could misbehave in the classroom because they perceive cultural differences between themselves and their teachers or do not identify with their teachers (Monroe, 2005; Villegas and Irvine, 2010), respond to perceived pressures to "act white" (Fordham and Ogbu, 1986; Fryer, 2006), or react to teachers' lower expectations of them (Gershenson, Holt and Papageorge, 2016).

The second explanation for the racial disparities in discipline is more related to implicit biases and perceptions. Teachers determine whether the severity and frequency of student misbehavior merit an office referral. If teachers are subconsciously inclined to be more lenient toward same-race students (Gregory and Mosely, 2004; Gregory, Skiba and Noguera, 2010), demographic matches could matter for student disciplinary outcomes. Indeed, recent evidence demonstrates that educators' implicit bias may differ depending on race, even among young children. In a study about potential biases of preschool teachers, Gilliam, Maupin, Reyes, Accavitti and Shic (2016) asked educators to read a standardized vignette of a preschooler with challenging behavior and rate its severity. When educators were provided with students' family background information, educators decreased the severity ratings when teacher and child race matched and increased the ratings when race did not match.[21]

In summary, the literature suggests several mechanisms through which the school reforms could have affected school discipline rates. From heightened competition to the spread of No-Excuses discipline models, the school reforms could have altered student's behavior and/or incentivized schools to expel students. The following chapters provide empirical evidence on the net effect of the entire package of reforms on school discipline.

3. Research Design

[21] A third plausible mechanism is that schools serving black and low-income students tend to use more strict disciplinary practices. However, this mechanism is less likely to operate because the race and income composition of New Orleans students did not change significantly after Katrina.

The objective of this paper is to measure the effect of the New Orleans school reforms on school discipline, as measured by students' expulsions and out-of-school suspensions. To accomplish this goal, it constructs a synthetic control group of students, with pre-reform discipline rates that resemble that of New Orleans students.

3.1. The Synthetic Control Method

The Synthetic Control Group (SCG) method is designed to estimate the effects of an intervention that has been given to only one treatment group (Abadie et al., 2010). In the case of the New Orleans school reforms, this requires creating a synthetic control group for New Orleans that best approximates: i) the counterfactual discipline rates in New Orleans in the absence of the post-Katrina school reforms, and ii) the pre-treatment evolution of other characteristics that may be related to school discipline rates.

Following Abadie et al. (2010), I create a synthetic control group for New Orleans students, based on students in other school districts. The SCG method picks the weighted combination of all other Louisiana school districts that minimizes the mean squared prediction errors of the outcome variables and a set of observable characteristics in the pre-intervention period. This set of characteristics, or predictors of discipline rates, consists of lagged values of the discipline rates (in years 2002, 2004 and 2005), and the average number of students in each school district.

I define X_1 as a vector of observable characteristics in New Orleans before the school reforms were introduced and X_0 as the matrix of these characteristics in other school districts. The SCG method first chooses a vector V that weights characteristics according to their predictive power on the outcome. Then, the SCG method chooses a vector of weights, W, that minimizes the expression outlined in equation (1). This expression represents the distance between the pre-treatment characteristics of New Orleans and those of other school districts.

$$\|X_1 - X_0 W\| = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)} \qquad (1)$$

The synthetic control estimator of the effect of the school reforms is given by $Y_1 - Y_o W^*$, which is the weighted difference in post-reform discipline outcomes between New Orleans and its synthetic control.

With the SCG framework, it is not possible to use traditional statistical inference approaches, because of the small-sample nature of the data. However, it is possible to conduct falsification exercises or "placebo studies". These placebo studies are based on the premise that we should not observe estimated effects of similar or even greater magnitudes in districts where the school reforms did not take place.

Following Abadie, Diamond and Hainmueller (2015) and McCelland and Gault (2017), I artificially assign treatment status to each school district in the donor pool (which I also refer to as placebo districts), conduct synthetic control group analyses using each placebo treatment, and obtain differences between each district and its synthetic control (which I refer to as the estimated effect). I then compare New Orleans with the placebo districts, based on how large these estimated effects are relative to the average pre-reform differences.

To do this, I calculate the pre-reform Root Mean Squared Prediction Error (RMSPE), which measures the magnitude of the gap in discipline outcomes between each school district and its synthetic counterpart *before* the reforms took place. The RMSPE is defined in equation 2, where Y_{1dt} measures the discipline outcome of district d and $\sum w_{jd}^* Y_{jt}$ measures that of its corresponding synthetic control.

$$RMSPE_d = \sqrt{\frac{1}{T}\sum_{t<2005}\left(Y_{dt} - \sum_{j\neq d} w_j^* Y_{jt}\right)^2} \quad (2)$$

Then, I calculate the effect size $R_{dt} = \frac{Y_{dt} - \sum_{j\neq d} w_j^* Y_{jt}}{RMSPE_d}$ or the ratio of the difference between the district and its synthetic control in year t and the pre-reform $RMSPE_d$. This ratio accounts for the fact that a large post-reform difference between the district and its synthetic control is not indicative of a large effect if the pre-reform

$RMSPE_d$ is also large.[22] This is because a large $RMSPE_d$ means that the synthetic control is not a close reproduction of discipline outcomes prior to the reforms. I compare the size of the effect of the reforms in New Orleans to that of other placebo districts. If there is an actual effect of the school reforms on a given year, then only a handful of placebo districts would have effect sizes R_{dt} larger than that of New Orleans.[23]

3.2. Threats to Identification

The SCG creates a control group whose discipline rates resemble that of New Orleans in the pre-treatment period. The identifying assumption is that, absent the school reforms, discipline rates would have also been similar during the post-treatment period, both in levels and in trends. There are five general threats to this assumption that serve as alternative potential causes for a change in school discipline (Harris & Larsen, 2016).

First, the population of the city might have changed (The Data Center, 2014; Vigdor, 2008). In the process of rebuilding the city, city leaders decided to shut down and eventually replaced most of the major public housing projects. For this and other reasons, low-income residents may have not returned. Since low-income students tend to be disciplined at higher rates, this by itself could have improved discipline in schools.[24]

---

[22] Performing statistical inference based on R_{dt} is equivalent to performing statistical inference year by year, based on the ratio of RMSPE post-reform to RMSPE pre-reform, which is what Abadie, Diamond and Hainmueller (2015) recommend. I use R_{dt}, as opposed to the ratio of RMSPE, because I am interested in assessing how the effects of the reform change over time.

[23] Because the sample size of this study is small, as it is typically in SCG method studies, I do not perform inference based on the percent of placebo districts with effect sizes larger than that of New Orleans. This would require establishing thresholds above which a result is considered *true*, which can be confused with standard statistical inference, and therefore be misleading. Instead, I follow Billmeier and Nannicini (2013) and Shores, Candelaria and Kabourek (2019), and document instances where several of the placebo districts display effect sizes larger than New Orleans.

[24] Controlling for observable characteristics such as students' eligibility for free/reduced price lunch would partially overcome this threat. However, this variable is not a perfect measure of family income, and it also does not capture other unobserved family characteristics that shape students' behavior in school.

Second, when Louisiana families evacuated, they generally placed their children in schools near their temporary residences. There is evidence that New Orleans evacuees experienced larger gains in school quality in these "interim schools" relative to non-New Orleans evacuees (Sacerdote, 2012). These higher quality schools could also have had discipline policies and practices that improved student behavior. If these gains did not fade out, then some of the later changes in discipline might reflect the discipline policies of these interim schools rather than the New Orleans reforms.

Third, the accountability instituted by NCLB a few years prior to Katrina could have increased out-of-school suspensions and expulsions, regardless of the school reforms. As shown by prior research (Figlio, 2006; Jacob, 2005; Koretz, 2009), accountability can induce some schools to manipulate high-stakes measures or reallocate resources in detriment of lower-stakes outcomes. Schools might also change their student population by removing distracting and/or low performing students. NCLB could have induced such strategic behavior in pre-Katrina New Orleans schools. The school district might have been disproportionately affected by this because of the large share of low-performing schools and the large proportion of minority students.[25]

Fourth, Hurricane Katrina was one the worst disasters in American history. It forced the evacuation of entire populations and created persistent trauma and anxiety for evacuees (DeSalvo et al. 2007; Elliot and Pais, 2006; Weems et al. 2010). Some of these psychological effects were driven by evacuees' poor labor and housing outcomes post-hurricane (Elliot and Pais, 2006).[26] These poor post-Katrina labor outcomes did not seem to reflect poor pre-Katrina labor outcomes. Instead, they reflected the severity of damage of property, as the poor labor outcomes were most observed among those who had lived in heavily flooded areas that got severe

[25] NCLB exerted strong pressure on low-performing schools and established minimum standards for the performance of minority students.

[26] The ability to cope with stress and recover from the disaster was strongly related with labor and housing outcomes post-Katrina. Adults who were unemployed and who had not returned to their original homes one month after the hurricane were more likely to experience Post-Traumatic Stress Disorder (Elliot and Pais, 2006).

housing damage (Groen and Polivka, 2008). While most of the psychological evidence pertains to adults, there is also evidence of trauma and disruption among children more than two years after the Hurricanes (Brown et al., 2011). Adults' stress and anxiety, together with children's trauma and disruption, could have induced students to misbehave.

Finally, the number of reported incidents is a function of both the number of incidents that occur and the probability that the incident will be reported. Ideally, we could separate the two, but this is not possible in general or with the data available in the present study. The data comes from reports of discipline incidents that schools send to the Louisiana Department of Education, and the accuracy and consistency of the data depends on how each individual school tracks its discipline incidents. If the probability of reporting a discipline incident changed with the school reforms, then the estimated effect would be biased.

3.3. Addressing Threats: The Pooled Sample

The first approach to address some of the threats explained above consists of using pooled cross sections of student cohorts who were in the same grades pre- and post-hurricane (e.g. comparing discipline incidents for the 2005 cohort of 5th-12th graders with the 2012 cohort of 5th-12th graders). I call this the pooled sample. I aggregate the information of these students, creating an artificial school district of students in 5th-12th grades.

Creating a synthetic control group based on the discipline outcomes of pre-Katrina cohorts assumes that the unobserved factors affecting discipline are the same among pre-reform and post-reform cohorts. Using pre-reform discipline outcomes allows me to control for any pre-Katrina strategic behavior induced by NCLB. This law was introduced a few years before Katrina and could have increased discipline rates even in the absence of the city-wide school reforms (see Section 3.1). NCLB places particular pressure on low-performing schools, and these schools were prevalent in New Orleans before Katrina. To the extent that pre-Katrina districts with high discipline rates were also low-performing districts, constructing

a synthetic control district based on pre-Katrina discipline outcomes would overcome this threat.[27]

The pooled cross section sample also has the advantage of being a large sample that includes all students who were in New Orleans (and its synthetic control group) pre- or post-Katrina. This sample allows estimating coefficients with precision and drawing conclusions that are generalizable to the entire student population. However, using this sample implies relying on observable demographic information to account for population change. The second sample, which follows in the next section, assesses this issue.

To isolate the effect of the trauma and disruption created by the hurricane, the ideal approach would be restricting the donor group (i.e., those districts from which the SCG draws possible controls) to the hurricane-affected districts. However, performing this restriction under the SCG framework is suboptimal, because only a handful of school districts were affected by the hurricanes, and some of them are significantly smaller than New Orleans. This means that there are only a few comparable hurricane-affected districts to construct the synthetic control.[28] Nevertheless, I use this restriction as a robustness check and include the results in the Appendix and describe them in the next section.

To gauge the role of reporting effects in discipline incidents, I estimate the results separately for discipline incidents where the probability of bias is theoretically lower. The separation assumes that more specific incidents are reported more consistently over time and more severe incidents are reported more accurately. The literature on victimization supports this assumption, as victims tend to report severe offenses at higher rates than non-severe offenses (Bachman, 1998; Birbeck et al., 1993; Goudriaan et al., 2004; Hart and Rennison, 2003; Kilpatrick, et al., 1987; Lizotte 1985; Skogan 1976, 1984). Anecdotal conversations with school

[27] Non-disciplined students in Louisiana perform 0.2 standard deviations *above* the mean, while disciplined students, perform between 0.2 and 0.4 standard deviations *below* the mean (Barrett, McEachin, Mills and Valant, 2017; Table 1).

[28] There would be only three large enough hurricane-affected districts to include in the donor group. These are Jefferson, Calcasieu and St. Tammany parishes. The other four hurricane-affected districts had student populations in 2005 that were less than a fifth of that of New Orleans. These smaller parishes are Cameron, Plaquemines, St. Bernard and Vermilion.

leaders and educators also support this assumption. Any correlation between treatment and reporting probabilities should be smaller in these cases.

I first separate suspensions for specific incidents from those for non-specific incidents. Suspensions for incidents that are specific or well-defined (such as leaving the classroom without permission) should be reported more consistently over time, when compared to suspensions for offenses that are non-specific or ambiguously defined (such as willful disobedience). I further divide suspensions for specific incidents between serious and non-serious. Suspensions for serious offenses, such as fighting in school, should be reported more accurately than those for non-serious offenses, such as leaving the classroom without.

3.4. Addressing Threats: The Panel Sample of Returnees

Using the pooled sample as described in the previous section only addresses some of the concerns. The change in population and the interim schools where New Orleans students enrolled could still bias the results. To address this, a second approach consists of using a panel of students who evacuated because of Katrina and then returned to their pre-hurricane district for at least one year post-hurricane (e.g. comparing a students' change in discipline incidents between 2005, when she was a 3rd grader, and 2010, when she was a 8th grader).

A panel analysis allows studying a fixed group of individuals and thereby accounts for unobserved differences, among which are the changes in population caused by the hurricane. It also allows controlling for the influences of interim schools on students' behavior, as all returnee students enrolled temporarily in other schools and then came back to their original school districts. This sample, however, has an important disadvantage. It represents only evacuees who returned to their original school district, which is a small, non-random subsample of the original population. This limits statistical power and generalizability.

As I show below, the combination of results using the pooled sample and the panel sample provides a consistent pattern of results for expulsion and out-of-school suspension rates.

4. Data

The objective of this study is to estimate the effect of the New Orleans school reforms on exclusionary discipline rates. To accomplish this, I use student-level longitudinal data from the Louisiana Department of Education (LDOE). The administrative data includes all Louisiana public school students from 2001 through 2015, and includes their demographic characteristics, grade level and the schools where they enrolled (I call this the "enrollment data").

The enrollment data is complemented with LDOE data at the discipline incident level. The discipline data includes each instance of discipline reported by schools, identifying which student was involved, the type of offense, type of sanction (expulsion or suspension), place of sanction (in-school or out-of-school) and the sanction's start and end date. In order to merge this discipline data with the enrollment data, discipline incidents are aggregated at the student-year level. If a student was sanctioned in more than one school within a school year, all discipline incidents are assigned to the school that sanctioned her the most (as measured by the total number of discipline incidents within the school year).[29] The main dependent variables are: i) an indicator for whether the student was expelled during the school year; and ii) an indicator for whether the student was ever suspended out-of-school during the school year.

I omit in-school suspensions as a discipline outcome because New Orleans schools are not required to report them by law. The average rate of in-school suspensions in New Orleans is many times lower than in the rest of the state (Figure A.1. in the Appendix). Because of this difference, the total rate of suspensions (in- and out-of-school) appears to be significantly lower in the city, as compared with the rest of the state. Although it is possible that New Orleans schools refrain from using in-school suspensions because of capacity and personnel constraints,

[29] An ideal scenario would be to construct a database at the student-school-year level, where discipline incidents are assigned to the school where they happened. This database would be merged with enrollment data at student-school-year level. However, I am limited by data availability, since the enrollment data for years 2001-2012 only reports the last school attended by a student within the school year. Therefore, I cannot observe all the schools attended by non-disciplined students during this period (i.e. students who are not found in the discipline database).

conversations with former teachers and school leaders revealed that in-school suspensions are indeed used, but are not reported.

In some specifications, suspension variables are further disaggregated by type of offense: specific and non-specific. By specific offenses, I mean offenses that are clear and well defined, such as leaving school without permission or getting involved in a fight. Non-specific offenses, on the other hand, are less serious and ambiguously defined, such as willful disobedience or disrespect for the authority. This disaggregation has the purpose of discerning changes in schools' deliberate use of exclusionary discipline (likely reflected in non-specific offenses) from changes in students' behavior (likely reflected in specific offenses).

Specific offenses can be severe (categorized as serious) or non-severe (categorized as non-serious). Serious offenses correspond to habits that can injure others or offenses that need to be reported to law enforcement authorities, while non-serious offenses correspond to less severe habits, such as leaving the classroom without permission or being habitually tardy. This further disaggregation is done to check if the results are driven by measurement error in the out-of-school suspensions variable. Serious offenses are likely to be reported with more accuracy and, therefore, are more reliable.

Table 1 presents the disaggregation of out-of-school suspensions, together with their frequency by type of offense. The most frequent offenses leading to out-of-school suspensions are those classified as non-specific (61%, Column 2), which include willful disobedience (23%), violating rules habitually (13%) and disrespecting the authority (12%). Specific offenses, on the other hand, are equally divided between serious (Column 3) and non-serious offenses (Column 4), each representing roughly 20% of overall out-of-school suspensions. The most frequent type of serious offense corresponds to fighting in school (13%), followed by displaying habits that can injure others (3%).

5. Sample and the Synthetic Control Group

This paper implements a synthetic control group method based on two samples: the pooled sample of student cohorts and the panel sample of returnees. Each

sample assesses different threats to identification, as explained in Section 3. The following sections describe each sample and the corresponding synthetic control group that serves as a comparison for New Orleans. Unless otherwise noted, I treat all New Orleans schools as belonging to one school district, regardless of the entity in charge of running or supervising the school (OPSB or RSD).

5.1. Pooled Sample

The pooled sample is drawn from the population of Louisiana public school students enrolled in grades 5-12 from 2001 through 2015, excluding year 2003.[30] Table 2 presents statistics of this population, separating New Orleans from the rest of Louisiana. The population consists of approximately 1.4 million students, out of which 8% were enrolled in New Orleans schools. Table 2 reveals that New Orleans' demographic composition is very different from the rest of Louisiana. Ninety percent of the students enrolled in New Orleans from 2001 through 2015 were African-American (compared to 40% in the rest of the state) and 70% were eligible for free-reduced price lunch (compared to 56%).

The pooled sample focuses on middle and high school students because Louisiana schools start using exclusionary discipline as early as grade 5. As Figure A.2 in the Appendix reveals, in 2005 more than 20% of New Orleans' 5th grade students were suspended out-of-school (Panel A). New Orleans' expulsion rate is lower in grade 5 (does not exceed 1%), but increases significantly in later grades, reaching 2% in grades 8 and 9 (Panel B).

As explained in section 3.3, the comparison group consists of a synthetic control group of students in grades 5 through 12. To build this comparison group, I first eliminate from the donor pool small school districts, defined as those with less than one fifth of the New Orleans student population in 2005.[31] I aggregate their information, creating artificial school districts of students in 5th-12th grades. After this process, there are left 16 school districts in the donor pool. For each outcome, I

[30] The year 2003 is excluded from the sample because New Orleans displayed an atypical increase in expulsion rates, coinciding with the enactment of NCLB. Appendix B explains this with more detail.
[31] Those eliminated had fewer than 7,000 students in grades 5-12 in 2005. New Orleans had 36,908 such students in 2005.

construct a separate synthetic control group of students, based on the pre-intervention discipline rates in 2002, 2004 and 2005 and the number of students in the district.

The list of weights for each predictor and outcome combination is presented in in Panel A of Table 3. The pre-reform outcomes receive the highest weight, while the district size has little weight, regardless of the outcome considered.[32] The weights are then used to choose the school districts that contribute to the synthetic control group. The list of weights for each school district and outcome combination is presented in Panel A of Table 4. For all outcomes in the pooled sample, there are 3 to 4 school districts that contribute to the synthetic control (e.g. have positive weights). Of these contributing districts, there are 1 to 2 with large weights. For instance, Caddo and Terrebone Parishes contribute 45% and 35% to the synthetic control group when the outcome is expulsions (Column 1), while Jefferson Parish contributes 89% to the synthetic control when the outcome is out-of-school suspensions (Column 2). While these are large weights, the major contributing district always changes with the outcome, which alleviates the concern that the results are driven by one specific district. Nonetheless, in section 7 I perform robustness checks that eliminate Caddo Parish from the donor pool and obtain similar results.

Panel A in Table 6 compares the pre-reform characteristics of New Orleans to those of the synthetic control group, and also to those of a student-weighted average of the remaining school districts in Louisiana that are medium to large sized. Columns 1-3 present statistics specific to the expulsions outcome, while columns 4-6 present statistics specific to the out-of-school suspensions outcome. Overall, the results in Table 6 suggest that the synthetic control group provides a much better comparison for New Orleans than the average of other large school districts. The synthetic New Orleans is very similar to the actual New Orleans in terms of pre-reform discipline rates, both for expulsions and out-of-school suspensions. The synthetic control, however, cannot match closely the district size, since it has far

[32] The small weight attached to district size might be due to the scale of the variable, which is artificial.

fewer students than New Orleans.[33] This is consistent with the small weight that the SCG gives to the district size predictor, relative to the lagged outcomes (Table 3).

Panel A in Figure 1 shows that before the school reforms, New Orleans and the rest of Louisiana's large school districts experienced different paths in their expulsion rates. However, the synthetic control group can accurately reproduce the pre-reform expulsion rate path for New Orleans, as shown in Panel B. Out-of-school suspensions had similar paths in New Orleans and the rest of Louisiana before Katrina, but differed in their levels (Figure 3, Panel A). The synthetic control group also has out-of-school suspension levels similar to New Orleans (Figure 3, Panel B).

One potential disadvantage of the pooled sample is that the population in New Orleans might have changed with the hurricane. In particular, low-income families might have not returned to New Orleans after the storm (see Section 3.2 for more details). However, Table 7 shows otherwise. The percent of students who belong to minority groups or who are eligible for free or reduced-price lunch is similar between those students who returned and those who did not. This is true for both New Orleans and other hurricane-affected districts. This piece of evidence supports the idea that the public student population in hurricane-affected areas (including New Orleans) did not change significantly after Katrina. Harris and Larsen (2016) provide other evidence reinforcing this point. Using Census data for households with students in the public-school system, they show that after the storm the percent of children in poverty decreased only 1 percentage point, relative to other hurricane-affected districts.[34]

Nonetheless, I perform additional analyses using only the sample of students who returned to their original school district after the hurricane, as explained in the following section.

5.2. Panel Sample of Returnees

[33] New Orleans has around 39,000 students in grades 5-12, while the synthetic control group specific to expulsion rates has 18,000 students. In the case of out-of-school suspensions, the synthetic control group's size is closer to New Orleans, with 28,000 students (Table 6, Panel A).

[34] The authors also show that the average level of education increased in New Orleans (relative to other hurricane-affected districts), while the median household income dropped. The former might bias results down, while the latter might bias results up. Harris and Larsen (2015) also show that the effects of these two changes on student achievement are likely to compensate each other.

The panel sample consists of the cohort of students who were in grade 7 in 2005, and returned to their original school district after Katrina for at least one year. These students would have graduated (on-time) in 2010. This specific cohort of students was chosen for two reasons. First, it allows testing the effect of the school reforms on high school discipline, given that these students would have been enrolled in high school in 2009, when expulsion rates reached a peak (Figure 1). Second, these students are old enough to have had a record of discipline incidents pre-Katrina, given that discipline incidents in Louisiana start as early as grade 5 (Figure A.2 in the Appendix). This is necessary because lagged discipline rates are used as a predictor of discipline outcomes to construct the synthetic control group.

The comparison group consists of a synthetic control group of returnees. To build this comparison group, I drop students who never returned to their pre-Katrina school districts and aggregate the information of those who did return. I also eliminate from the sample small school districts, defined as those with less than one fifth of New Orleans' returnee population in 2005.[35] I aggregate returnees' information, creating artificial school districts of 2005 7th grade returnees. After this process, there are 17 schools districts remaining in the donor pool. Similar to the pooled sample, I construct a separate synthetic control group of returnees for each outcome, based on pre-intervention discipline outcomes in years 2002, 2004 and 2005, and the number of students.

The weights for each predictor and outcome combination are qualitatively similar to that of the pooled sample (Table 5), in that the pre-reform outcomes receive the largest weights. The school district weights are also comparable to the pooled sample, since a small number of districts contribute to the synthetic control.[36] However, in contrast to the pooled sample, there is one major contributing district that is common across outcomes. The weight of the school district in Caddo Parish is large in the synthetic control group corresponding to total

[35] Those eliminated had less than 500 7th grade returnees in 2005. New Orleans had around 2,600 such students in 2005.

[36] An exception happens with the outcome out-of-school suspensions for specific offenses, where there are 16 school districts contributing to the synthetic control. However, among these school districts, there is one (Caddo Parish) that receives a larger weight.

out-of-school suspensions, out-of-school suspensions for specific offenses, and for non-serious offenses (Columns 2, 4 and 6 in Table 5). This could constitute a concern, because Caddo Parish could be driving the results. However, I show in the robustness checks section that omitting Caddo from the donor group yields the same conclusions.

Panel B in Table 6 shows that the synthetic control group provides a much better comparison for New Orleans returnees than the average of other school districts. This is consistent with the findings for the pooled sample (Panel A). The synthetic New Orleans is very similar to the actual New Orleans in terms of pre-reform discipline rates. This point is further confirmed in Panel B of Figures 4 and 5, showing that the synthetic control resembles expulsion and out-of-school suspension rates in New Orleans before Katrina.

6. Results

Figure 6 illustrates the effect of the school reforms on the expulsion rate. In the figure, the bold line represents the difference in expulsion rates between New Orleans and its synthetic control, while the gray lines represent the same difference for the placebo districts. As mentioned in section 3.1, the placebo exercise consists of applying the SCG to other districts comparable to New Orleans. If the New Orleans results arise because of chance, then several placebo districts would display larger effect sizes. The figures include placebo districts that belong to the donor pool, with a size comparable to that of New Orleans (i.e. with at least one fifth of the New Orleans student population in 2005).[37]

Figure 6 presents the results for expulsions both for the pooled sample (Panel A) and the panel sample (Panel B). The bold line in the two panels reveal that New Orleans expulsion rate increased temporarily with the school reforms. By 2009, the percent of students expelled had increased between 1.5 (pooled sample) and 2.8

[37] Panel A in these figures (pooled sample) includes placebo districts with at least 7,000 students in grades 5-12 in 2005. Panel B (panel sample) includes placebo districts with at least 500 students in the 7th grade cohort in 2005, who returned to their original school district at least for one year after Katrina.

(panel sample) percentage points, representing a 1- to 3-fold increase, relative to the New Orleans average in 2005 (0.011).

In order to assess if the large increase in 2009 is the result of chance, Column 1 in Tables 8 and 10 present the annual estimated effects (i.e. the difference between New Orleans and its synthetic control), while Column 1 in Tables 9 and 11 present the number of placebo districts with annual effect sizes above that of New Orleans. The effect size is measured as the absolute value of the estimated effect in a given year divided by the pre-reform RMSPE.[38]

All placebo districts in the panel sample (Table 11 and Panel B in Figure 6) display effect sizes in 2009 far below that of New Orleans. The results for the pooled sample are consistent (Table 9 and Panel A in Figure 6), and only 2 out of the 16 placebo districts have effect sizes in 2009 above that of New Orleans. The magnitude of the effects on the expulsion rate is not implausible. The effect sizes are consistent with the complaints raised by community stakeholders and the 2010 federal lawsuit filed against the LDOE because of discipline practices in New Orleans (see Section 2.2 for further details).

The estimates reveal a large increase in the expulsion rate in 2009 both in the panel and pooled samples. The large effects on expulsion rates are driven by schools that had not yet been converted into charter schools and were temporarily run directly by RSD. This is evidenced in Figure 7, which shows the treatment effects by education sector. For these calculations, I estimate the difference between expulsion rates of students enrolled in that sector and that of New Orleans' synthetic control. As the figure shows, the RSD-direct sector experienced a 4-fold increase in expulsions, while other sectors (including the charter sector) displayed slight decreases. These results reveal that the increase in expulsions was not caused by new charter schools that followed strict discipline regimes. Instead, the changes brought by the reforms lead to an increase in exclusionary expulsion policies among schools that were transitioning to charter schools, with little enforcement of the RSD code of conduct that governed them. This phenomenon could also have been

[38] The RMSPE measures the lack of fit between the path of discipline outcomes for any particular school district and its synthetic control group. See Section 3.1 for more details.

reinforced by the RSD, which encouraged schools to adopt No-Excuses policies, including their discipline practices, and hired a security firm and a former New Orleans Police Department commander to ensure safety at RSD schools (see Section 2.2 for more details).

The increase in the expulsion rate quickly disappeared. As both Panels in Figures 6 and 7 show, the peak in expulsions started reversing immediately in 2010, coinciding with the federal lawsuit against LDOE alleging unlawful discipline practices against special education students (described in section 2.2). By 2011, the increase of expulsions in the RSD direct-run sector had completely disappeared and by 2012 expulsions had decreased 1.5 percentage points. This effect size is found in only 3/16 placebo districts (Column 1 in Table 9). It is not clear if this sharp decrease was the result of the pressure exerted by community stakeholders and the federal lawsuit, or the stabilization of the school reforms. However, it is possible to rule out alternative explanations, such as the creation of the centralized expulsion system. By the time the centralized system was implemented at the beginning of the 2012-2013 academic year, the increase in expulsions had already disappeared.

The effects of the reforms on out-of-school suspensions are qualitatively similar to expulsions, but much smaller in magnitude and also found in several placebo districts. Figure 8 shows that the percent of students suspended out-of-school initially increased between 17% (pooled sample) and 26% (panel sample) in 2009.[39] However, none of these effects seem to be specific to New Orleans, as several placebo districts display similar effect sizes. As with expulsions, the coefficients varied by sector. Figure 9 shows that by 2009 the percent of students suspended out-of-school increased between 52% (pooled sample) and 62% (panel sample) in schools directly run by RSD. [40] While the estimates are not different from those found in the placebo districts, the results on out-of-school suspensions corroborate the findings regarding expulsions. The increase in discipline rates was

[39] In the pooled sample, the increase was 4.2 percentage points in 2009, relative to New Orleans average of 23.8% in 2005. In the panel sample, the increase was 7.5 percentage points, relative to 28.4% average in 2005 (Column 2 in Tables 8 and 10).

[40] In both the pooled and the panel samples, the increase was close to 15 percentage points in 2009. This corresponds to a 52% increase in the pooled sample (relative to an average of 23.8% in 2005) and a 62% increase in the panel sample (relative to a 28.4% average in 2005).

driven by the RSD direct-run schools that were in transition to become charter schools.

I estimate the effects on out-of-school suspensions separately by type of infraction to test for the potential influence of reporting effects. Suspensions for offenses that are specifically defined and serious are more likely to be reported equally across schools and over the years. According to Figure 10, the percent of students suspended for non-specific offenses did not change significantly with the reforms (e.g. was not much different than that of the placebo districts), regardless of the sample used. In contrast, the percent of students suspended for specific offenses displayed sustained increases (Panel A in Figure 11). The increases ranged between 38% and 54% post-reform and were found in only 2-3 of the 16 placebo districts (Column 4 in Tables 8 and 9).[41] However, in the panel sample, the effect sizes are smaller (between 17% and 22%) and they are also found in several (6/16) placebo districts (Column 4 in Tables 10 and 11).[42]

Suspensions for specific offenses are further divided into serious and non-serious offenses.[43] Figures 12 and 13 present the results of this further disaggregation and show large increases in the suspension rate for serious offenses, as compared to those for non-serious offenses. Serious offenses increased between 61% (panel sample) and 90% (pooled sample),[44] and effects of these sizes can be found only in 1 of the 16 placebo districts (Column 5 in Tables 9 and 11).[45] Suspensions for non-serious offenses, in contrast, seemed to only increase between 15% (panel sample) and 25% (pooled sample), and it is difficult to distinguish if

[41] By 2009 the percent of students suspended for specific offenses had increased 4.7 percentage points, reaching 6.5 percentage points in 2014. The average in 2005 was 11.9% (Column 4 in Tables 7 and 8).

[42] In the panel sample, specific offenses increased 3.8 percentage points in 2009. The average in 2006 was 16.9%.

[43] Serious offenses correspond to fights and habits that can injure others. Non-serious offenses include infractions such as leaving the classroom without authorization or being habitually tardy/absent. Section 4 explains with more detail the classification of offenses.

[44] By 2009, suspensions for serious offenses had increased 1.4 percentage points in both the pooled and panel samples, relative to a 2005 average of 1.5% in the pooled sample and 2.1% in the panel sample (Column 5 in Tables 8 and 10).

[45] Column 6 in Tables A.1 and A.2 shows that only 6% of the placebo districts (or 1 out of 16) displayed effect sizes larger than that of New Orleans.

these effects are obtained by chance.[46] According to the pooled sample, only 3-4 out of the 16 placebo districts have larger effect sizes, which is contradicted by the results in the panel sample, where 9/17 placebo districts have effects sizes above that of New Orleans (Column 6 in Tables 9 and 11).

In summary, reporting errors do not seem to drive the increase in suspensions. When I focus on suspensions for serious offenses, which tend to be recorded more accurately, the effects are even higher. This is true both for the pooled and the panel sample.

7. Robustness Checks

It is clear that the effects of the reforms may be confounded with the devastation and disruption created by Katrina. Evidence shows that New Orleans was disproportionately affected by the hurricanes, compared to the rest of Louisiana. The hurricane could have created trauma among New Orleans students, causing them to misbehave and be expelled from schools. However, if this were the case, expulsions would have increased in the immediate aftermath of Katrina; and as trauma receded, expulsions would have decreased slowly over time. The results suggest otherwise. The post-Katrina expulsion rate peaked in 2009, three years after the reforms initiated, and dropped immediately after. Nonetheless, I perform additional analyses to account for the hurricane's impact by including only other hurricane-affected districts in the comparison group. The results of these analyses are presented in Figures A.3 and A.4 in the Appendix. Although performing inference is challenging, because the donor pool consists of only three hurricane-affected districts that are large enough to be included, the results are consistent with the findings detailed above. They show that expulsions increased sharply in 2009 and decreased rapidly thereafter, regardless of the sample used (Panel A in both Figures).

[46] Suspensions for non-serious offenses increased 2.8 (2.4) percentage points in the pooled (panel) sample in 2009, relative to a 2005 average of 10% (15%).

Given that the initial increase in expulsions was driven by changes in RSD direct-run schools, I looked more closely at whether the differences in infractions were due to the types of students attending RSD direct run schools versus other schools. To do this, I analyze the sample of students who were in 4th grade in 2005 for the first time and who had returned to New Orleans by 2009.[47] I classify them according to their education sector in 2009, and look at their achievement and discipline outcomes both pre and post-Katrina. I find that, before Katrina, students in RSD direct-run schools had higher suspension rates and performed worse in Math and ELA exams, as compared to their counterparts in other sectors (Figure A.5). However, these same characteristics did not change significantly in RSD-direct run schools in 2009. There is no evidence of a spike in the level of disadvantage of students in RSD direct-run schools in that year, that would have explained the large increases in exclusionary discipline rates.

Nonetheless, I perform additional analyses to account for the differences in the type of students attending RSD direct-run schools. I treat each education sector (RSD direct, RSD charter, etc.) as a separate school district and compare each of them to a separate synthetic control that resembles the pre-Katrina characteristics of its students.[48] To do this, I first classify panel sample students (2005 7th grade cohort of returnees) according to their education sector in 2009 and create artificial pre-reform education sectors based on the pre-reform outcomes of these students. Then, I use the synthetic control method to find a control group for each sector. The set of predicting variables includes pre-Katrina out-of-school suspension rates and 6th and 7th grade Math and ELA test scores, to account for the differences in the type of students in each sector. Figure A.6 in the Appendix shows similar results. Expulsion rates increase substantially in RSD-direct schools in 2009 when

[47] This sample is different than the analysis panel sample, which consisted of returnees who were in the 7th grade in 2005. I choose a different sample because it allows examining achievement patterns both pre- and post-Katrina. The post-Katrina achievement data is not available for the analysis panel sample, when students were in high school and took tests less often.

[48] Students sorted into the education sectors as a result of the school reforms. This makes the definition of education sectors endogenous. This is not a concern, however, because the objective of this exercise is to find control groups that resemble the type of students that sort in RSD-direct schools, as opposed to calculate heterogeneous effects by education sector,

compared to the RSD-direct synthetic control, while schools in other sectors experience no significant change in their expulsion rates (i.e. their increase in expulsions rate is no larger than the increase in other placebo districts). This increase in the RSD direct-run sector expulsions reverses immediately after, similar to the findings in Section 6.[49]

Given that one single district (Caddo Parish) was the major contributor to the synthetic control in the panel sample (Table 5), I checked if the temporary increase in expulsions was driven by this single control district. I repeat the estimations dropping Caddo Parish from the donor pool and find similar results. Panel A in Figure A.7. confirms that the expulsion rate increased significantly in 2009 and dropped immediately after.

8. Discussion and Conclusions

This paper investigates the effect of the New Orleans school reforms on school's exclusionary discipline rates. It provides evidence of temporary large increases in expulsions in 2009, ranging from 1 to 3 times that of its synthetic control. The increase is consistent with complaints raised by community stakeholders about discipline rates in some schools and the lawsuit filed alleging high discipline rates of special education students in New Orleans schools.

The evidence suggests that the effects were caused by schools temporarily run by RSD. These schools did not belong to the charter sector, meaning that No-Excuses charter schools had no role in the increase. Instead, the increase in expulsions could be the adverse result of transitioning to an entirely new system. During the transition, there was apparently little oversight over discipline policies of schools directly run by RSD. These schools expected to be converted into charter schools, which meant a change in the school administration, teachers and, in some cases, the student body. The expectation of change may have given these schools little incentive to avoid using exclusionary discipline when faced with students'

[49] Panels B-D in Figure A.6. also show that The rate of suspensions out-of-school also increase more in RSD direct-run schools, and that increase is sustained through 2010.

misbehavior. This adverse reaction might have been exacerbated by the RSD leadership, which encouraged schools to adopt No Excuses policies, with high academic expectations and strict approaches for school discipline and safety.

The large increase in the expulsion rate was temporary. Expulsions started to decrease in 2010, coinciding with the lawsuit that alleged high discipline rates of special education students in New Orleans schools. Although the stabilization of the school reforms could have also played a role, expulsions would have decreased slowly, had that been the case. After the increase in expulsions, RSD implemented the New Orleans centralized expulsion system in 2012, which established common criteria for expelling a student and a hearing process that enforces the application of those criteria to expulsions. This system, however, had no role in decreasing expulsions back to its pre-Katrina levels, since it was established after the effect had already disappeared.

This paper also provides some qualitatively similar but less convincing evidence of increases in the out-of-school suspension rate. The rate of suspensions for specific offenses (i.e. those that are well defined, such as getting involved in fights or being habitually tardy) increased between 40 and 60% in the years after the reform. These increases, however, are not found when using the panel sample or when focusing on suspensions for non-specific offenses. The inconsistent results are partially explained by schools reporting suspensions with different accuracy levels. This creates measurement error and could attenuate the effects. To assess this, I focus only on suspensions for serious offenses, which tend to be reported with more accuracy, and find consistent increases in the suspension rate ranging from 60 to 90% in both the pooled and panel samples. The results regarding suspensions for non-serious offenses are smaller in magnitude and might be created by chance, but are consistent in terms of their direction.

A potential limitation of this paper is that the effects of the reforms may be confounded with the devastation and disruption created by Katrina. The evidence shows that New Orleans was disproportionately affected by the hurricane compared to the rest of Louisiana. The hurricane could have created trauma among New Orleans students, causing them to misbehave and be expelled from schools.

However, if this had been the case, expulsions would have increased in the immediate aftermath of Katrina and, as trauma receded, would have decreased slowly over time. The evidence in this study suggests otherwise. The post-Katrina expulsion rate peaked in 2009, three years after the reforms initiated, and dropped immediately after. Nonetheless, I address this limitation by performing robustness checks that include only hurricane-affected districts in the donor pool. The results hold when applying this restriction, but are generally noisy.

Because New Orleans represents the most intense test-based and market-based accountability reform ever implemented in the United States, the results from this study draw valuable lessons for other large urban school districts going through intense changes in their structures. A charter-based school system brought large effects on the academic achievement of its students (Harris and Larsen, 2016; Bross, Harris and Liu, 2016), but the drastic change also created large temporary increases in exclusionary discipline rates. The increase in discipline rates can be ameliorated, however, with public pressure and legal challenges.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *Journal of the American Statistical Association, 105*(490), 493–505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science, 59*(2), 495–510.

Abdulkadiro, A., Angrist, J., Dynarski, S., Kane, T., & Pathak, P. (2011). Accountability and flexibility in public schools: evidence from Boston's charters and pilots. *The Quarterly Journal of Economics*, *126*, 699-748.

American Psychological Association (APA). (2008). Are zero tolerance policies effective in the schools? An evidentiary review and recommendations. *American Psychologist*, *63*(9), 852-862.

Angrist, J.D., Pathak, P.A., & Walters, C.R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics*, *5*(4), 1-27.

Bachman, Ronet (1998) "The Factors Related to Rape Reporting Behavior and Arrest." *Criminal Justice and Behavior* 25:8-29.

Barret, N., & Harris, D.N. (2015). *Significant changes in the New Orleans teacher workforce.* New Orleans, LA: Education Research Alliance for New Orleans, Tulane University.

Barret, N., McEachin, A., Mills, J., & Valant, J. (2016). *Race and discipline in the U.S. south: The origins of the black-white gaps in exclusionary discipline in Louisiana* (Unpublished manuscript).

Billmeir and Nannicini (2013). "Assessing Economic Liberalization Episodes: A Synthetic Control Approach". *Review of Economics and Statistics* 95(3), 983-1001.

Birbeck, Christopher, Luis G. Gabaldon, and Gary LaFree (1993) "The Decision to Call the Police: A Comparative Study of the United States and Venezuela." *International Criminal Justice Review* 3(1):25-43.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In J. Heckman & E. Leamer (Eds.), *Handbook of econometrics* (Vol. 5, pp. 3705-3843). New York: Elsevier.

Boyd, A., Maranto, R., and Rose, C. (2014) The Softer Side of 'No Excuses': A view of KIPP schools in action. *Education Next* 14(1), 48-52.

Bross, W., Harris, D. N., & Liu, L. (2016). *The effects of performance-based school closure and charter takeover on student performance.* New Orleans, LA: Education Research Alliance for New Orleans, Tulane University.

Brown, T. H., Mellman, T. A., Alfano, C.A., & Weems, D.F. (2011). Sleep fears, sleep disturbance, and PTSD symptoms in minority youth exposed to Hurricane Katrina. *Journal of Traumatic Stress, 24*(5), 575-580.

Buras, K. L. (2012). It's all about the dollars. In Watkins, W.H. (Ed.), *The assault on public education* (166). New York, NY: Teachers College.

Carr, S. (2008, April 17). Recovery district taking tips from KIPP schools – Methods stress structured learning college expectations. *The Times-Picayune*. Retrieved from http://infoweb.newsbank.com/resources/doc/nb/news/1201BDFBD32D9778?p=AWNB

Carr, S. (2010). New Orleans charter schools work to sustain teachers' energy, results. *NOLA.com*. Retrieved from http://www.nola.com/education/index.ssf/2010/03/new_teachers_working_long_hard.html

Carter, S. (2000). No Excuses: Lessons from 21 High-Performing, High-Poverty Schools. Heritage Foundation, Washington, DC.

Charpentier, C. (2008, December 10). Parent blasts New Orleans charter school's expulsion process. *The Times-Picayune*. Retrieved from http://www.nola.com/news/index.ssf/2008/12/parent_blasts_new_orleans_char.html

Cheng, A., Hitt, C., Kisida, B., & Mills, J. N. (2017). "No Excuses" charter schools: A meta-analysis of the experimental evidence on student achievement. *Journal of School Choice, 11*(2), 209-238.

Chiang, H. (2017). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, *93*(9), 1045-1057.

Cullen, J., & Reback, R. (2006). Tinkering toward accolades: school gaming under a performance accountability system. In T. Gronberg & D. Jansen (Eds.)*, Improving school accountability: check-ups or choice* (Vol. 14, Advances in Applied Microeconomics, pp. 1-34). Amsterdam: Elsevier Science.

Cullen, J., Jacob, B., & Levitt, S. (2006). The effect of school choice on participants: evidence from randomized lotteries. *Econometrica*, *74*(5), 1191-1230.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review*, *95*, 158-165.

Dee, T. & Jacob, B. (2011). The impact of no Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446

DeSalvo, K. B., Hyre, A. D., Ompad, D. C., Menke, A., Tynes, L. L., & Muntner, P. (2007). Symptoms of posttraumatic stress disorder in a New Orleans workforce following Hurricane Katrina. *Journal of Urban Health*, *84*(2), 142-152.

Dobbie, W., & Fryer, R. (2011). Are high quality schools enough to close the achievement gap? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics 3*, 158-187.

Dobbie, W., & Fryer, R. (2016). Charter Schools and Labor Market Outcomes. National Bureau of Economic Research Working Paper Series No. 22502.

Downey, D. B., & Pribesh, S. (2004). When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education*, *77*, 267-282.

Elliott, J. R., & Pais, J. (2006). Race, class, and Hurricane Katrina: Social differences in human responses to disaster. *Social Science Research, 35*, 295–321.

Figlio, D. (2006). Testing, Crime and Punishment. *Journal of Public Economics*, 90, 837–851.

Figlio, D. & Gletzer, L. (2006). Accountability, Ability and Disability: Gaming the System. Gronberg, T. (ed.) Advances in Microeconomics. Elsevier.

Figlio, D., & Rouse, C. (2006). Do accountability and voucher threats improve low performing schools? *Journal of Public Economics, 90*, 239–255.

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of 'acting white'." *The Urban Review, 18*(3), 176-206.

Fryer, R. (2006). Acting White: The social price paid by the best and brightest minority students. Education Next 6(1): 52-59.

Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, *52*, 209-224.

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). *Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions?.* New Haven, CT: Yale Child Study Center.

Goudriaan, H., Lynch, J. & Nieuwbeerta, P (2004). Reporting to the Police in Western Nations: A Theoretical Analysis of the Effects of Social Context. *Justice Quarterly* 21(4):933-69.

Goodman, J. T. (2013). Charter management organizations and the regulated environment: Is it worth the price? *Educational Researcher*, *42*(2), 89–96.

Greene, J. (2016). Ed Reform is Animal Farm [Blog post]. Retrieved from https://jaypgreene.com/2016/06/28/ed-reform-is-animal-farm/

Gregory, A., & Mosely, P. M. (2004). The discipline gap: Teachers' views on the over-representation of African American students in the discipline system. *Equity & Excellence in Education*, *37*, 18–30.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*, 59–68.

Groen, J. & Polivka, A. (2008). The effect of Hurricane Katrina on the labor market outcomes of evacuees. *American Economic Review: Papers and Proceedings*, *98*(2), 43–48.

Gross, B., Tuchman, S., & Yatsko, S. (2016). *Grappling with discipline in autonomous schools: new approaches from D.C. and New Orleans.* Seattle, WA: Center on Reinventing Public Education. Retrieved from http://www.crpe.org/sites/default/files/crpe-grappling-with-discipline-6.2016-final.pdf.

Harris, D. N. (2015). Good news for New Orleans: Early evidence shows reforms lifting student achievement. *Education Next, 15*(4). Retrieved from http://educationnext.org/good-news-new-orleans-evidence-reform-student-achievement/cowe

Harris, D. N., & Larsen, M. (2016). *The effects of the New Orleans post-Katrina school reforms on student academic outcomes*. New Orleans, LA: Education Research Alliance for New Orleans, Tulane University.

Harris, D. N., & Larsen, M. (2018). *What Effect did the New Orleans School Reforms Have on Student Achievement, High School Graduation, and College Outcomes?* New Orleans, LA: Education Research Alliance for New Orleans, Tulane University.

Hart, T. & Rennison, C. (2003). *Reporting Crime to the Police*, 1992-2000. Washington, DC: Bureau of Justice Statistics (NCJ195710).

Hoffman, J. V., Assaf, L. C., & Paris, S. G. (2001). High-stakes testing in reading: Today in Texas, tomorrow? *The Reading Teacher*, 482–492.

Holbein, J. B., & Ladd, H. F. (2017). Accountability pressure: regression discontinuity estimates of how No Child Left Behind influenced student behavior. *Economics of Education Review*, *58*, 55-67.

Horn, J., and Wilburn, D. (2013). The mismeasure of education. Charlotte, NC: Information Age Publishing.

Hoxby, C. (1999). The effects of school choice on curriculum and atmosphere. In S. Mayer & P. Peterson (Eds.), *Earning and learning: How schools matter* (pp. 280-314). Washington, D.C.: Brookings Institution Press.

Imberman, S. (2011). Achievement and behavior in charter schools: Drawing a more complete picture. *Review of Economics and Statistics*, *93*(2), 416-435.

Jacob, B. (2005). Accountability, incentives, and behavior: the impact of high-stakes testing in the Chicago public schools. *Journal of Public Economics, 89*, 761–796.

Jones, B. D. (2007). The unintended outcomes of high-stakes testing. *Journal of Applied School Psychology*, *23*(2), 65–86.

Kilpatrick, D., Saunders, B., Veronen, L., Best, C. & Von, J. (1987). Criminal Victimization: Lifetime Prevalence, Reporting to Police, and Psychological Impact." *Crime and Delinquency* 33(4):479-89.

Koretz, D. (2009). *Measuring up: What educational testing really tells us.* Cambridge, MA: Cambridge University Press.

Lack, B. (2009). No excuses: A critique of the Knowledge Is Power Program (KIPP) within charter schools in the USA. *The Journal for Critical Education Policy Studies*, 7(2), 127–153

Ladd, H. F. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review*, 18(1), 1–16.

Lindsay, C. & Hart, C. (2017). Exposure to same-race teachers and student disciplinary outcomes for black students in North Carolina*. Educational Evaluation and Policy Analysis*, *39*(3), 485–510.

Lizotte, A. (1985). The Uniqueness of Rape: Reporting Assault Violence to the Police. *Crime and Delinquency* 31(2):169-90.

McCelland and Gault (2017). The Synthetic Control Method as a Tool to Understand State Policy. Urban Institute Research Report.

Medler, A. (2016, August 24). Considering discipline as a student outcome could change the debate [Blog post]. Chicago, IL: NACSA. Retrieved from http://www.qualitycharters.org/2016/08/considering-discipline-student-outcome-change-debate/

Mock, B. (2010, October 6). New Orleans accused of failing disabled students. *Newsweek*. Retrieved from http://www.newsweek.com/new-orleans-accused-failing-disabled-students-73991

Monroe, C. R. (2005). Understanding the discipline gap through a cultural lens: Implications for the education of African American students. *Intercultural Education*, *16*(4), 317-330.

Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics, 92*(2), 263–283.

P.B., et al. vs. Pastorek, et al., 2:10-cv-04049 (E.D. La. 2010)

Petrilli, M. J. (2016, June 1). *Forum: Discipline practices in America's charter schools.* Retrieved from https://edexcellence.net/articles/forum-discipline-practices-in-americas-charter-schools

Public Broadcasting Service (2005, November 1). New Orleans schools before and after Katrina. *NewsHour* [Television broadcast]. Arlington, VA: Public Broadcasting Service. Retrieved from http://www.pbs.org/newshour/bb/education-july-dec05-neworleans_11-01/

Reckdahl, K. (2007, September 19). School openings a success, leader tells BESE members – Vallas gives report on RSD's goals. *The Times-Picayune*. Retrieved from http://infoweb.newsbank.com/resources/doc/nb/news/11BC4AF676A99760?p=AWNB

RSD Louisiana Recovery School District (2008). *Positive school culture: 2008-2009 student code of conduct.* Retrieved from http://www.soeanola.com/pdf/student-code-of-conduct-2008-2009.pdf

Sacerdote, B. (2012). When the saints come marching in: Effects of Katrina evacuees on schools, student performance and crime. *American Economic Journal: Applied*, *4*(1), 109-135.

Samuels, C. (2010, July 29). Suit filed against New Orleans district on behalf of students with disabilities. *Education Week*. Retrieved from

http://blogs.edweek.org/edweek/speced/2010/07/suit_filed_against_new_orleans.html?r=2011872236&qs=New+Orleans+Special+Education+Lawsuit

Shores, Candelaria and Kabourek (2019) "Spending More on the Poor? A Comprehensive Summary of State-Specific Responses to School Finance Reforms from 1990-2014" CEPA Working Paper No. 19-01. Center for Education Policy Analysis, Stanford University.

Simon, D. (2007a, June 25). School security firm defends its cost – superintendent seeks new charter deal. *The Times-Picayune.* Retrieved from http://infoweb.newsbank.com/resources/doc/nb/news/119FF2FDCAD54120?p=AWNB

Simon, D. (2007b, December 4). Eight new charters endorsed for N.O. - state board expects to vote Thursday. *The Times-Picayune*. Retrieved from http://infoweb.newsbank.com/resources/doc/nb/news/11D55B1DEB00A6D8?p=AWNB

Sims, P., & Rossmeier, V. (2015). *The state of public education in New Orleans: 10 years after hurricane Katrina.* New Orleans, LA: Cowen Institute for Public Education Initiatives at Tulane University. Retrieved from http://www.speno2015.com/images/SPENO.2015.small.single.pdf

Skogan, W. (1976). Citizen Reporting of Crime: Some National Panel Data. *Criminology* 13(4): 535-549.

Skogan, W. (1984). Reporting Crime to the Police: The Status of World Research. *Journal of Research in Crime and Delinquency* 21(2):113-37.

Steele, J., Vernez, G., Gottfried, M., & Schwam-Baird, M. (2011). *The transformation of a school system: Principal, teacher, and parent perceptions of charter and traditional schools in post-Katrina New Orleans.* Santa Monica, CA: RAND Corporation.

Sullivan, E., & Morgan, D. (2010). Pushed out: Harsh discipline in Louisiana schools denies the right to education, a focus on the recovery school district in New Orleans. *National Economic and Social Rights Initiative (NESRI) & Families and Friends of Louisiana's Incarcerated Children (FFLIC).* Retrieved from https://www.nesri.org/sites/default/files/Pushed%20Out.pdf

The Data Center. (2014). *Who lives in New Orleans and metro parishes now?.* New Orleans, LA.

Thernstrom, A., & Thernstrom, S. (2003). *No excuses: Closing the racial gap in learning.* New York, NY: Simon and Schuster.

Thompson, J. (2011, July 25). New Orleans charter schools need to respect teachers' experience, students' dignity. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/john-thompson/new-orleans-charter-schoo_b_908171.html

Vigdor, J. (2008). The economic aftermath of Hurricane Katrina. *Journal of Economic Perspectives, 22*(4), 135-154.

Villegas, A. M., & Irvine, J. J. (2010). Diversifying the teaching force: An examination of major arguments. *Urban Review*, *42*, 175–192.

Weems, C. F., Taylor, L. K., Cannon, M. F., Marino, R. C., Romano, D. M., Scott, B. G., & Triplett, V. (2010). Posttraumatic stress, context, and the lingering effects of the Hurricane Katrina disaster among ethnic minority youth. *Journal of Abnormal Child Psychology*, *38*, 49–56.

Weiher, G. & Tedin, K. (2002). Does choice lead to racially distinctive schools? Charter schools and household preferences. *Journal of Policy Analysis and Management*, *21*(1), 79-92.

Wheelock, A., Haney, W., & Bebell, D. (2000). What can student drawings tell us about high-stakes testing in Massachusetts? *The Teachers College Record.*

Withman, D. (2008). Sweating the Small Stuff: Inner-City Schools and the New Paternalism. Thomas B. Fordham Institute.

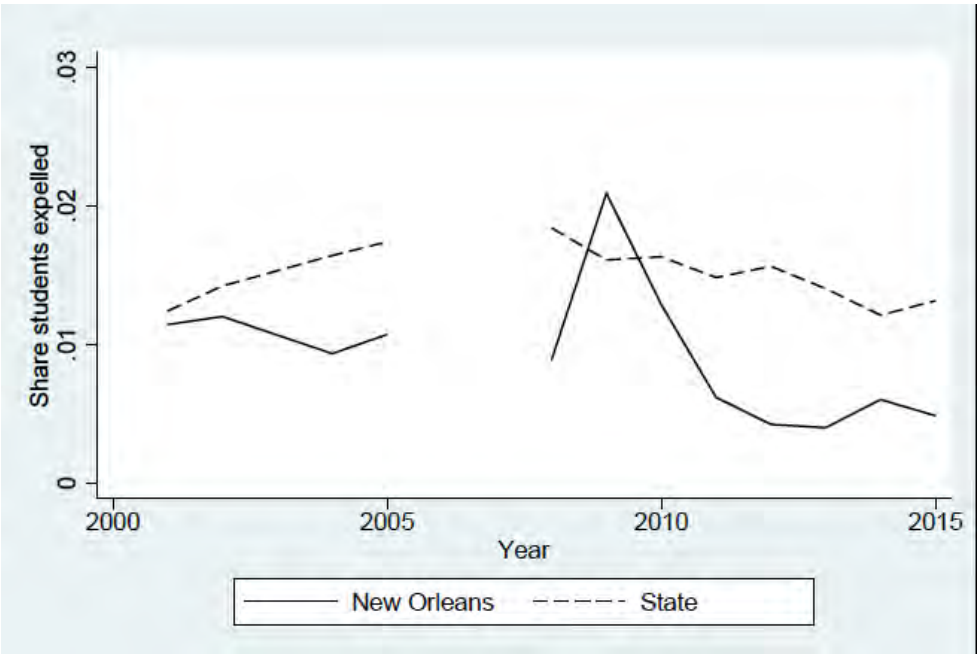
Williams, J. (2014). Harvey charter school looks to give fresh start to Woodmere families. *NOLA.com.* Retrieved from http://www.nola.com/education/index.ssf/2014/11/harvey_charter_school_looks_to_give_fresh_start_to_woodmere_families.html

Wright, A. C. (2015). Teachers' perceptions of students' disruptive behavior: The effect of racial congruence and consequences for school suspension (Unpublished manuscript). University of California, Santa Barbara.

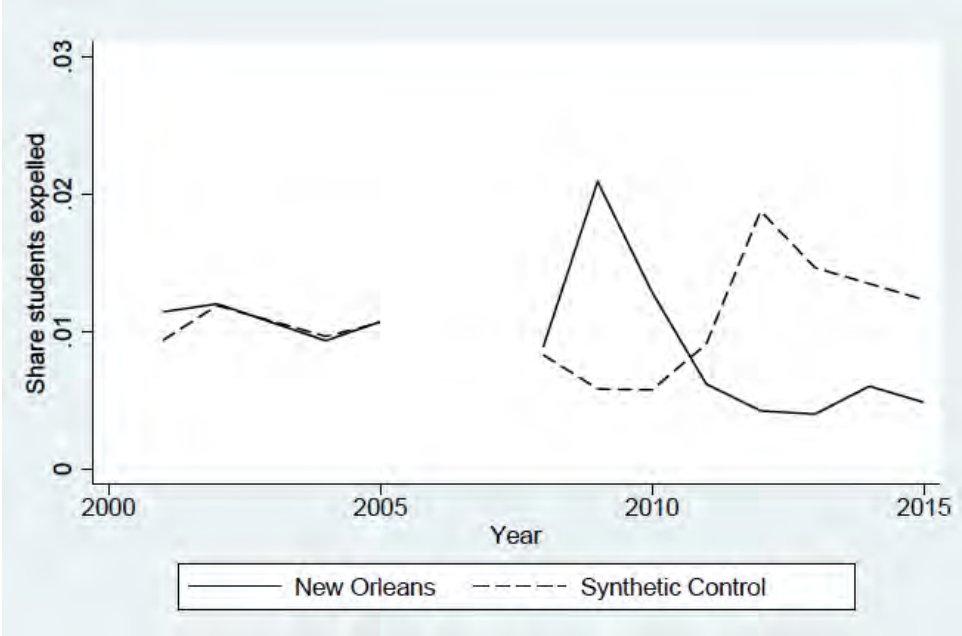
Figures and Tables

Figure 1. Pooled sample: Share of students expelled

Panel A. New Orleans and the rest of the state

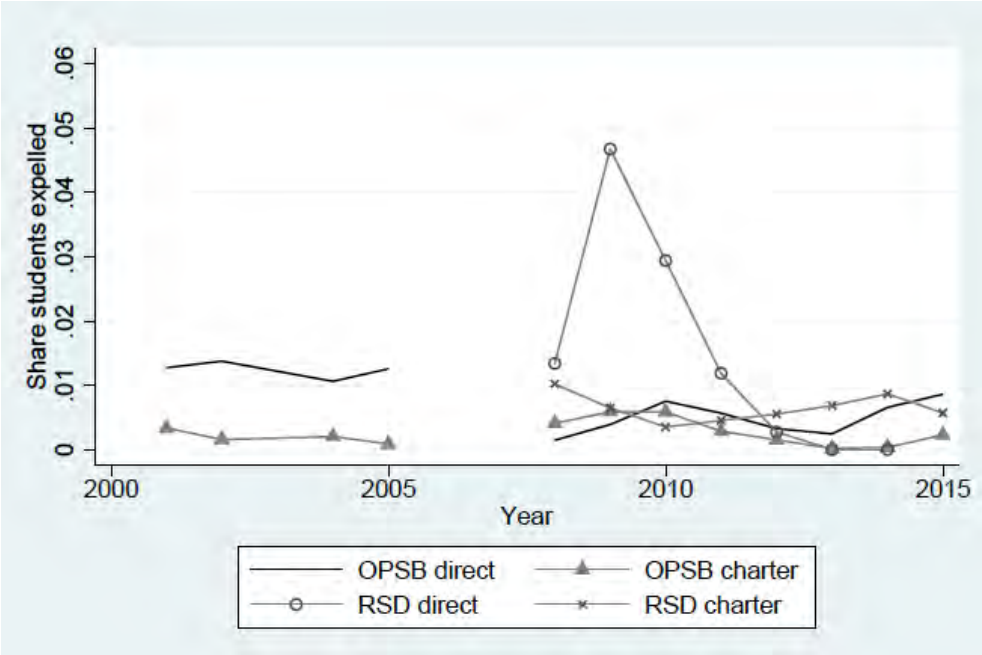


Panel B. New Orleans and the synthetic control group



Notes: The sample consists of all students in grades 5-12, enrolled in Louisiana public schools. In Panel A, State corresponds to large school districts in Louisiana (with more than 7,000 students in grades 5-12 in 2005). The 2006 and 2007 years are not included because of low quality of the data in the aftermath of Katrina.

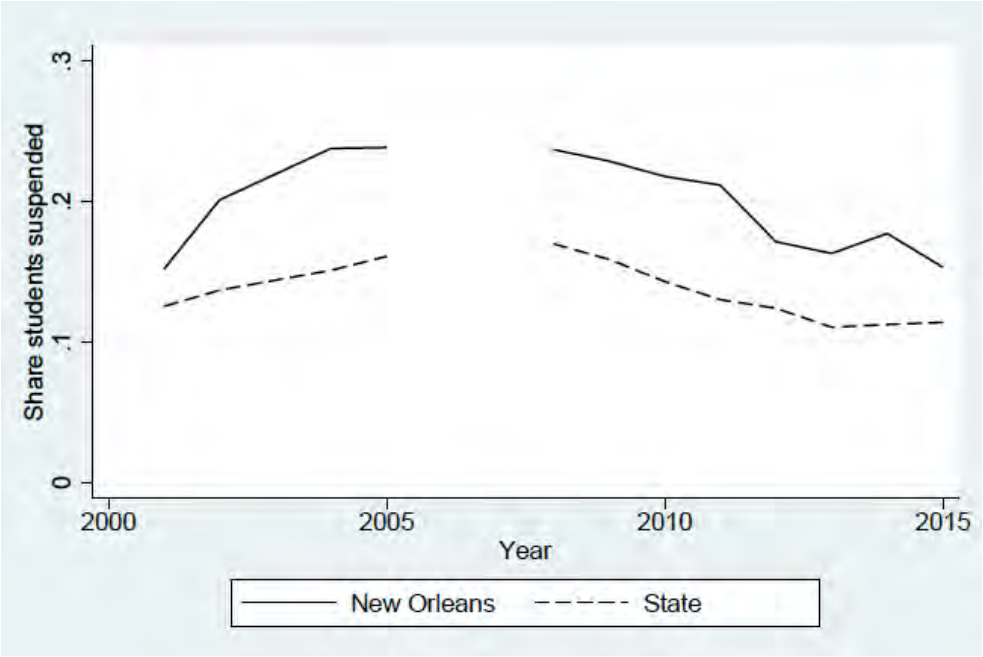
Figure 2. Pooled Sample: Share of students expelled in New Orleans, by education sector



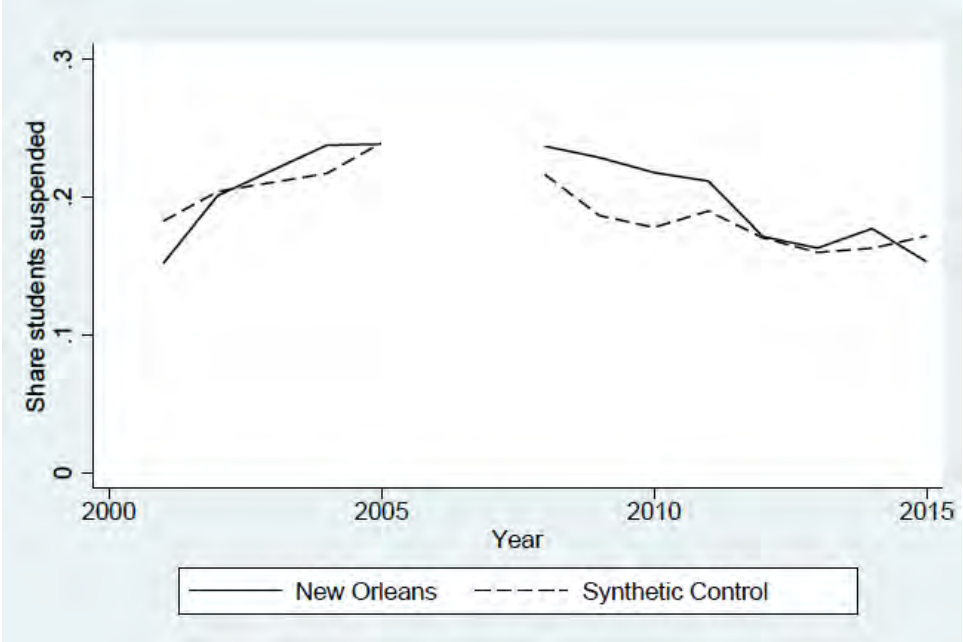
Notes: The sample consists of all students in grades 5-12 in New Orleans. RSD direct (RSD charter) denotes schools that were directly run (overseen) by the Recovery School District. OPSB direct (charter) denotes schools that were directly ran (overseen) by the Orleans Parish School Board. More details in the notes section in Figure 1.

Figure 3. Pooled Sample: Share of students suspended out-of-school

Panel A. New Orleans and Rest of the State



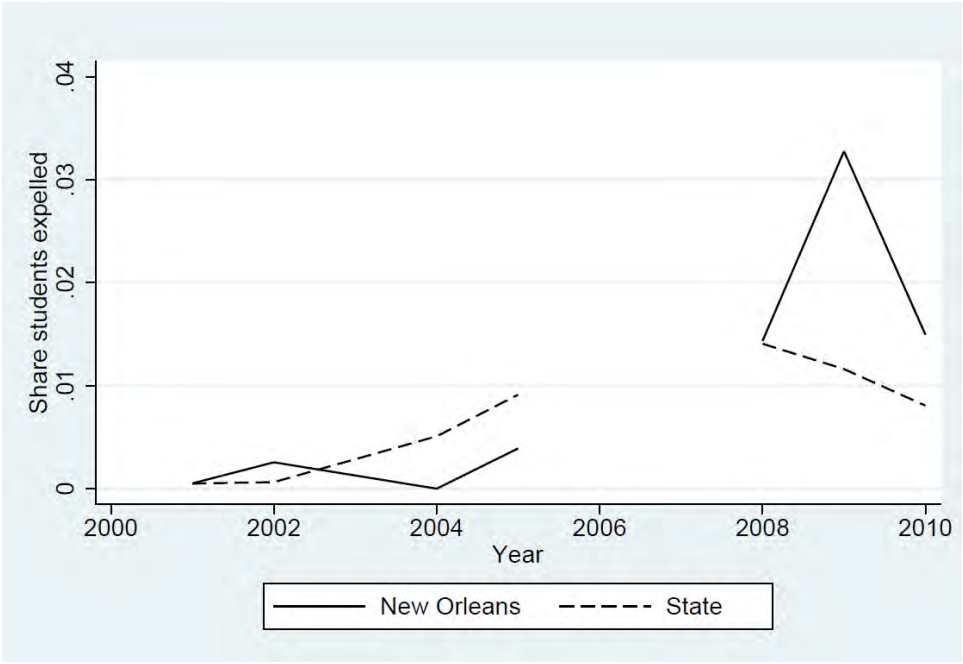
Panel B. New Orleans and the Synthetic Control Group



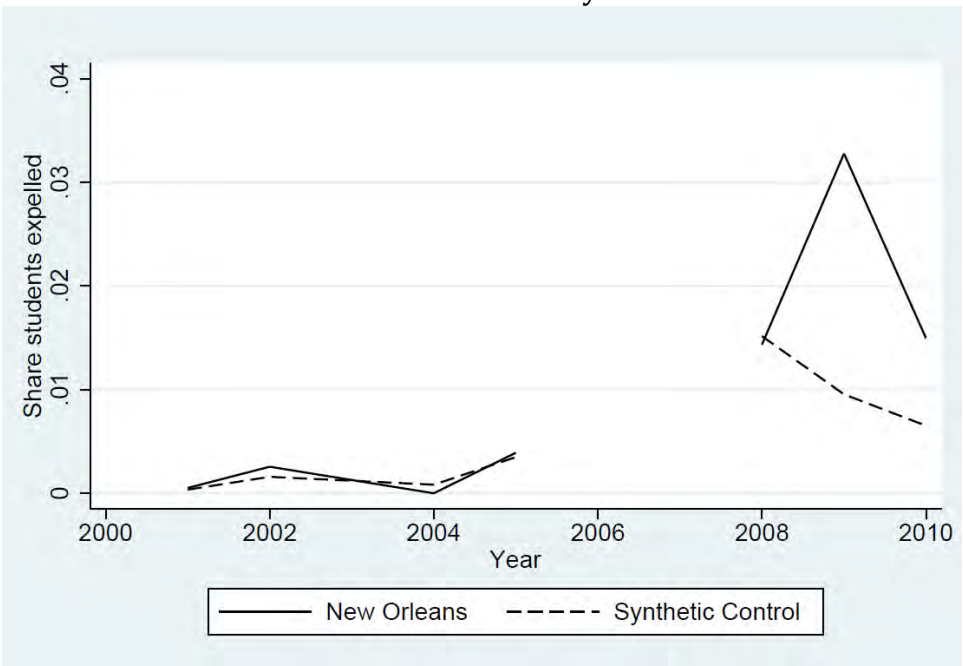
Notes: For details refer to notes in Figure 1

Figure 4. Panel sample: Share of students expelled

Panel A. New Orleans and the rest of the state



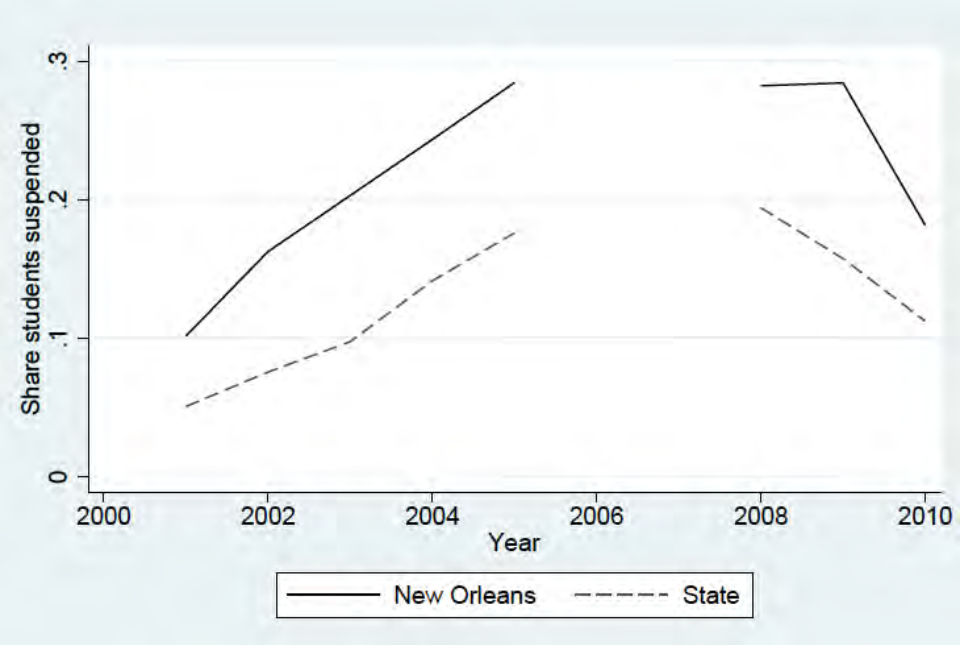
Panel B. New Orleans and the Synthetic Control



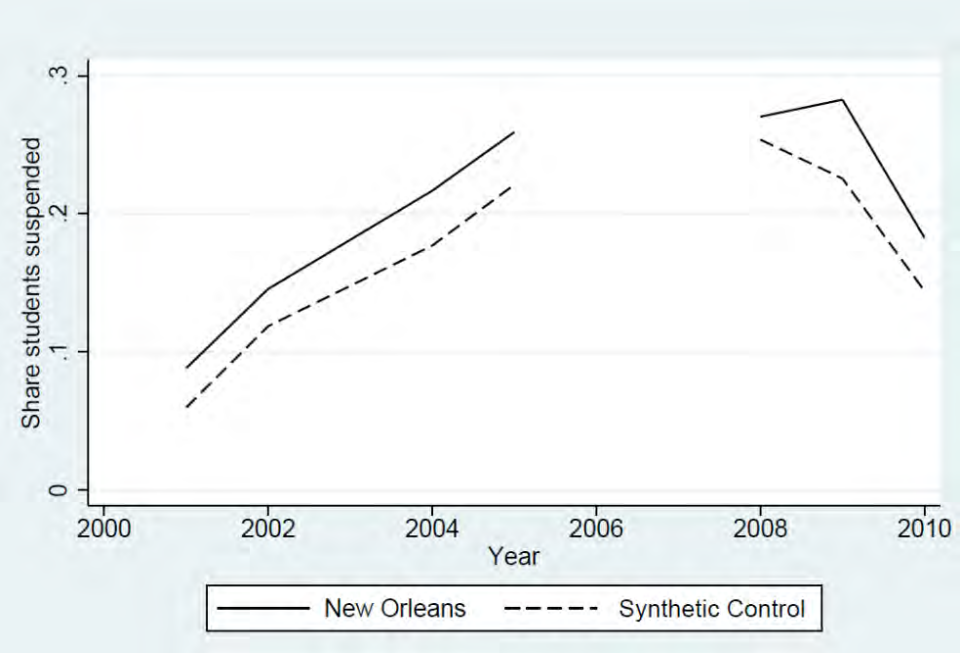
Notes: The sample consists of artificial school districts created with students who were in the 7th grade in 2005 and returned to their original school district.. The synthetic control group is created using several predictors for the outcome, such as lagged values of the outcome, school district size, average test scores and demographic characteristics.

Figure 5. Panel Sample: Share of Students Suspended Out-of-school

Panel A. New Orleans and the Rest of the State



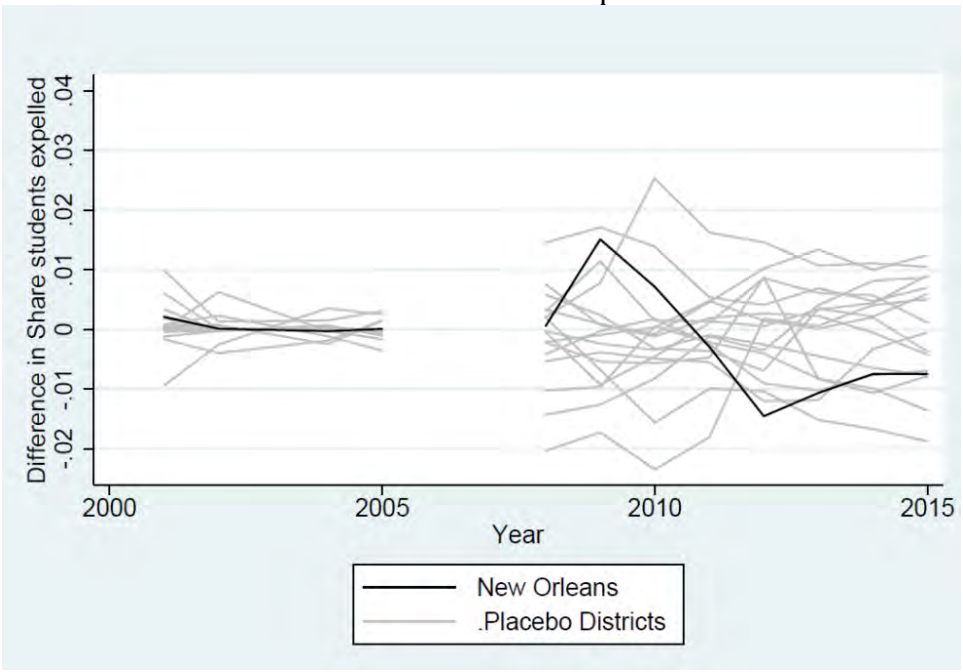
Panel B. New Orleans and the Synthetic Control



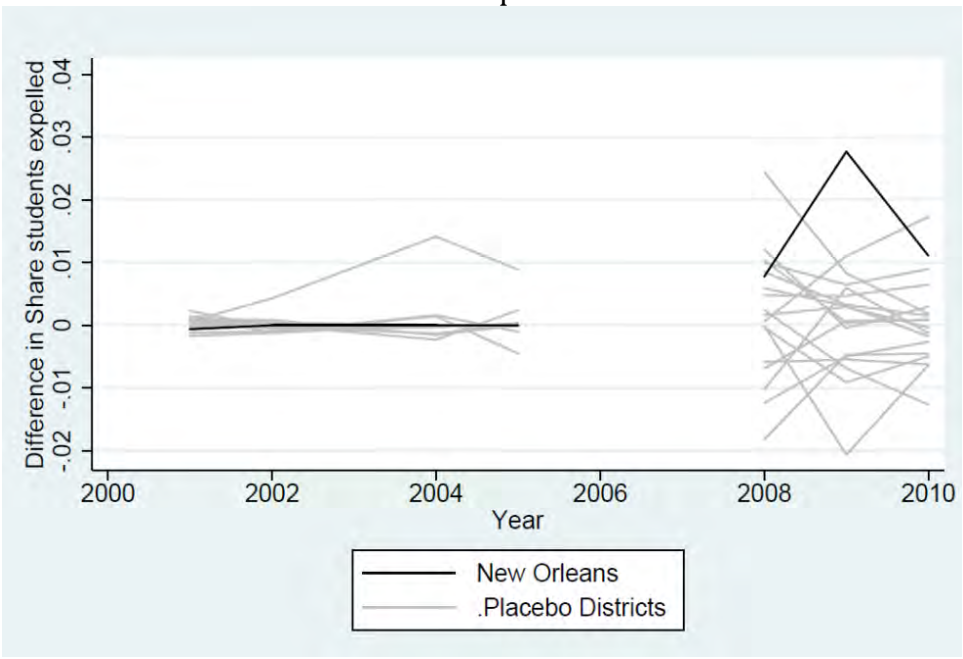
Notes: See notes in Figure 4

Figure 6. Effects on the Share of Students Expelled

Panel A. Pooled Sample



Panel B. Panel Sample of Returnees

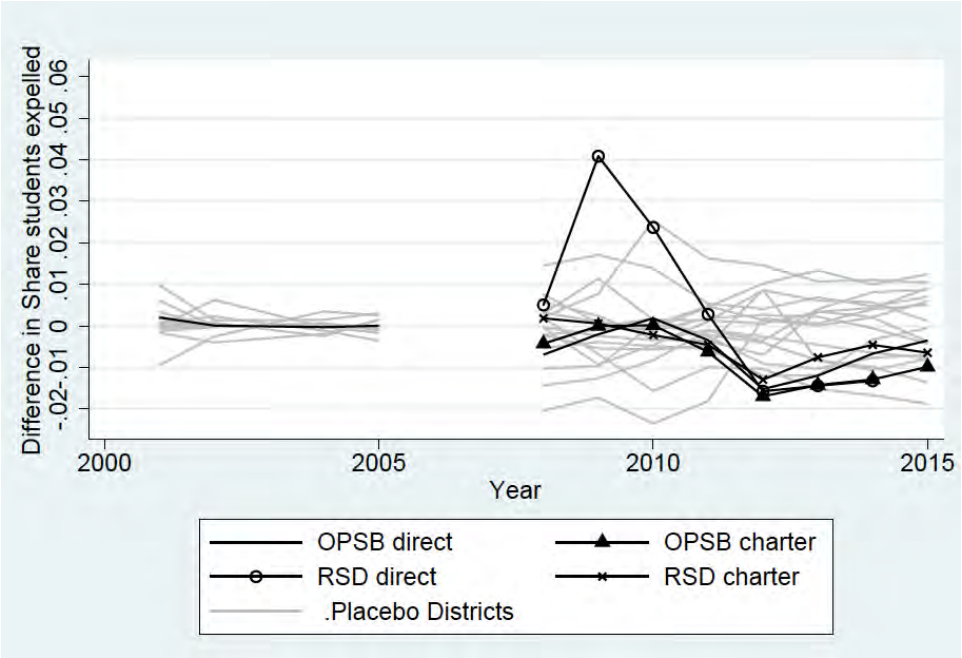


Notes: The sample in Panel A consist of all students in large school districts enrolled in grades 5-12 between 2001 through 2015. The sample in Panel B consists of large district students who were enrolled in the 7th grade district 2005 and who returned to their pre-Katrina school district at least for one year. The bold line represents the difference in expulsions rate between New Orleans and its synthetic control, while the gray lines represent the same difference for all the placebo districts. The placebo includes large school districts with at least one fifth of the student population in New Orleans in 2005.

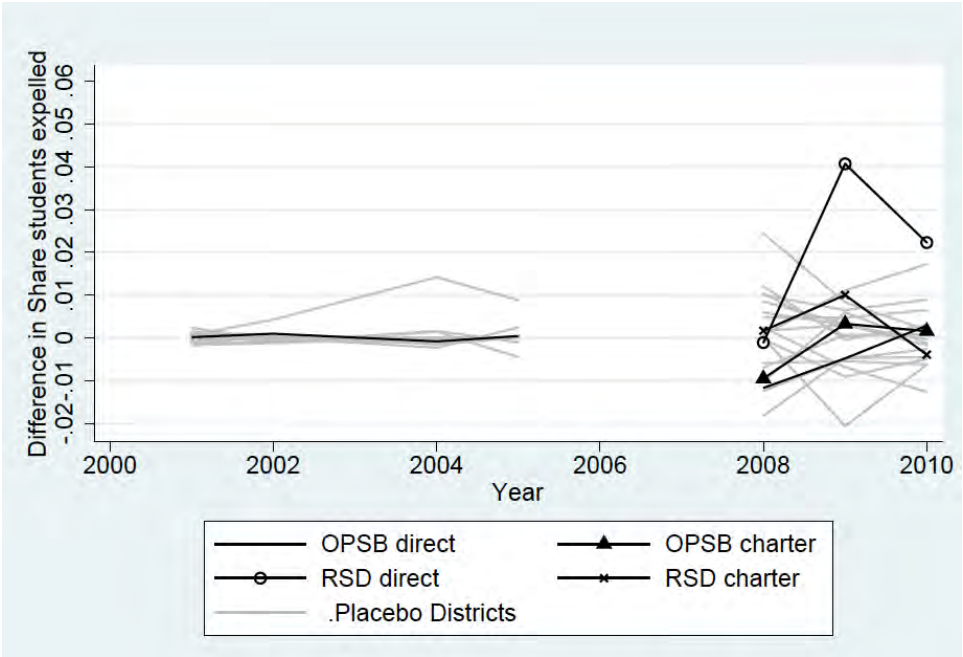
.

Figure 7 Effects on share of students expelled, by education sector

Panel A. Pooled sample



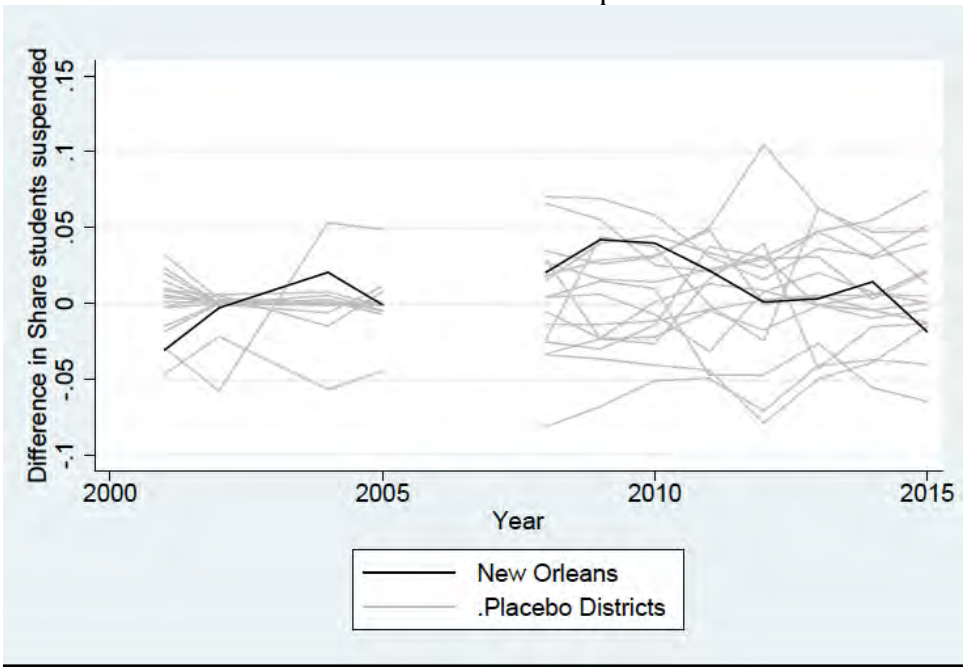
Panel B. Panel sample of returnees



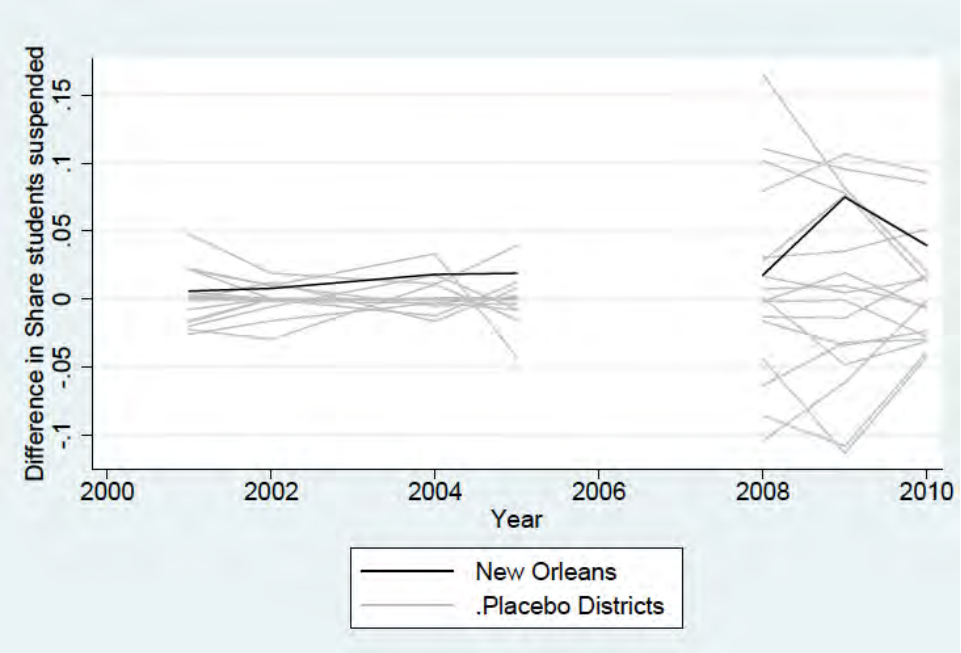
Notes: In the post-reform years, the line for each education sector represents the difference in expulsions rate between schools in that sector and New Orleans (overall) synthetic control. In the pre-reform years, the OPSB direct line represents the difference between all New Orleans schools (most of which were OPSB direct) and New Orleans synthetic control. Refer to notes in Figure 6 for more details.

Figure 8. Effects on the share of students suspended out-of-school (SOS)

Panel A. Pooled sample



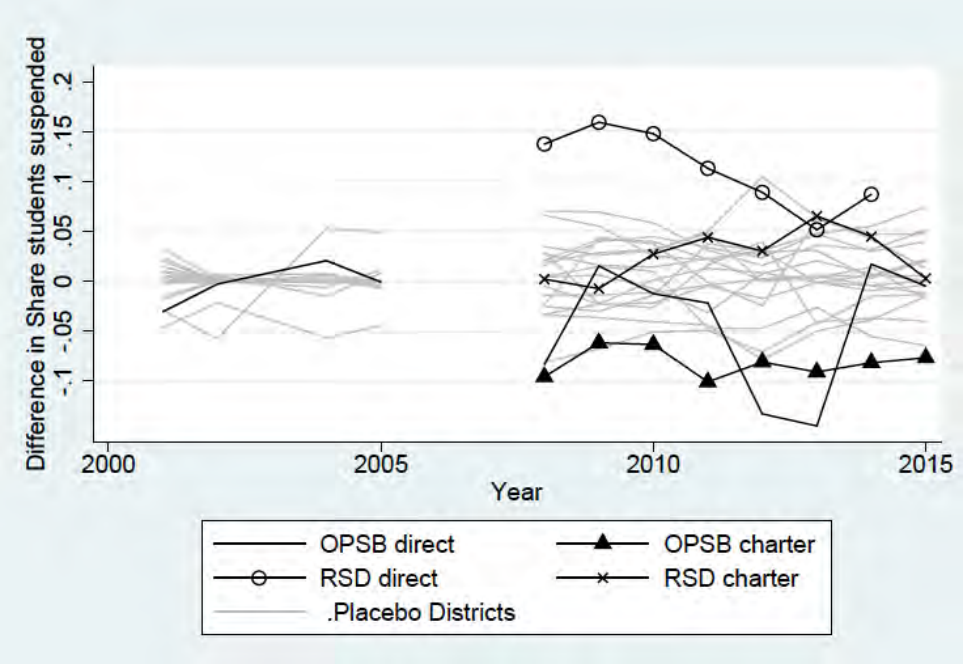
Panel B. Panel Sample



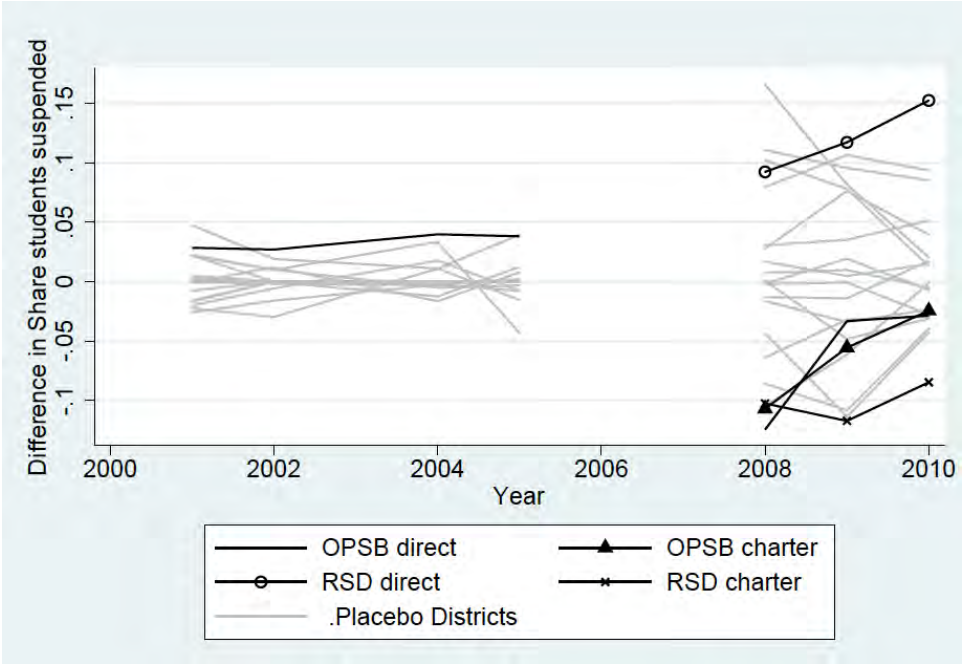
Note: Refer to notes in Figure 6 for more details

Figure 9. Effects on share of students SOS, by education sector

Panel A. Pooled sample



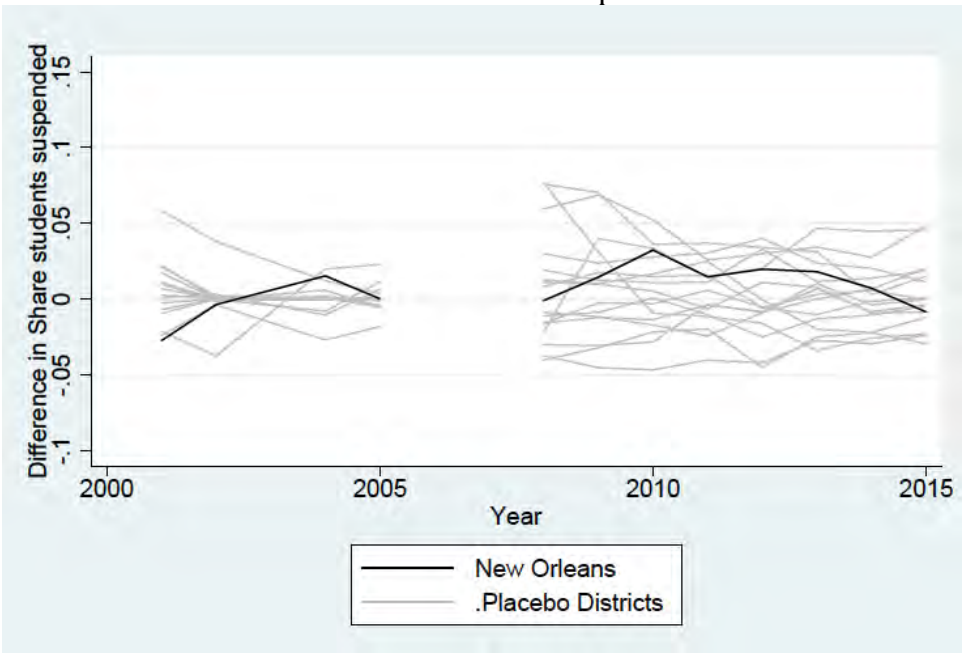
Panel B. Panel sample of returnees



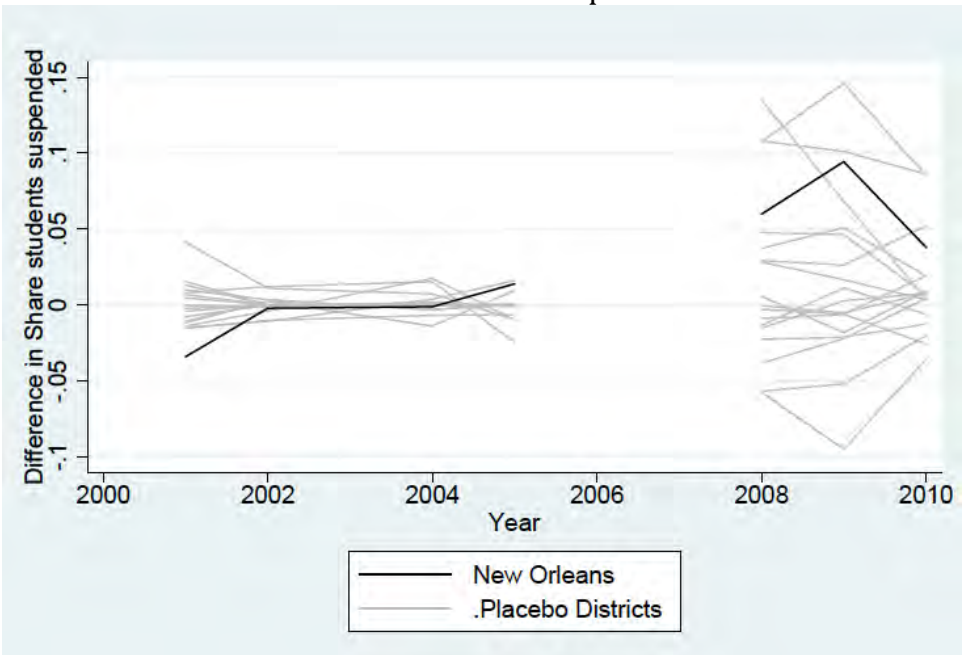
Refer to notes in Figure 7

Figure 10. Effects on SOS for non-specific offenses

Panel A. Pooled sample



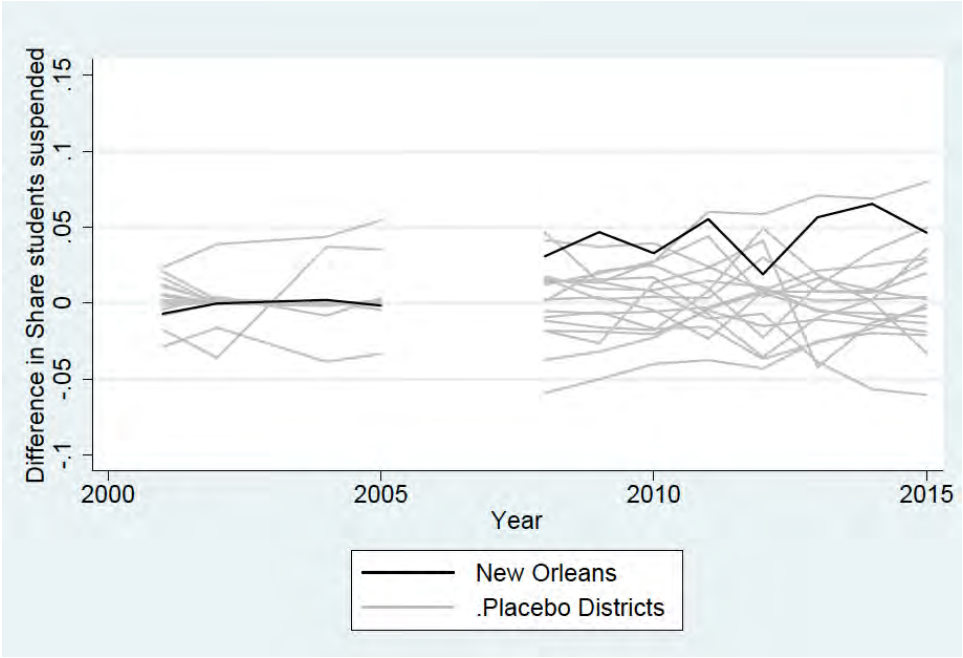
Panel B. Panel sample



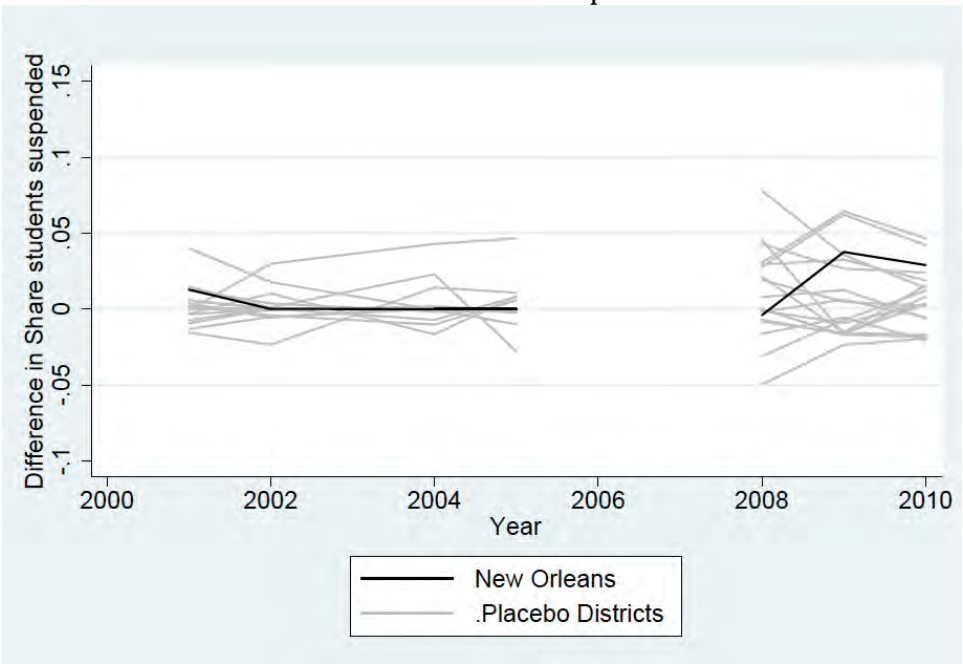
Notes: Non-specific offenses correspond to offenses that are less severe and non-specific or ambiguously defined, such as willful disobedience or disrespect for the authority. Refer to notes in Figure 6 for more details

Figure 11. Effects on SOS for specific offenses

Panel A. Pooled sample



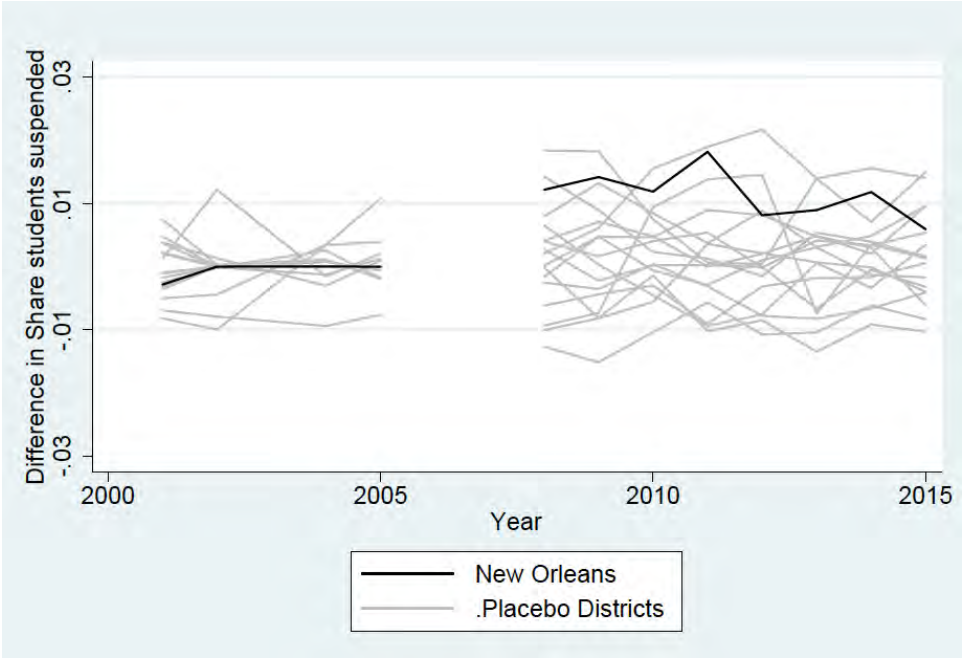
Panel B. Panel Sample



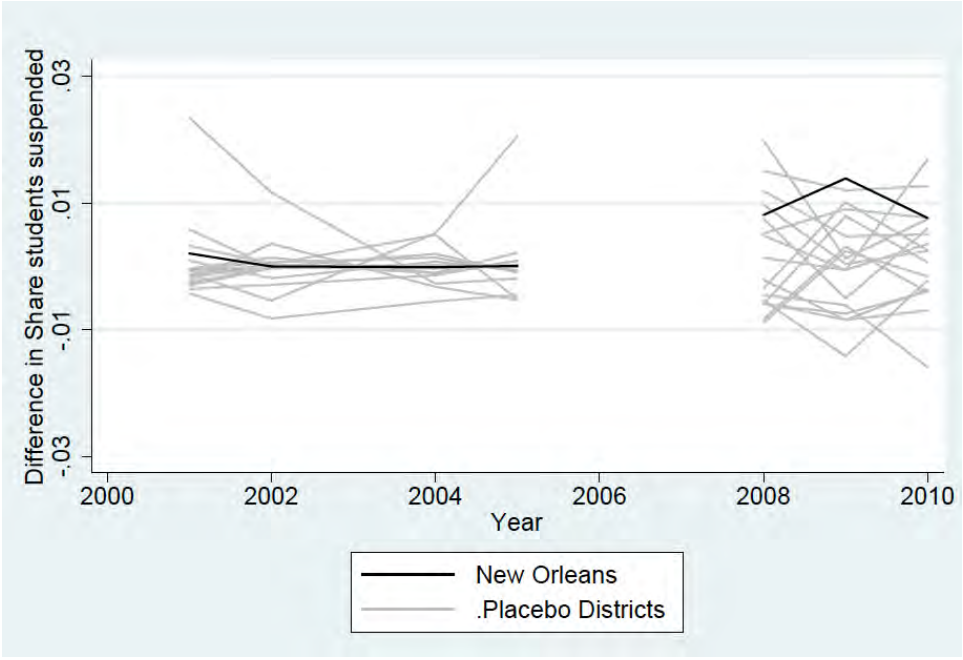
Notes: Specific offenses are well-defined offenses that can vary in severity, such as getting involved in fights or being habitually tardy. Refer to notes in Figure 6 for more details

Figure 12. Effects on SOS for specific and serious offenses

Panel A. Pooled sample



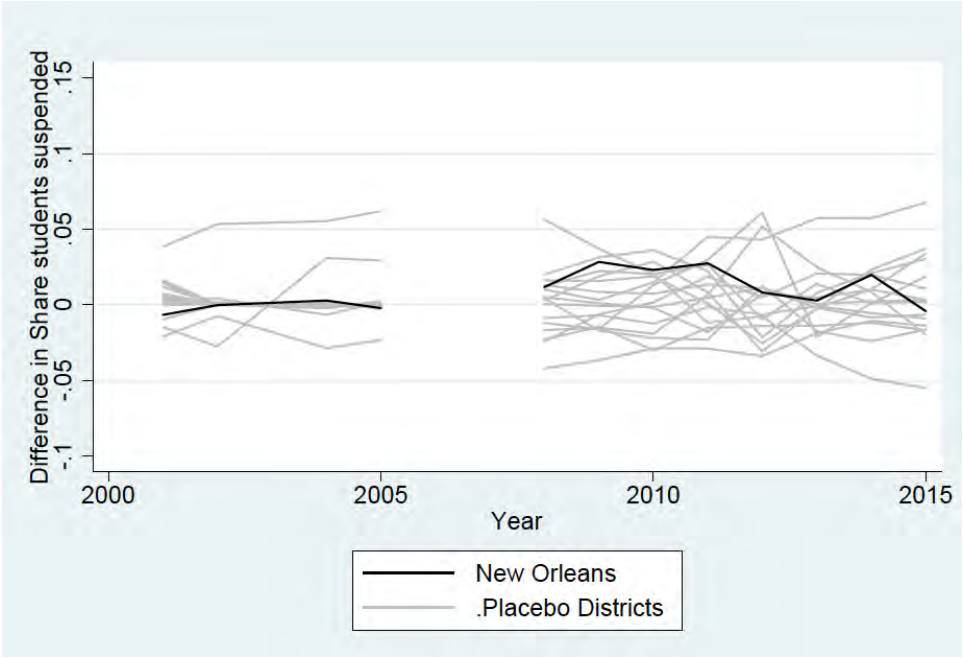
Panel B. Panel Sample of Returnees



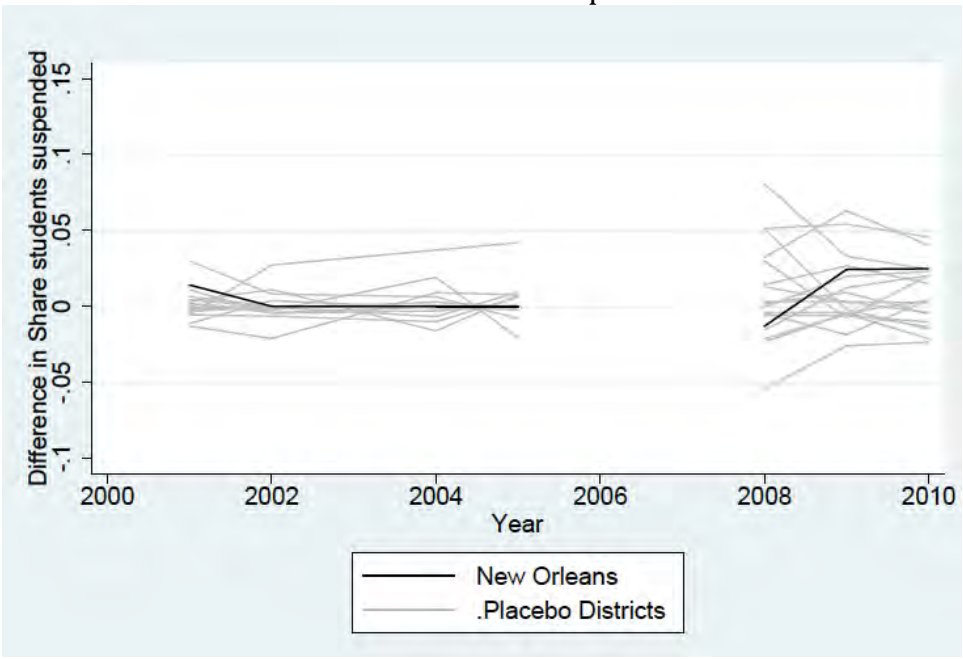
Notes: Specific and serious offenses correspond to well-defined offenses that are high in severity, such as fights or incidents reportable to law enforcement agencies. For more details refer to notes in Figure 6

Figure 13. Effects on SOS for specific and non-serious offenses

Panel A. Pooled sample



Panel B. Panel sample



Notes: Specific and non-serious offenses correspond to well-defined offenses that are low in severity, such as leaving the classroom without permission or using profane language. Refer to notes in Figure 6 for more details

Table 1. Number and Percent of Suspensions by Infraction Type

		Percent of Total		
	N. suspensions	Non-specific Offenses	Specific Offenses	
			Serious	Non-serious
Infraction Type	(1)	(2)	(3)	(4)
Willful disobedience	905,061	23.7		
Habitually violates rule	505,765	13.2		
Fights in school	495,681		13.0	
Disrespects authority	483,167	12.6		
Any other serious offense	285,716	7.5		
Leaves school/classroom	267,153			7.0
Profane	244,298			6.4
Habitually tardy/absent	210,297			5.5
Injurious habits	116,850		3.1	
Immoral practices	59,228	1.6		
Tobacco	42,491			1.1
Stealing	35,097		0.9	
Failure to Serve Assigned Conseq	28,039	0.7		
Unauthorized use of Technology	23,910	0.6		
Substance governed by Law	17,887		0.5	
Throws missiles to injure others	17,923		0.5	
Other	87,372	0.5	1.0	0.8
Total	3,720,640	60.8	19.9	21.6

Notes: The unit of observation is suspensions, so some students have multiple observations within the same year. The sample consists of all infractions that led to suspensions among students in grades 5-12 from 2001 through 2015. Columns 2-4 classify suspensions in mutually exclusive categories according to the infraction specificity and severity: Non-specific, Specific, serious and non-serious. The numbers in these columns represent the number of suspensions of that type as a percent of the total number of suspensions. The row Other aggregates all infractions within a specific category that occurred at very low frequencies. Possessing weapons, for instance, is included as a serious offense in row Other.

Table 2. Pooled Sample: Student Population Characteristics

	New Orleans	Rest LA
Number of observations		
Student-year	421,339	5,630,287
Students	116,643	1,269,506
Schools	257	1,485
Districts	1	65
Demographics		
Black	0.906	0.413
Hispanic	0.020	0.028
White	0.048	0.532
Lunch	0.696	0.561
Special education	0.090	0.100
Discipline Variables		
Share students suspended	0.200	0.252
Share suspended out-of-school	0.186	0.144
Share suspended in-school	0.030	0.157
Share expelled	0.009	0.014

Note: Sample consists of all students in grades 5th-12th in 2001-2015

Table 3. Synthetic weights of predictors, by outcome

Panel A. Pooled sample

	Expulsions	Out-of-school suspensions				
		Total	Non-specific	Specific		
				Total specific	Serious	Non-serious
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome 2002	0.303	0.509	0.477	0.508	0.420	0.505
Outcome 2004	0.354	0.251	0.270	0.245	0.282	0.241
Outcome 2005	0.342	0.232	0.247	0.244	0.297	0.250
N. students	0.002	0.007	0.006	0.003	0.000	0.004

Panel B. Panel sample

	Expulsions	Out-of-school suspensions				
		Total	Non-specific	Specific		
				Total specific	Serious	Non-serious
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome 2002	0.021	0.242	0.206	0.272	0.229	0.253
Outcome 2004	0.344	0.319	0.328	0.299	0.238	0.315
Outcome 2005	0.635	0.438	0.465	0.428	0.525	0.431
N. students	0.000	0.000	0.001	0.000	0.007	0.001

Notes: The rows represent the predictors, while the columns represent the outcomes. Columns 2-6 classify out-of-school suspensions (OSS) according to the infraction specificity and severity: Non-specific, specific and serious, specific and non-serious. The sample in Panel A consists of all students enrolled in grades 5-12 in New Orleans and other large school districts in Louisiana. The sample in Panel B consists of students in the 2005 3rd grade cohort who returned to their pre-Katrina school district.

Table 4. Pooled sample: Weights of school districts, by outcome

	Pct students expelled	Pct. students suspended out-of-school				
				Specific Offenses		
		All offenses	Non-specific offenses	All Specific Offenses	Serious offenses	Non-serious offenses
	(1)	(2)	(3)	(4)	(5)	(6)
Ascension	0.12	0	0	0	0.13	0
Bossier	0	0	0	0	0	0
Caddo	0.45	0	0	0.44	0	0.07
Calcasieu	0	0	0	0	0	0
East Baton Rouge	0	0	0	0.22	0.13	0
Iberia	0	0	0	0	0	0
Jefferson	0	0.89	0.68	0	0	0.75
Lafayette	0	0.09	0.26	0.28	0	0
Lafourche	0	0	0	0	0	0
Livingston	0	0	0	0	0	0.13
Orleans	0	0	0	0	0	0
Ouachita	0	0	0	0	0	0
Rapides	0	0.02	0.06	0.06	0	0.06
St. Landry	0	0	0	0	0.03	0
St. Tammany	0.08	0	0	0	0.72	0
Terrebonne	0.35	0	0	0	0	0
N. districts in the donor pool	16	16	16	16	16	16
N. districts in synthetic control	3	3	3	4	4	4

Notes: This Table lists only those districts that had a positive weight for at least one of the outcomes. The donor pool is the group of districts that the SCG method uses to create the synthetic control (16 districts total). This pool consists of large school districts, with at least a fifth of New Orleans' students population in 2005 (at least 7,000 students grades in 5-12 in 2005). The districts in the synthetic control are those districts that contribute to the synthetic control (e.g. that have a positive weight). Refer to Table 3 for further details.

Table 5 Panel sample: Weights of school districts, by outcome

	Pct students expelled	Pct. students suspended out-of-school				
				Specific Offenses		
		All offenses	Non-specific offenses	All Specific Offenses	Serious offenses	Non-serious offenses
	(1)	(2)	(3)	(4)	(5)	(6)
Ascension	0.65	0	0	0.002	0	0.019
Bossier	0	0	0	0.002	0	0
Caddo	0	1.00	0	0.735	0.10	0.798
Calcasieu	0	0	0	0.001	0	0
East Baton Rouge	0	0	0	0	0.14	0
Iberia	0	0	0	0.001	0	0
Jefferson	0	0	1.00	0.004	0.04	0
Lafayette	0	0	0	0.002	0.72	0
Lafourche	0	0	0	0.001	0	0
Livingston	0	0	0	0.001	0	0
Orleans	0	0	0	0	0	0
Ouachita	0	0	0	0	0	0.133
Rapides	0	0	0	0.002	0	0.049
St. Landry	0	0	0	0.002	0	0
St. Tammany	0	0	0	0.001	0	0
Tangipahoa	0.33	0	0	0.141	0	0
Terrebonne	0.02	0	0	0.104	0	0
N. districts in the donor pool	17	17	17	17	17	17
N. districts in the synthetic control	2	1	1	13	4	4

Notes: This Table lists only those districts that had a positive weight for at least one of the outcomes. The donor pool is the group of districts that the SCG method uses to create the synthetic control (17 districts in total). This pool consists of large school districts, with at least a fifth of New Orleans' students population in 2005 (at least least 500 7th grade students in 2005 who returned to their original school district post-Katrina). The districts in the synthetic control are those districts that contribute to the synthetic control (e.g. that have a positive weight). Refer to Table 3 for further details.

Table 6 Pre-reform characteristics in New Orleans and Synthetic Control

Panel A. Pooled sample

	Expulsions			Out-of-school suspensions		
	New Orleans	Synthetic Control	LA school districts	New Orleans	Synthetic Control	LA school districts
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome 2002	0.012	0.012	0.014	0.201	0.204	0.137
Outcome 2004	0.009	0.010	0.016	0.238	0.217	0.151
Outcome 2005	0.011	0.011	0.017	0.238	0.239	0.161
N. students	38,935	18,623	15,406	38,935	28,195	15,406

Panel B. Panel sample

	Expulsions			Out-of-school suspensions		
	New Orleans	Synthetic Control	LA school districts	New Orleans	Synthetic Control	LA school districts
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome 2002	0.003	0.003	0.002	0.163	0.155	0.075
Outcome 2004	0.002	0.002	0.010	0.243	0.225	0.142
Outcome 2005	0.005	0.005	0.018	0.284	0.266	0.169
N. students	2,397	1,730	1,592	2,397	2,885	1,592

Notes: The rows represent the predictors, while the columns represent the outcomes. Rows 1-3 present the average outcome in 2002, 2004 and 2005, while row 4 presents the average number of students during the period 2001-2005. Columns 1-3 compare the pre-reform characteristics of New Orleans (Column 1) to those of the synthetic control group specific to the expulsions outcome (Column 2), and also to those of a student-weighted average of the rest of school districts in Louisiana state (Column 3). Columns 4-6 make the same comparison against the synthetic control group specific to the out-of-school suspensions outcome.

Table 7. Characteristics of Returnees and Non-Returnees, 2005

2005 7th grade cohort	New Orleans		Hurricane-Affected Districts	
	All Pre-Katrina students	Returnees only	All Pre-Katrina students	Returnees only
	(1)	(2)	(3)	(4)
N. students	4780	2045	10697	6479
Demographic characteristics (share)				
Black	0.942	0.934	0.337	0.304
Hispanic	0.012	0.009	0.045	0.045
Other Race	0.017	0.020	0.035	0.034
White	0.029	0.037	0.582	0.617
Lunch	0.748	0.759	0.531	0.503
Special education	0.101	0.069	0.123	0.105
Discipline Outcomes (share)				
Expelled	0.007	0.004	0.025	0.011
Share suspended out-of-school	0.268	0.259	0.201	0.144
Non-specific offenses	0.176	0.166	0.133	0.088
Specific offenses	0.156	0.153	0.115	0.083
Serious offenses	0.019	0.018	0.032	0.025
Non-serious offenses	0.145	0.144	0.092	0.064
Test scores (std. dev)				
ELA test	-0.526	-0.469	0.129	0.259
Math test	-0.474	-0.417	0.165	0.284
Science test	-0.597	-0.577	0.199	0.309

Notes: The sample consists of students in the 7th in 2005. Columns (1) and (3) refer to all students that were enrolled in 2005, while columns (2) and (4) refer only to students who came back to their original school districts after Katrina. Hurricane-affected districts include Jefferson, Calcasieu, St. Tammany, Cameron, Plaquemines, St. Bernard and Vermilion.

Table 8. Pooled sample: Effects on discipline outcomes

Av. Treatment Effect in Year	Pct students expelled	Pct. students suspended out-of-school				
				Specific Offenses		
		All offenses	Non-specific offenses	All Specific Offenses	Serious offenses	Non-serious offenses
	(1)	(2)	(3)	(4)	(5)	(6)
2001	0.002	-0.031	-0.028	-0.007	-0.003	-0.007
2002	0.000	-0.003	-0.004	0.000	0.000	0.000
2004	0.000	0.021	0.015	0.002	0.000	0.003
2005	0.000	0.000	0.000	-0.002	0.000	-0.002
2008	0.000	0.02	-0.068	0.031	0.012	0.012
2009	0.015	0.042	0.017	0.047	0.014	0.028
2010	0.007	0.04	0.017	0.033	0.012	0.023
2011	-0.003	0.022	0.006	0.055	0.018	0.027
2012	-0.015	0.000	-0.063	0.019	0.008	0.008
2013	-0.011	0.003	-0.063	0.056	0.009	0.003
2014	-0.007	0.014	-0.014	0.065	0.012	0.028
2015	-0.007	-0.019	-0.002	0.046	0.006	-0.004
NOLA mean 2005	0.011	0.238	0.171	0.119	0.015	0.109
N. districts in synthetic control	4	3	3	4	4	4
N. placebo districts	16	16	16	16	16	16

Notes: The sample consists of students in grades 5-12 in 2001 through 2015, excluding 2003, 2006 and 2007. The table presents the difference in discipline outcomes between New Orleans and its synthetic control in the corresponding year. The number of districts in the synthetic control corresponds to those districts with a positive weight in the counterfactual group. The number of placebo districts is the same as the number of districts in the donor pool and consists of large school districts (e.g. with at least 7,000 students in grades 5-12 in 2005) whose average pre-Katrina RMSPE is less than 0.1.

Table 9. Pooled Sample: Number of placebo districts with effect sizes larger than that of New Orleans

Av. Treatment Effect in Year	Pct students expelled	Pct. students suspended out-of-school				
		All offenses	Non-specific offenses	Specific Offenses		
				All Specific Offenses	Serious offenses	Non-serious offenses
	(1)	(2)	(3)	(4)	(5)	(6)
2001	6	10	10	8	6	9
2002	11	7	4	10	13	9
2004	8	4	2	8	10	6
2005	11	11	14	8	11	8
2008	14	13	5	4	1	8
2009	2	11	14	3	1	3
2010	6	11	11	4	1	4
2011	8	11	15	2	0	4
2012	3	15	5	7	5	9
2013	4	14	3	4	4	10
2014	6	12	12	2	1	5
2015	6	9	14	4	5	10
N. districts in synthetic control	4	3	3	4	4	4
N. placebo districts	16	16	16	16	16	16

Notes: The effect size in a given year is defined as the ratio of the difference in discipline rates between the placebo and its synthetic control and the pre-reform RMSPE. The number of placebo districts consists of large school districts (e.g. with at least 7,000 students in grades 5-12 in 2005) whose average pre-Katrina RMSPE is less than 0.1.

Table 10. Panel Sample: Effects on Discipline Outcomes

Av. Treatment Effect in Year	Pct students expelled	Pct. students suspended out-of-school				
		All offenses	Non-specific offenses	Specific Offenses		
				All Specific Offenses	Serious offenses	Non-serious offenses
	(1)	(2)	(3)	(4)	(5)	(6)
2001	-0.001	0.006	-0.022	0.013	0.002	0.014
2002	0.000	0.008	0.014	0.000	0.000	0.000
2004	0.000	0.018	-0.001	0.000	0.000	0.000
2005	0.000	0.019	0.025	0.000	0.000	0.000
2008	0.008	0.018	-0.019	-0.004	0.008	-0.013
2009	0.028	0.075	0.089	0.038	0.014	0.024
2010	0.011	0.039	0.043	0.029	0.008	0.025
NOLA mean 2005	0.005	0.284	0.175	0.169	0.021	0.158
N. districts in synthetic control	3	1	1	15	4	4
N. placebo districts	17	17	17	17	17	17

Notes: The sample consists of students who were in the 7th grade in 2005 and who returned to their pre-Katrina school district at least for one year post-reform. Refer to notes in Table 8 for further details.

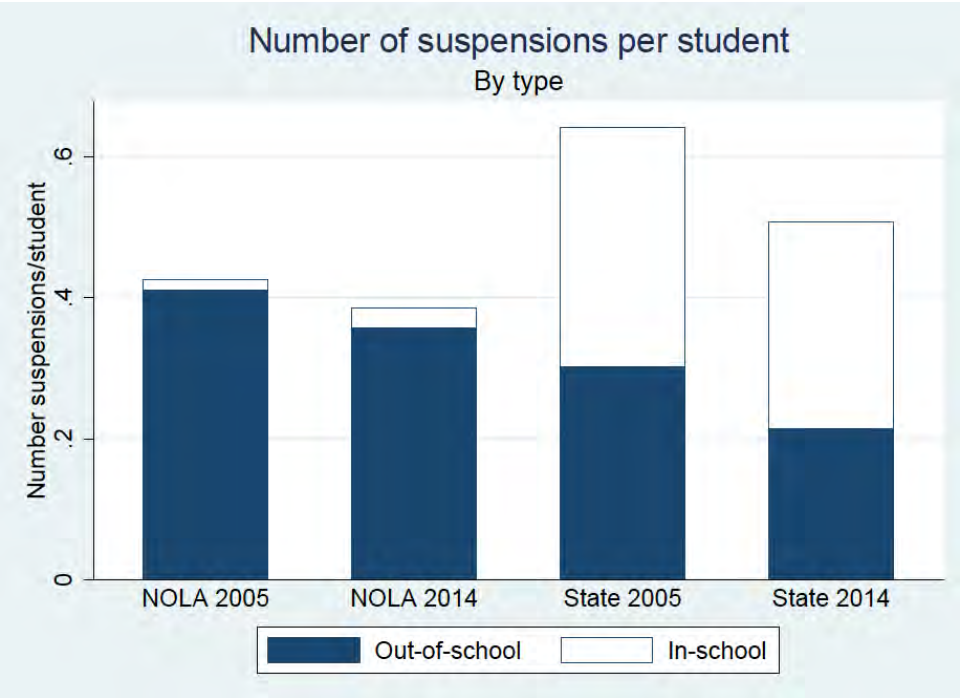
Table 11. Panel Sample: Number of placebo districts with effect sizes larger than that of New Orleans

Av. Treatment Effect in Year	Pct students expelled	Pct. students suspended out-of-school				
		All offenses	Non-specific offenses	Specific Offenses		
				All Offenses	Serious offenses	Non-serious offenses
	(1)	(2)	(3)	(4)	(5)	(6)
2001	2	14	11	1	3	2
2002	15	6	4	17	16	15
2004	17	3	12	15	14	15
2005	15	3	1	15	13	16
2008	3	13	13	13	4	7
2009	0	9	7	6	1	9
2010	0	9	8	6	4	8
N. districts in synthetic control	3	1	1	15	4	4
N. placebo districts	17	17	17	17	17	17

Notes: The effect size in a given year is defined as the ratio of the difference in discipline rates between the placebo and its synthetic control and the pre-reform RMSPE. The number of placebo districts consists of large school districts (e.g. with at least 500 students in grades 5-12 in 2005) whose average pre-Katrina RMSPE is less than 0.1.

Appendix A

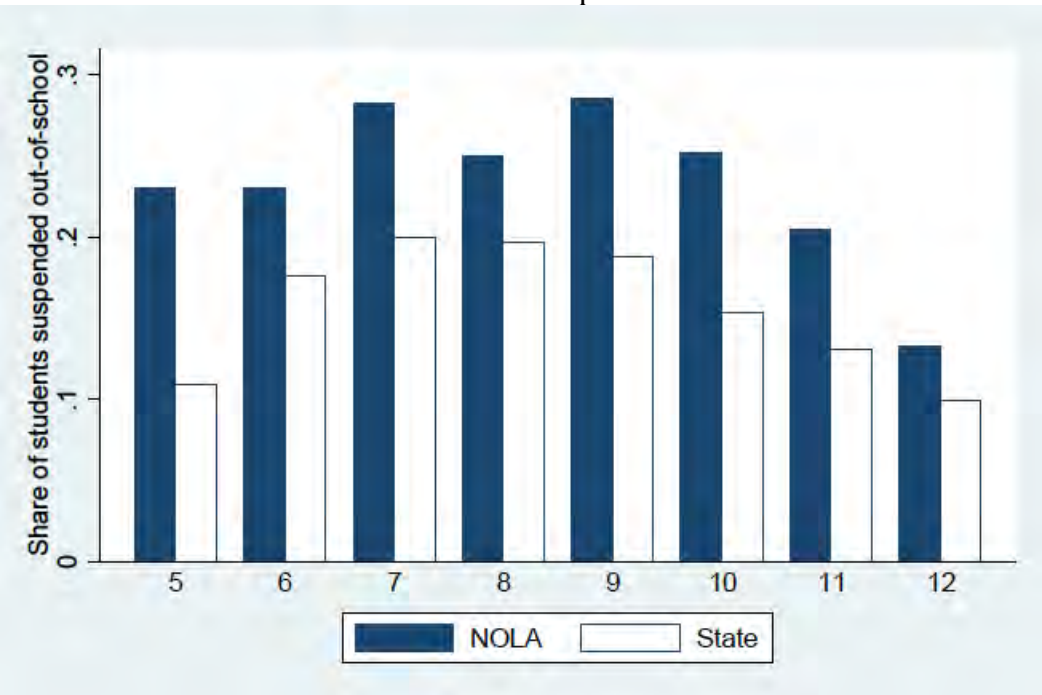
Figure A.1.



Notes: The bars represent the number of suspensions per student either in-school or out-of-school. For more details refer to notes in Figure 1

Figure A.2. Discipline rates by grade, 2005

Panel A. Share of Students Suspended Out-of-school



Panel B. Share of students expelled

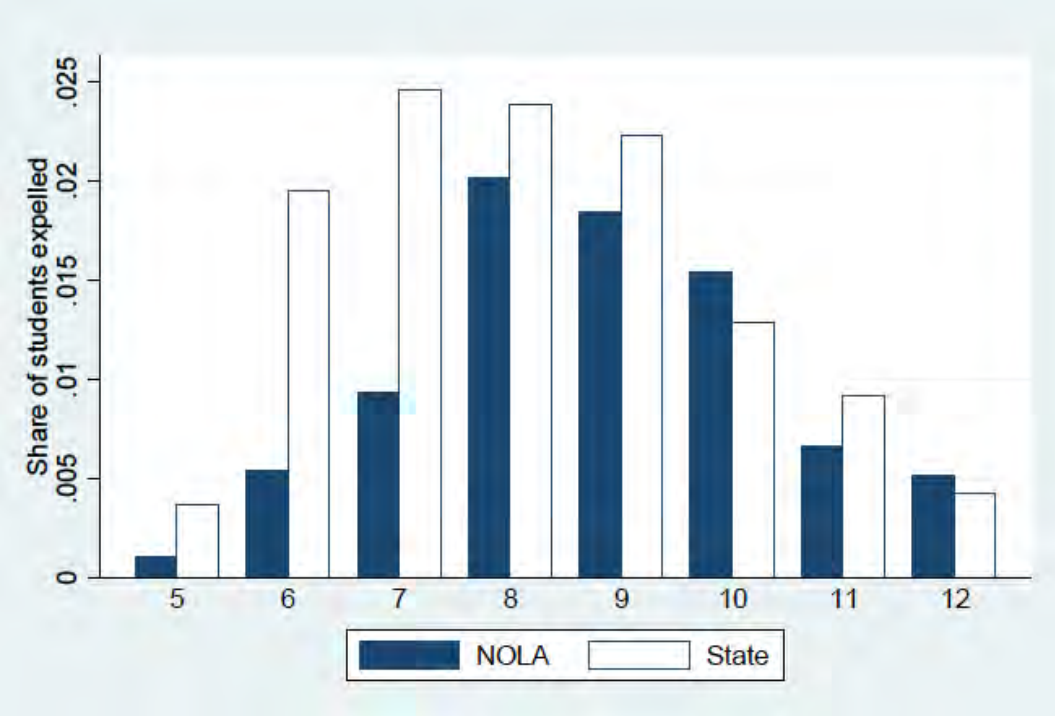
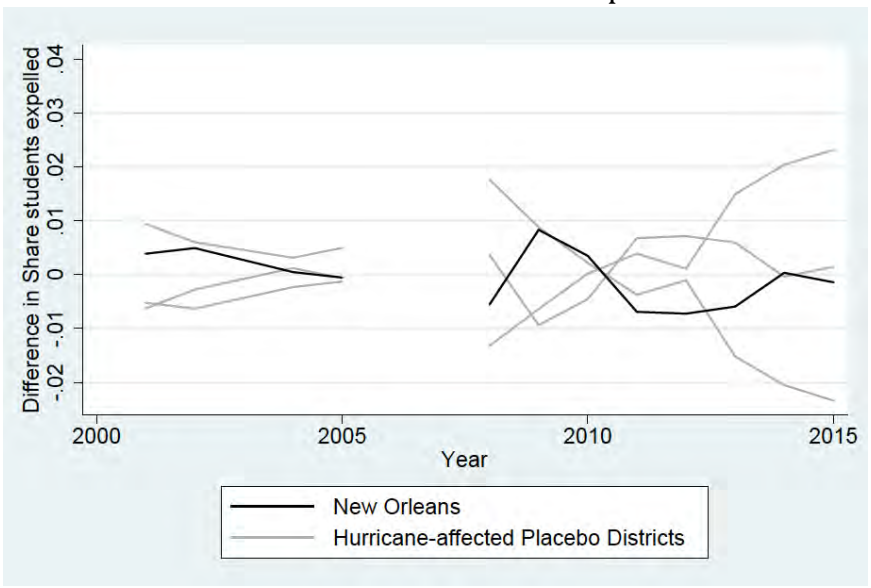
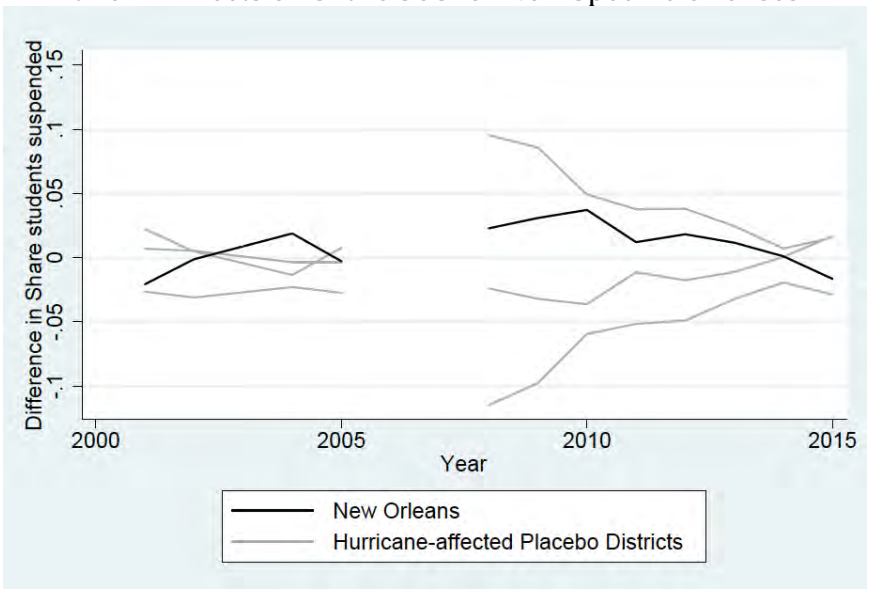


Figure A.3. Pooled Sample: Effects Including Hurricane-affected Districts Only

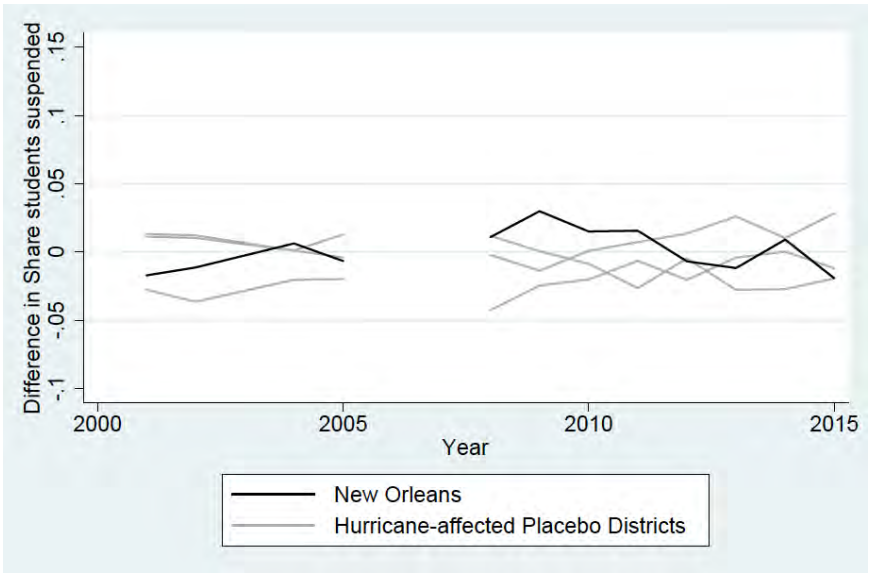
Panel A. Effects on Share of Students Expelled



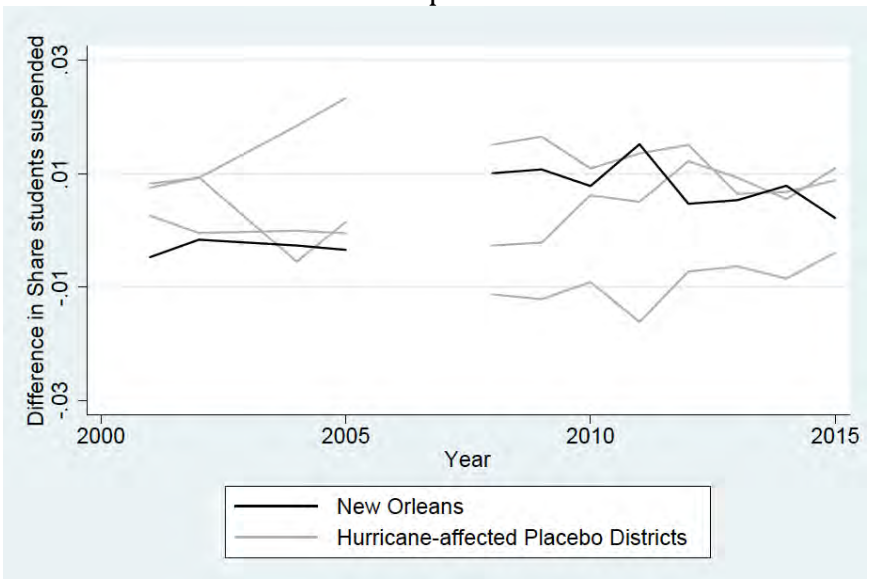
Panel B. Effects on Share SOS for Non-Specific Offenses



Panel C. Effects on Share SOS for Specific Offenses



Panel D. Effects on SOS for Specific and Serious Offenses

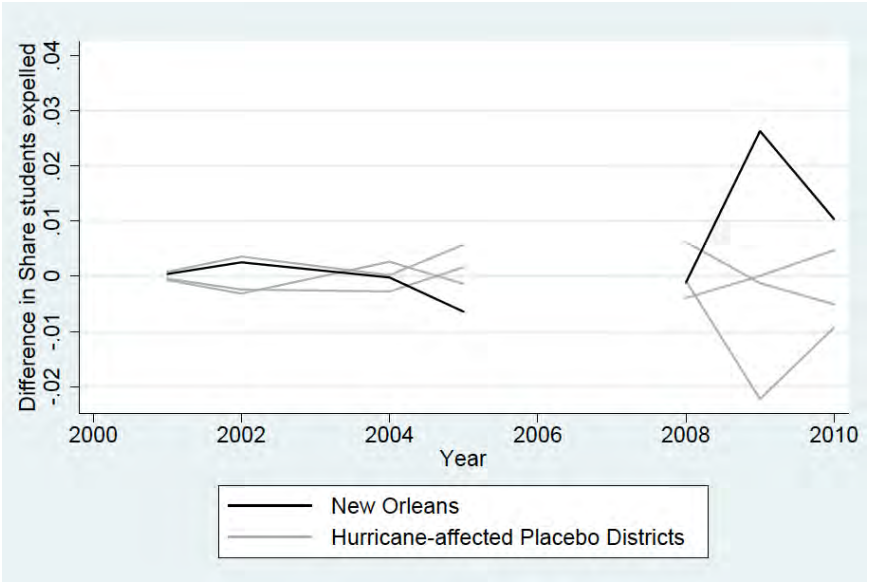


Notes: The sample consist of all students in hurricane-affected districts enrolled in grades 5-12 between 2001 through 2015. The bold line represents the difference in the discipline outcome between New Orleans and its synthetic control, while the gray lines represent the same difference for the rest of hurricane-affected placebo districts. The placebo includes hurricane-affected school districts with at least one fifth of the student population in New Orleans in 2005.

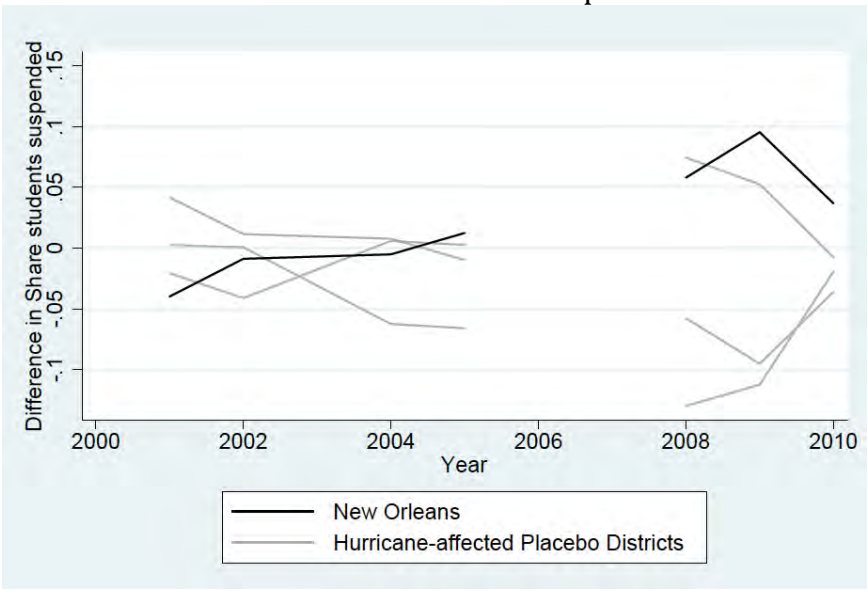
.

Figure A.4. Panel Sample: Effects Including Hurricane-affected Districts Only

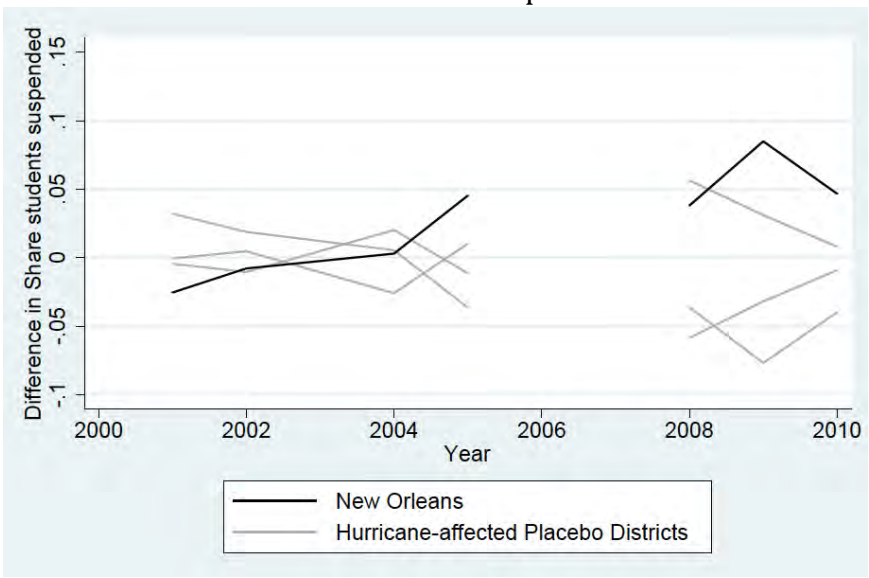
Panel A. Effects on Share of Students Expelled



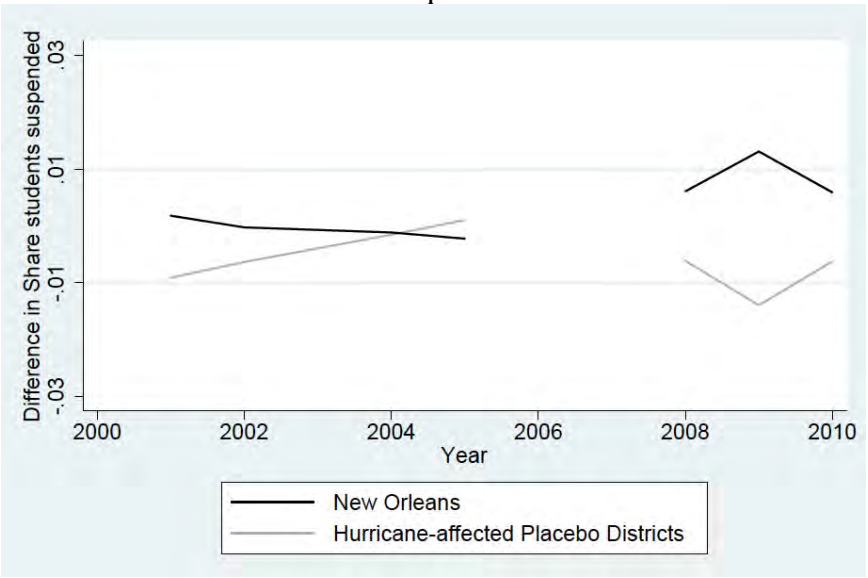
Panel B. Effects on Share SOS for Non-Specific Offenses



Panel C. Effects on Share SOS for Specific Offenses



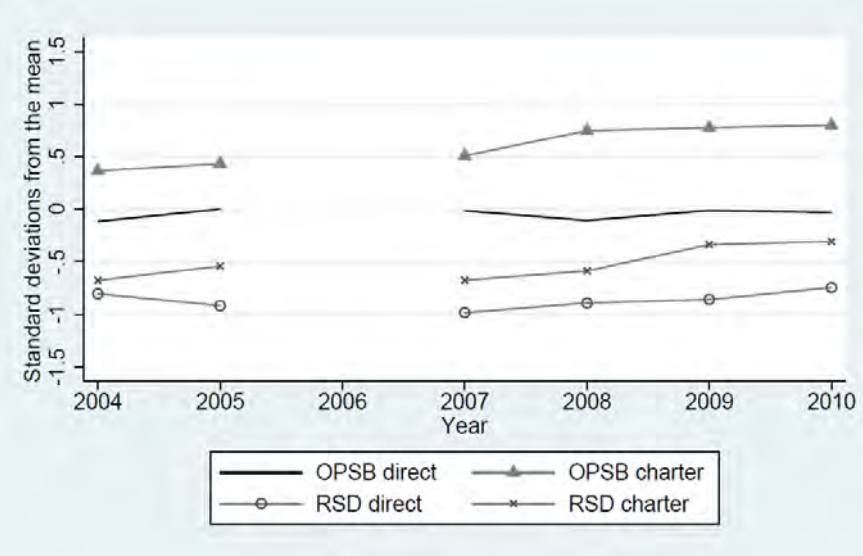
Panel D. Effects on SOS for Specific and Serious Offenses



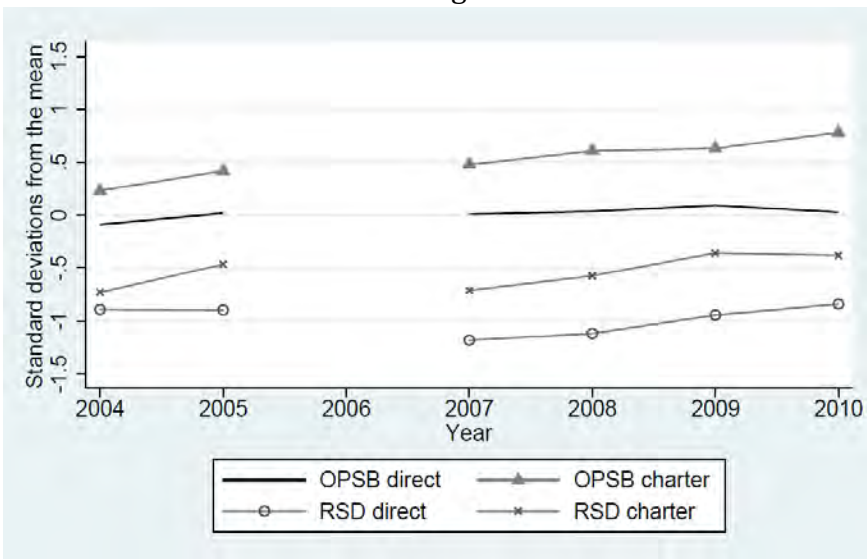
Notes: The sample consists of students who were in the 7th grade in 2005 and who returned to their pre-Katrina school district by 2009. Refer to notes in Figure A.3. for other details.

Figure A.5. Achievement and discipline outcomes by education sector (2005 4th grade cohort of returnees)

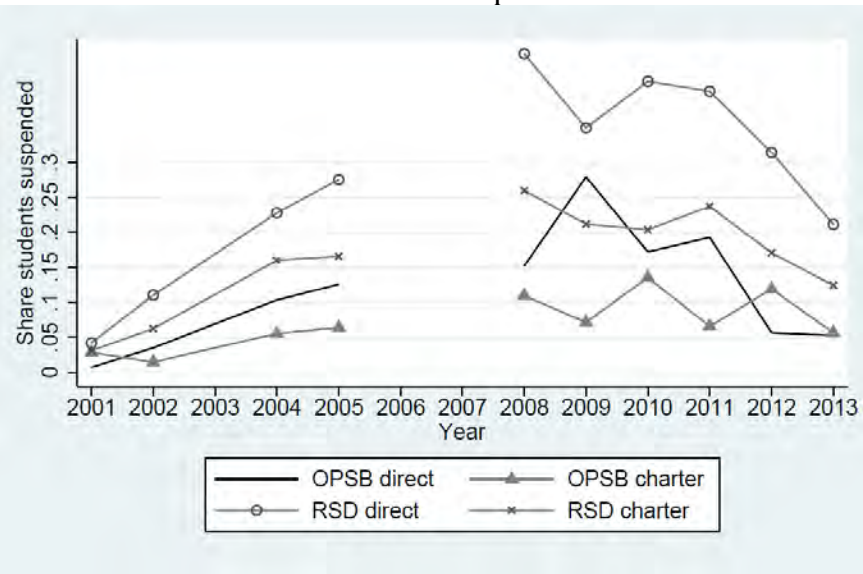
Panel A. Average Math test scores



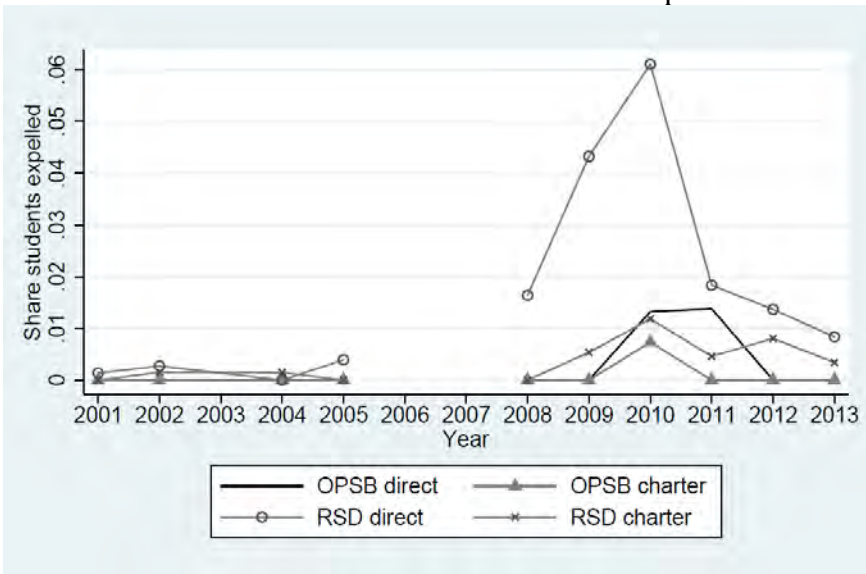
Panel B. Average ELA test scores



Panel C. Share Students Suspended Out-of-school



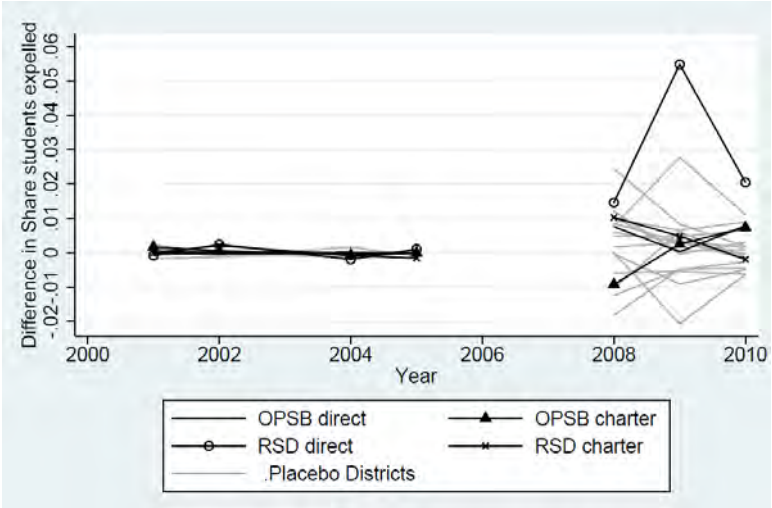
Panel D. Share Students Expelled



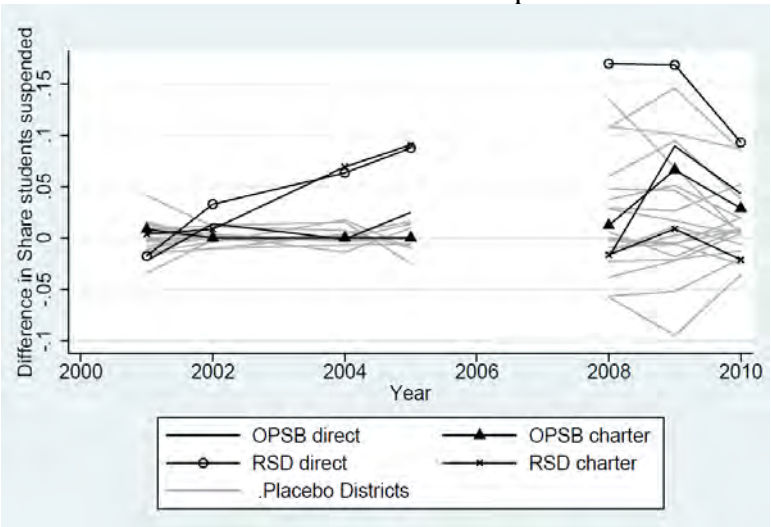
Notes: The sample consists of students who were in the 4th grade in 2005 for the first time, and returned to New Orleans by 2009. Students are classified according to the education sector they were enrolled in 2009. Test scores in Panels A and B are standardized by grade-year.

Figure A.6. Panel Sample: Effects on discipline outcomes, treating education sectors as separate districts

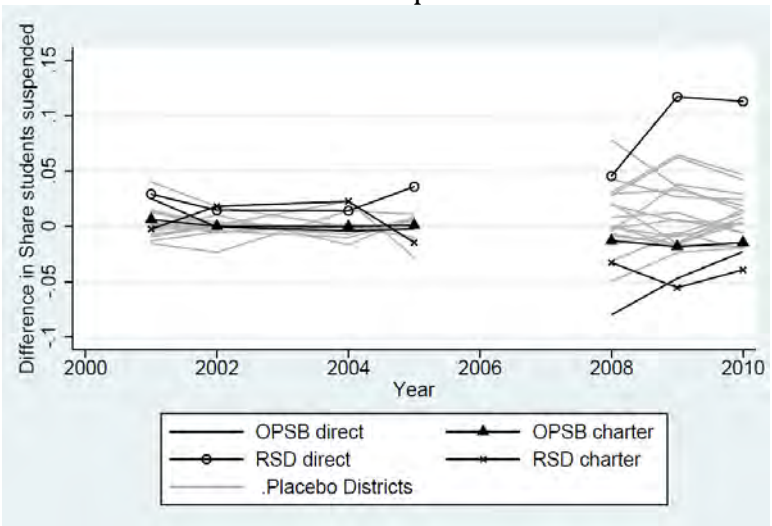
Panel A. Share of students Expelled



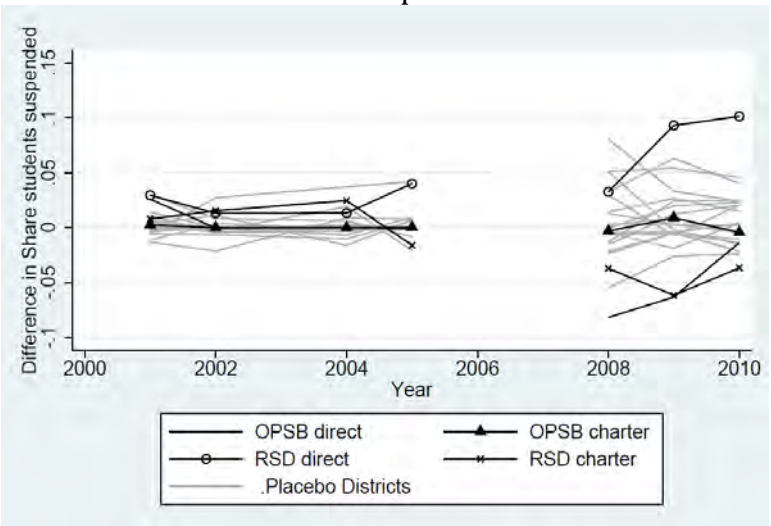
Panel B. Share SOS for non-specific offenses



Panel C. Share SOS for specific offenses



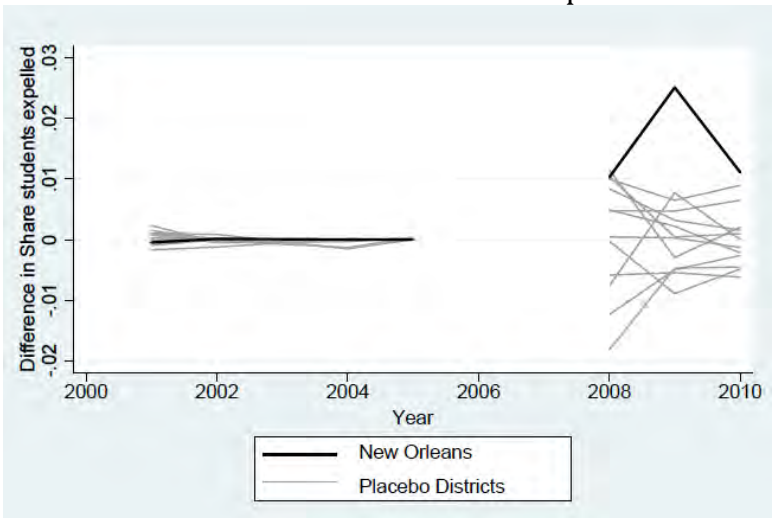
Panel D. Share SOS for specific and serious offenses



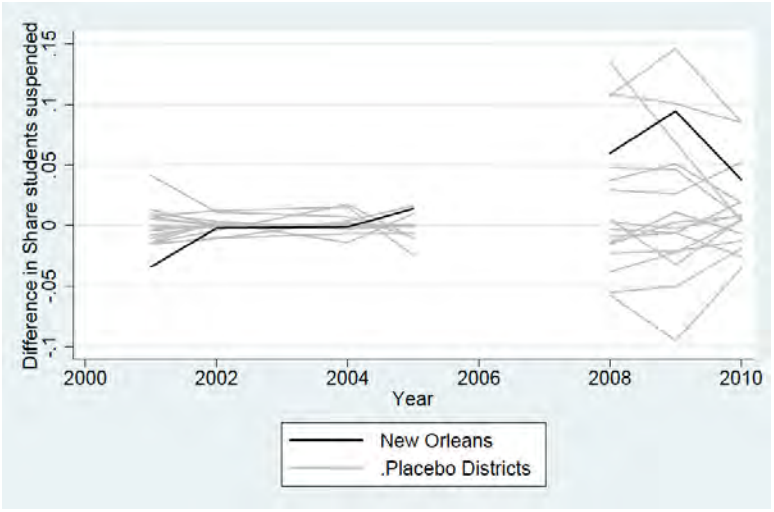
Notes: The sample consists of students who were in the 7th grade in 2005 and who returned to their pre-Katrina school district by 2009. Students are classified according to the education sector they were enrolled in 2009. The line for each education sector represents the difference in discipline outcomes between students who were in that sector in 2009 and the sector's synthetic control. The synthetic control is constructed based on pre-Katrina out-of-school suspensions rate, 6th and 7th grade average scores in math and ELA and the corresponding discipline outcome. The gray lines represent the same difference for all the placebo districts.

Figure A.7. Panel Sample: Effects omitting Caddo parish from the donor pool

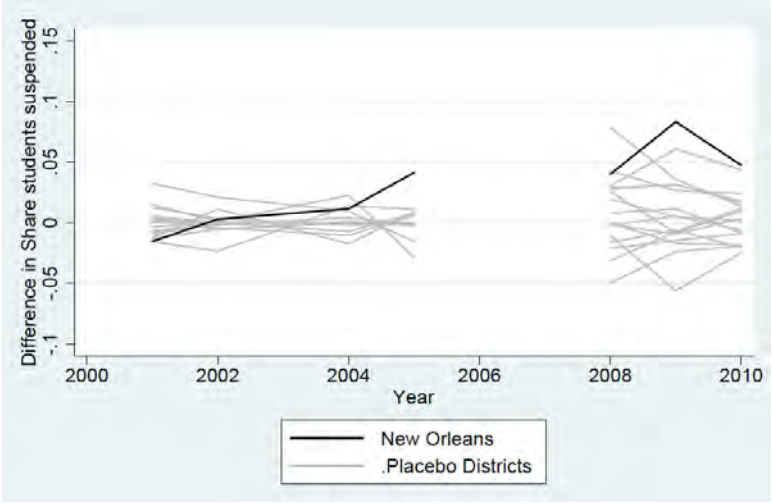
Panel A. Effects on share of students expelled



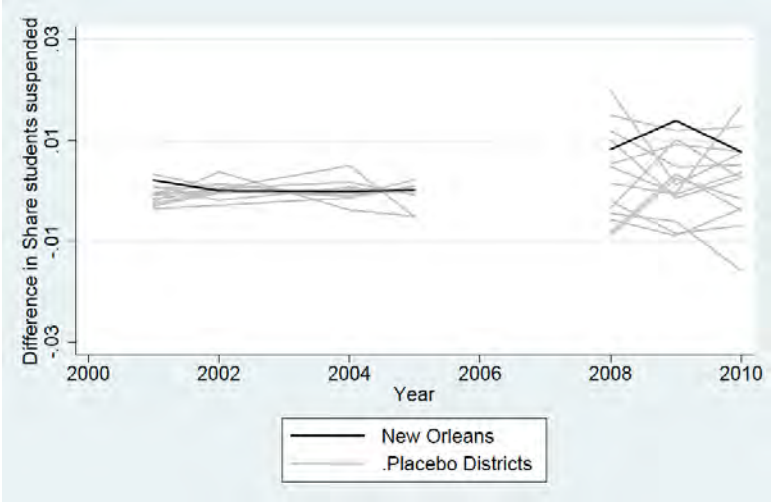
Panel B. Effects on share SOS for non-specific offenses



Panel C. Effects on share SOS for specific offenses



Panel D. Effect on SOS for specific and serious offenses



Appendix B. Excluding year 2003 from the sample

The year 2003 is excluded from the sample use to estimate the effect of the New Orleans school reforms on school discipline. This is because New Orleans displayed an atypical increase in 2003 expulsion rates, as shown in Figure B.1. This increase was probably due to the enactment of NCLB in January 2002. When NCLB was enacted, Louisiana had a pre-established accountability system in place that was based on test performance (among other factors). However, this pre-established system did not require schools to report and be accountable for the performance of students belonging to minority groups (the subgroup requirement). This changed when NCLB was enacted. This new requirement posed a significant challenge for New Orleans public schools, where black students compose the majority of students. School leaders might have initially reacted to this requirement by expelling low-performing students, before the administration of standardized achievement tests happened.

Figure B.2. provides evidence supporting this point. When doing a month-by-month comparison of New Orleans expulsions in 2003, relative to expulsions in the two adjacent years (2002 and 2004), expulsions in 2003 were consistently higher every month (Panel A). The difference is considerably larger during the month of March, which is precisely the time when students take standardized achievement tests. In the subsequent months, 2003 expulsions drop *below* those in 2002 and 2004. This pattern is not observed in the rest of Louisiana, where the racial composition of students is more balanced. As Panel B shows, the number of expulsions in a specific month is similar from 2002 through 2004. These pieces of evidence are consistent with the argument that New Orleans was particularly challenged by the new NCLB requirement, because of the demographic composition of its students, which might have initially led school to behave strategically.

Figure B.1 Share students expelled, including year 2003

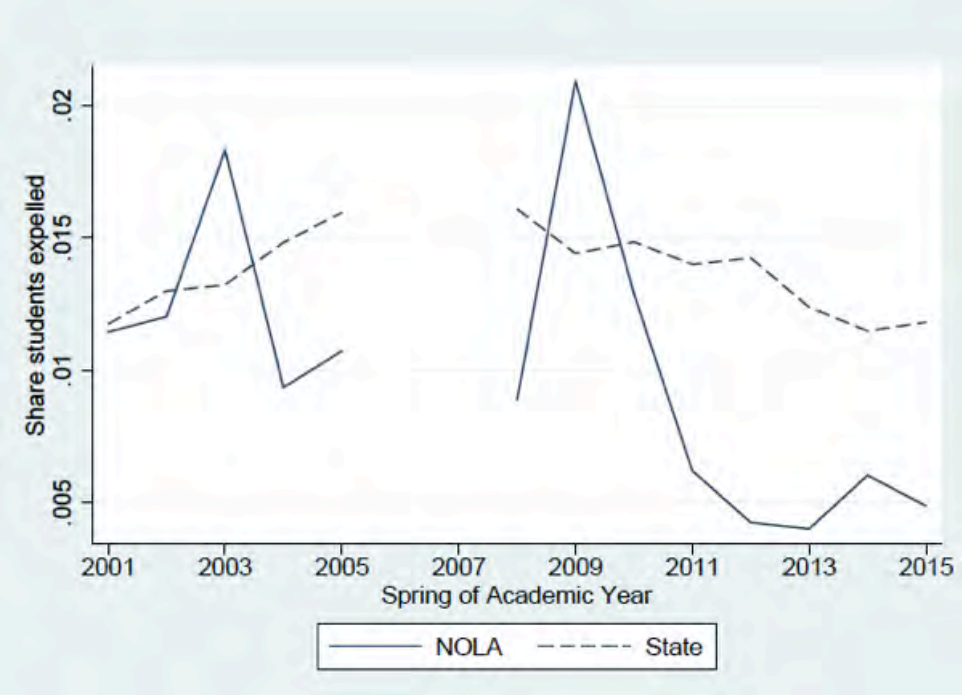
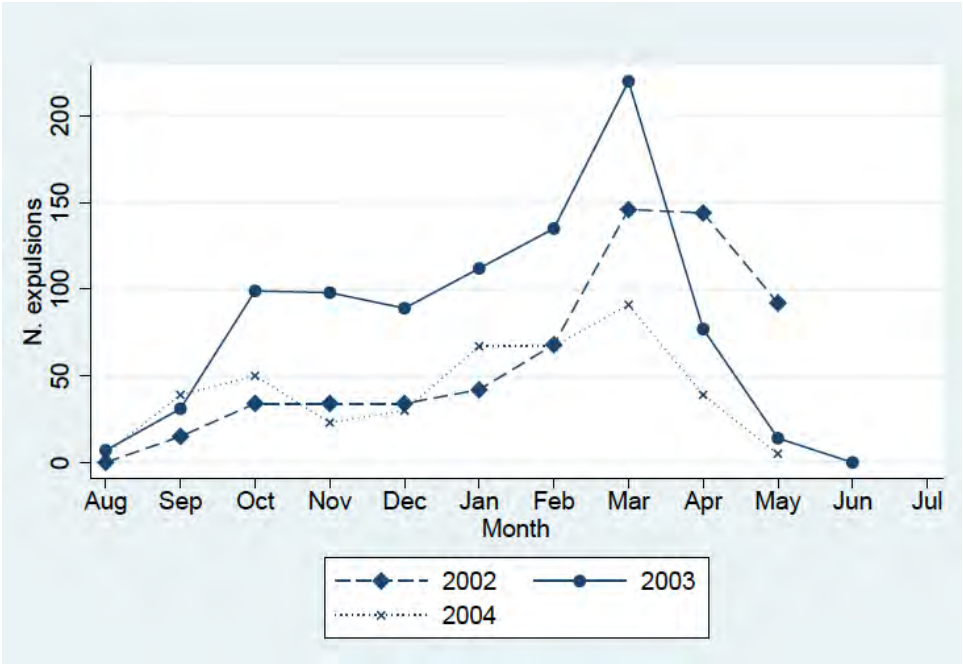
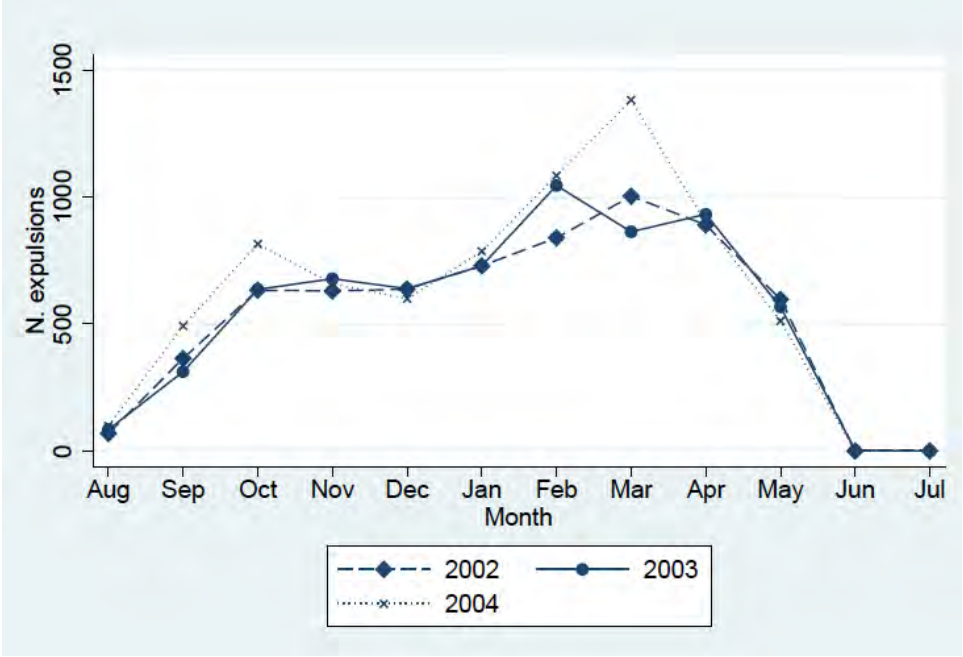


Figure B.2 Total number of expulsions by month, 2002-2004

Panel A. New Orleans



Panel B. Rest of the State



Notes: Number of expulsions corresponds to the total number of expulsions in a given month and year.