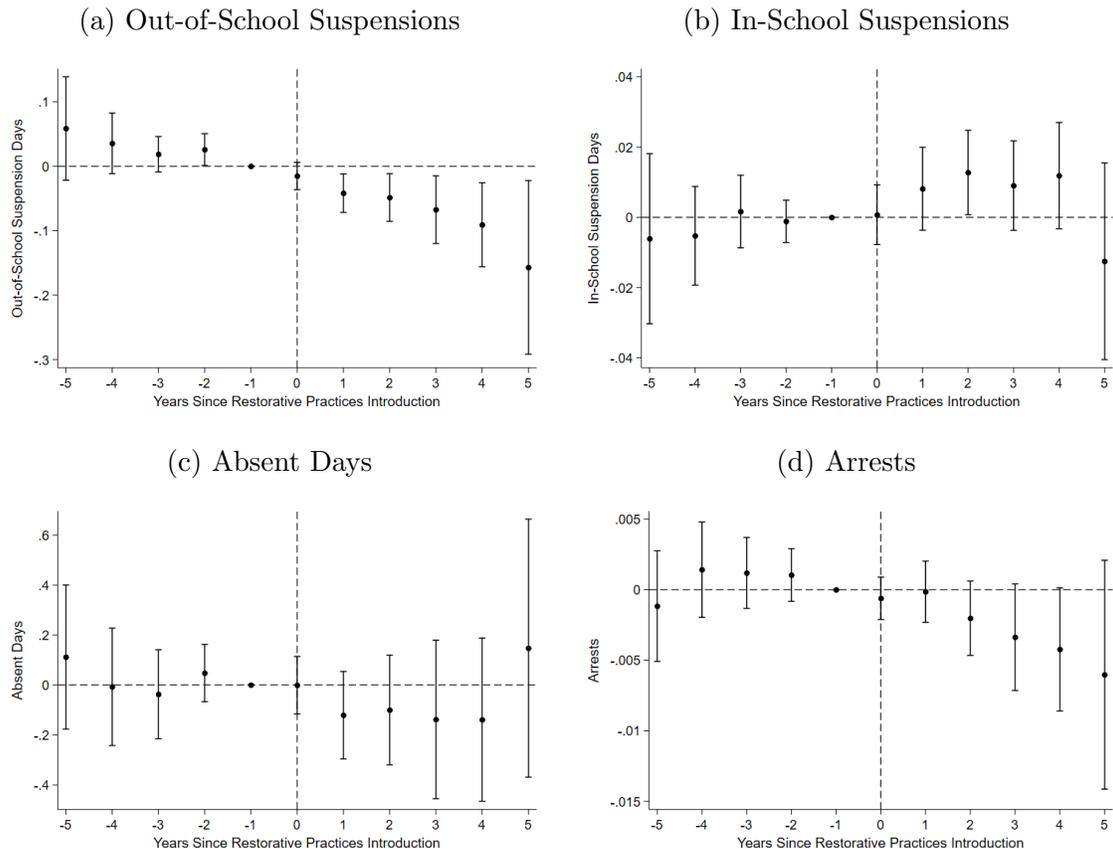


Appendices

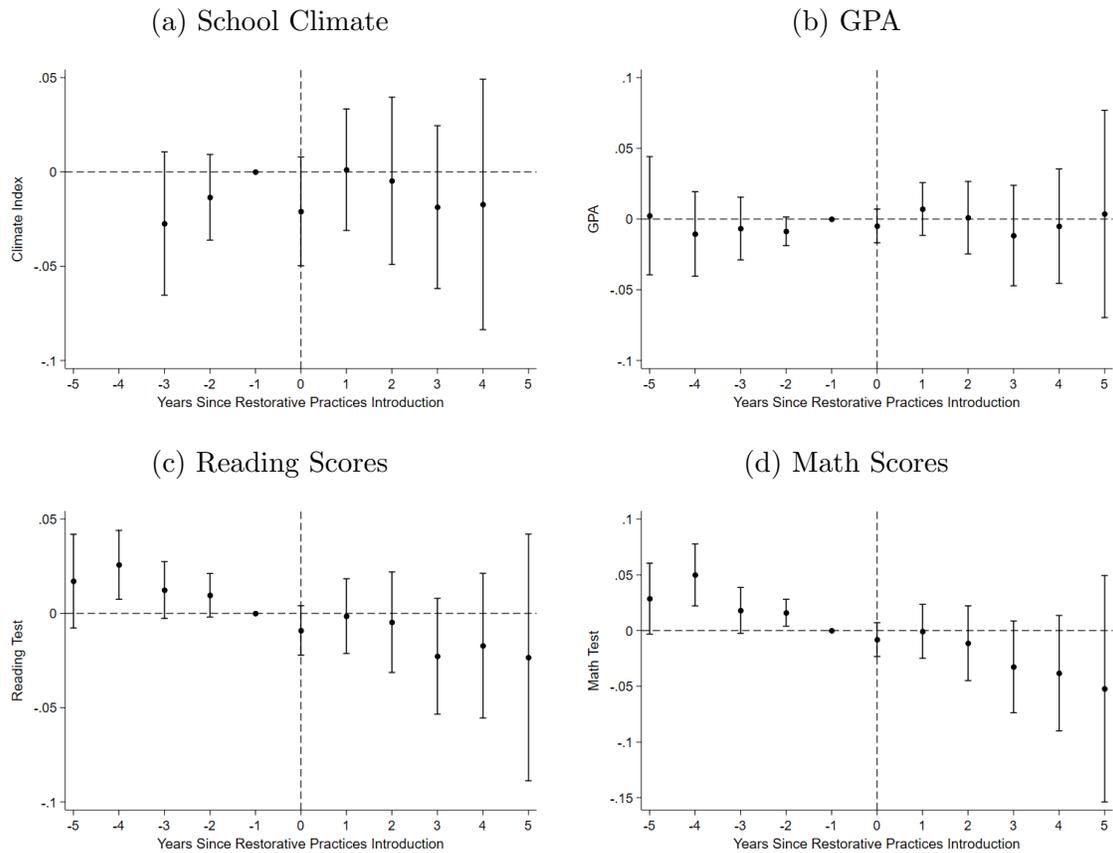
A Appendix Figures

Figure A1: Elementary School Event Studies: Behavioral Outcomes



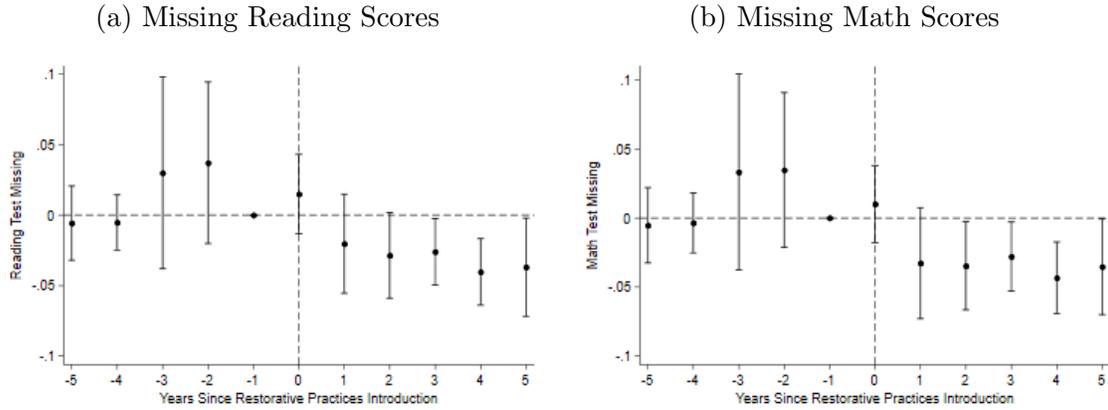
Notes: These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions, in-school suspensions, and absent days) and policing outcomes (overall arrests) over time in elementary schools. Observations are at the student-school year level. Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. Suspension and absence data are collected by Chicago Public Schools. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. An in-school suspension is defined as the removal of a student from their regular educational schedule for more than 60 minutes of the school day to an alternative supervised setting inside the school building. The absent days outcome is adjusted to equal total absent days minus out-of-school suspension days. Arrest data are collected by the Chicago Police Department. The arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Regressions for the absent days outcome include student member days in the corresponding school year as a control. Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

Figure A2: Elementary School Event Studies: School Climate and Learning



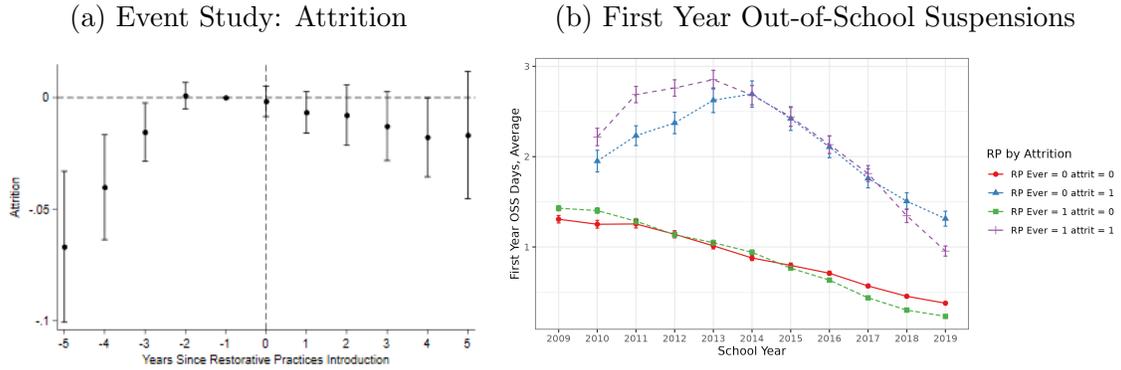
Notes: These figures show the event studies around the introduction of RP on students' perceptions of school climate and academic outcomes (GPA, reading test score, and math test score) over time in elementary schools. Observations are at the student-school year level. Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. Data for the school climate index begin two years after and ends one year before the data for the other outcome variables. Its graph therefore reflects one fewer estimated dynamic effect and two fewer placebo effects. GPA is calculated using semester final grades. Math and reading scores are standardized by test, school year, and grade. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

Figure A3: High School Event Studies: Test Score Missingness



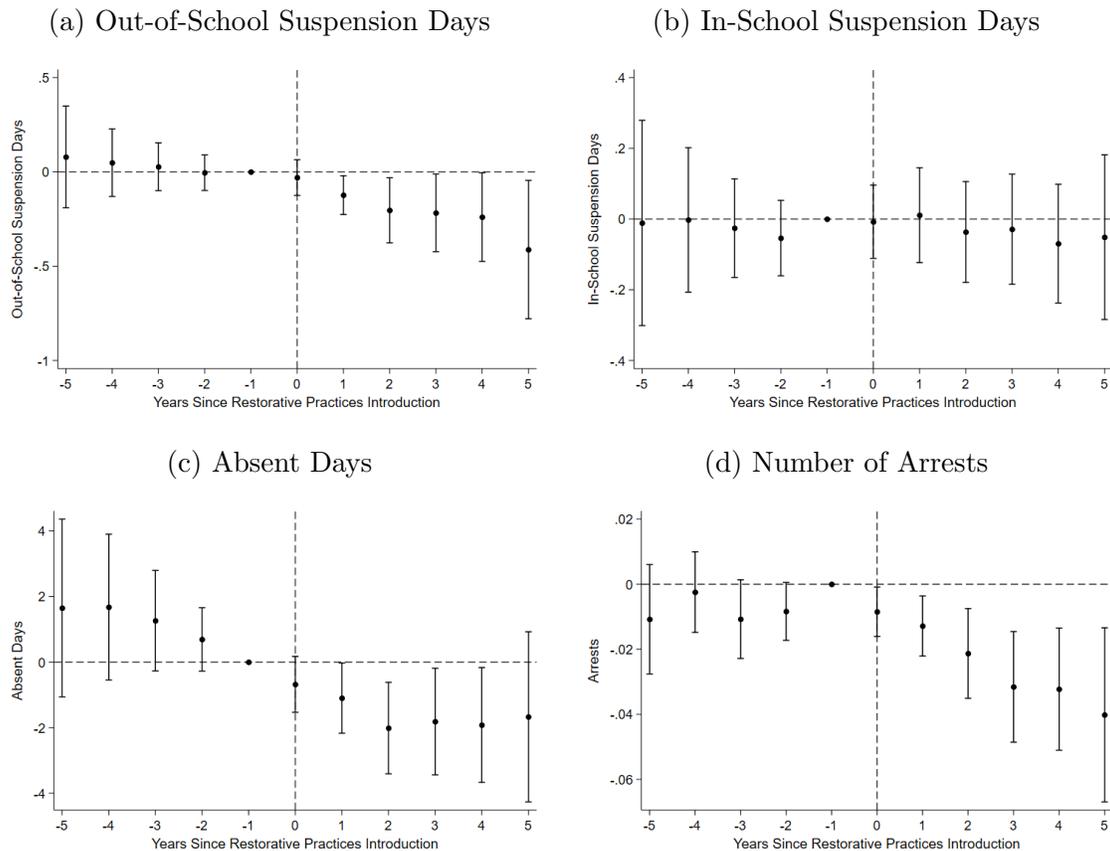
Notes: These figures show the event studies around the introduction of RP on the missingness of reading and math scores over time. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09. The sample includes all students in tested grade-cohort cells (students observed between SY09 and SY19 at any grade level between 9 and 11, except for 10th graders in SY15-SY16). Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

Figure A4: High School: Attrition Analysis



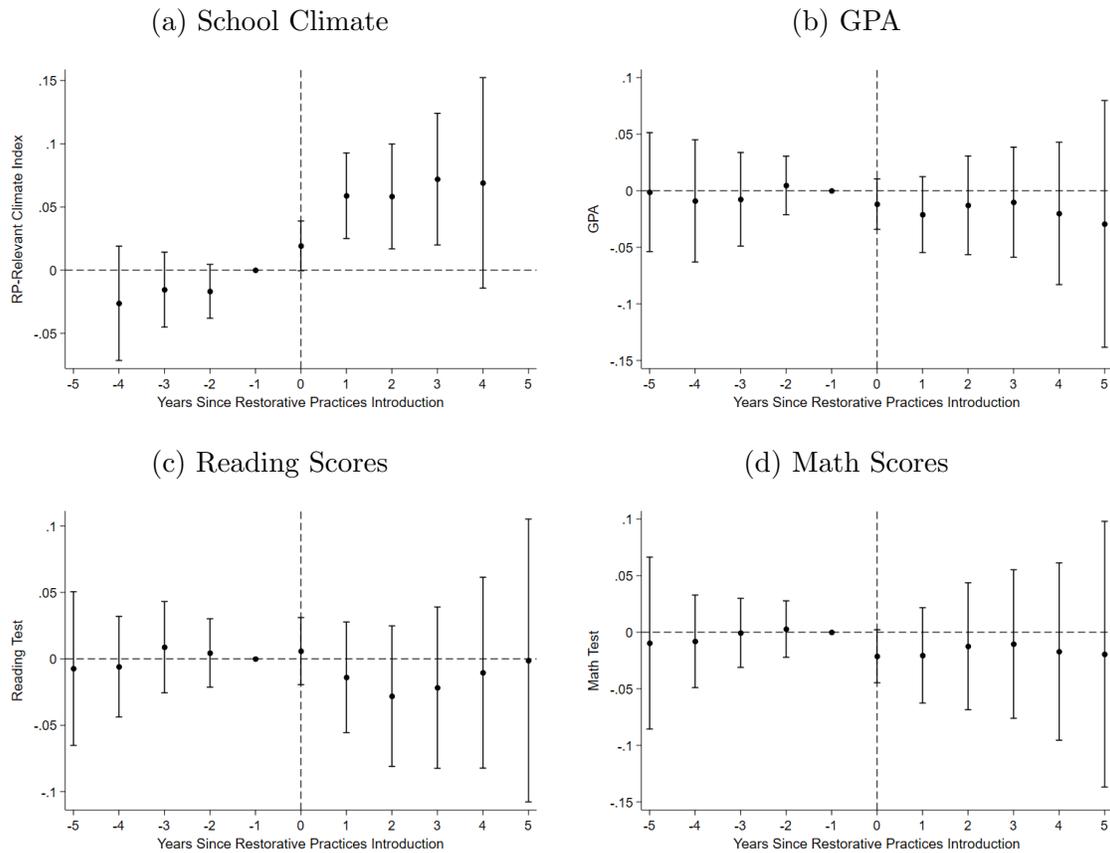
Notes: Panel A presents the event study for student attrition and panel B presents the average out-of-school suspension (OSS) days in a student’s first high school year by school treatment status and attrition status. The sample covers students in grades 9 to 12 between SY09 and SY19. To produce these plots, an artificial panel was first constructed that includes one observation for each student in each grade level between nine and twelve under the assumption that students progressed one grade level each year. For those student-grade observations that do not appear in our study sample due to student attrition, we code an attrition indicator variable equal to one. In panel B, the “attrit” variable for a student is coded as one if their attrition indicator ever equals one; the “RP Ever” variable for a student is coded as one if their initial high school ever adopted RP. Panel A includes the following covariates: student cohort fixed effects (based on grade and school year of entry), gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates in panel A are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. In each panel, bars represent 95% confidence intervals based on standard errors clustered by school.

Figure A5: High School Event Studies: Difference-in-Differences: Behavioral and Policing Outcomes



Notes: These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions, in-school suspensions, and absent days) and policing outcomes (overall arrests) over time in high schools. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Suspension and absence data are collected by Chicago Public Schools. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. An in-school suspension is defined as the removal of a student from their regular educational schedule for more than 60 minutes of the school day to an alternative supervised setting inside the school building. The absent days outcome is adjusted to equal total absent days minus out-of-school suspension days. Arrest data are collected by the Chicago Police Department. The arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Regressions for the absent days outcome include student member days in the corresponding school year as a control. Estimates are based on standard difference-in-differences models with entry school and school year fixed effects. Bars represent 95% confidence intervals based on standard errors clustered by school.

Figure A6: High School Event Studies: Difference-in-Differences: School Climate and Learning

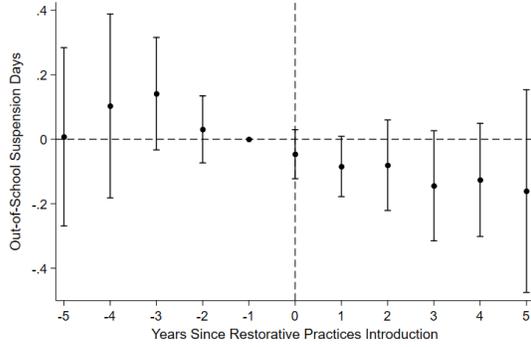


Notes: These figures show the event studies around the introduction of RP on students' perceptions of school climate and academic outcomes (GPA, reading test score, and math test score) over time in high schools. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. Data for the school climate index begin two years after and ends one year before the data for the other outcome variables. Its graph therefore reflects one fewer estimated dynamic effect and two fewer placebo effects. GPA is calculated using semester final grades. Math and reading scores are standardized by test, school year, and grade. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on standard difference-in-differences models with entry school and school year fixed effects. Bars represent 95% confidence intervals based on standard errors clustered by school.

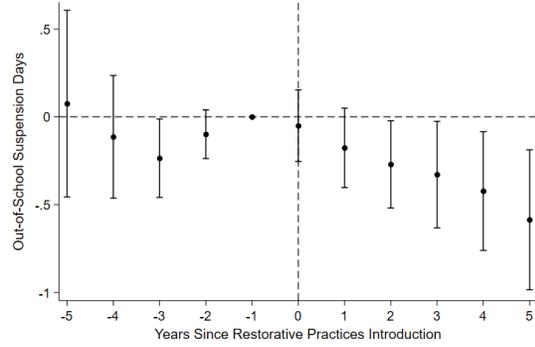
Figure A7: High School OSS Event Studies: Above- and Below-Median Predicted OSS

Panel A. All Students

(a) Below-Median Predicted OSS for Classmates

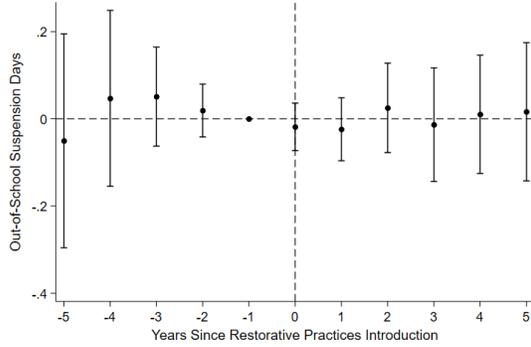


(b) Above-Median Predicted OSS for Classmates

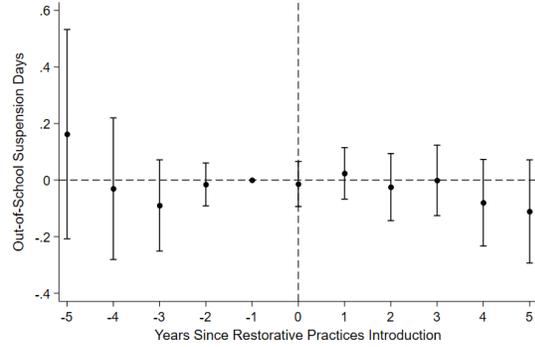


Panel B. Students Predicted to have Low-Suspension Propensity

(c) Below-Median Predicted OSS for Classmates



(d) Above-Median Predicted OSS for Classmates

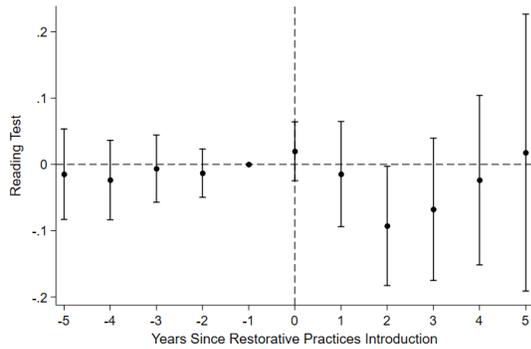


Notes: These figures show the event studies around the introduction of RP on out-of-school suspensions over time for all students and students with below-median predicted suspension days, by whether they belong to school-by-cohort cells that are above- versus below-median in predicted suspension days within a given cohort. Predictions for out-of-school suspension days for each student are constructed using a random forest algorithm as described in the text in Section VII. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

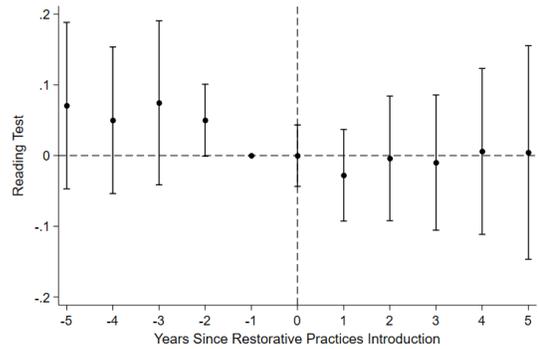
Figure A8: High School Event Studies: Academic Outcomes: Low Suspension Propensity Students

Panel A: Reading Scores

(a) Below-Median Predicted OSS for Classmates

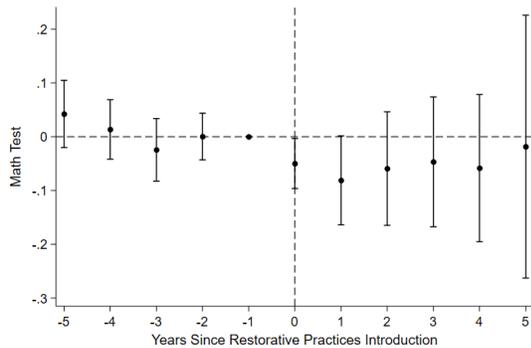


(b) Above-Median Predicted OSS for Classmates

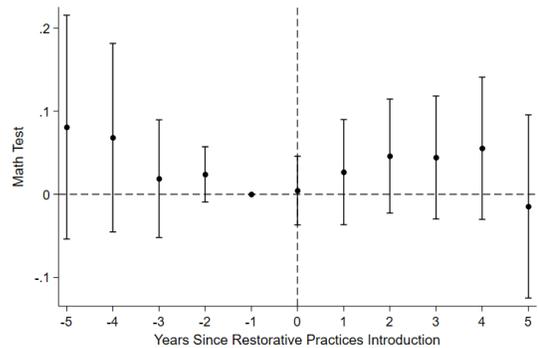


Panel B: Math Scores

(c) Below-Median Predicted OSS for Classmates



(d) Above-Median Predicted OSS for Classmates

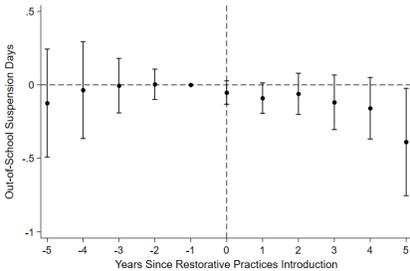


Notes: These figures show the event studies around the introduction of RP on student test scores over time for students with below-median predicted suspension days belonging to school-by-cohort cells that are above- versus below-median in predicted suspension days within a given cohort. Predictions for out-of-school suspension days for each student are constructed using a random forest algorithm as described in the text in Section VII. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

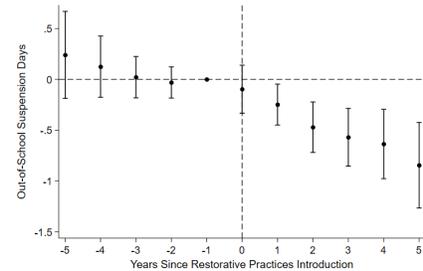
Figure A9: High School Event Studies: Treatment Heterogeneity by Race and School Share of Black Students: All Students

Panel A: Out-of-School Suspensions

(a) Below-Median Black Student Share

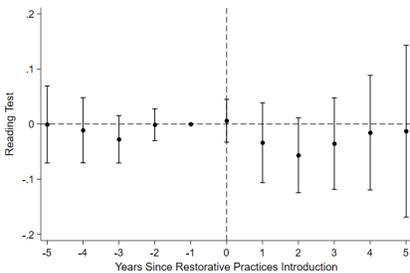


(b) Above-Median Black Student Share

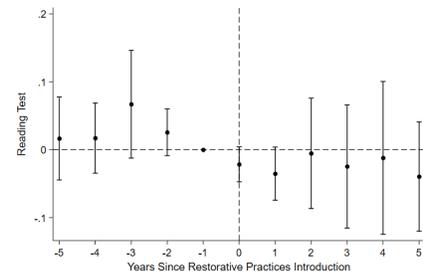


Panel B: Reading Scores

(c) Below-Median Black Student Share

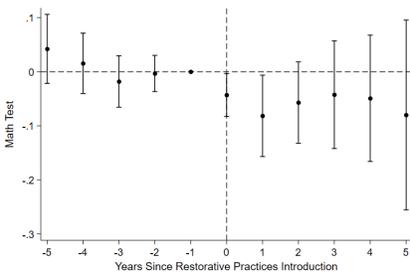


(d) Above-Median Black Student Share

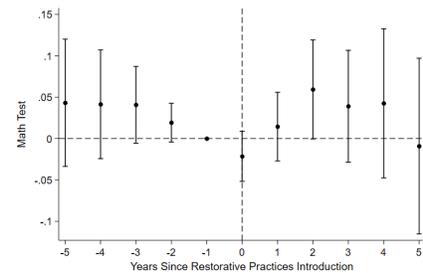


Panel C: Math Scores

(e) Below-Median Black Student Share



(f) Above-Median Black Student Share



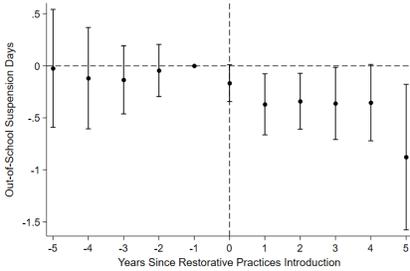
Notes: These figures show the event studies around the introduction of RP on out-of-school suspensions and test scores over time for students assigned to schools that are classified as having an above- or below-median share of Black students relative to the population of CPS schools based on SY13 enrollment and the median student's school-level share in SY13.

Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

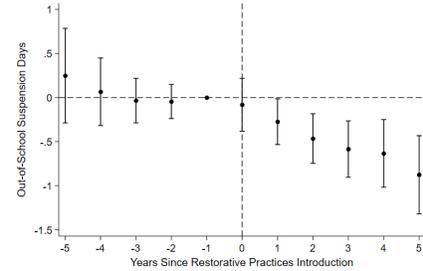
Figure A10: High School Event Studies: Treatment Heterogeneity by Race and School Share of Black Students: Black Students

Panel A: Out-of-School Suspensions

(a) Below-Median Black Student Share

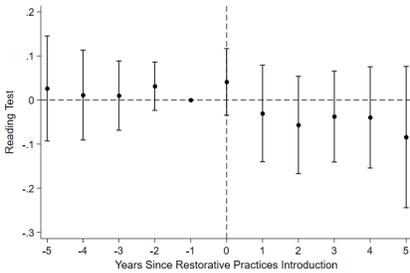


(b) Above-Median Black Student Share

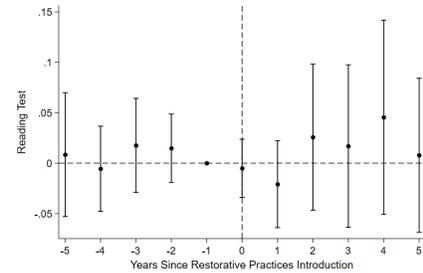


Panel B: Reading Scores

(c) Below-Median Black Student Share

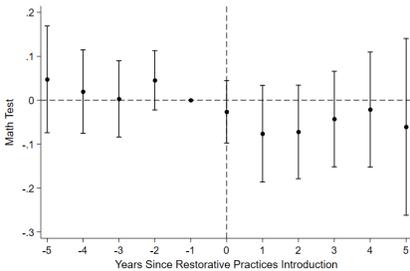


(d) Above-Median Black Student Share

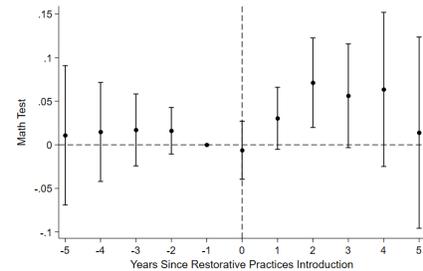


Panel C: Math Scores

(e) Below-Median Black Student Share



(f) Above-Median Black Student Share

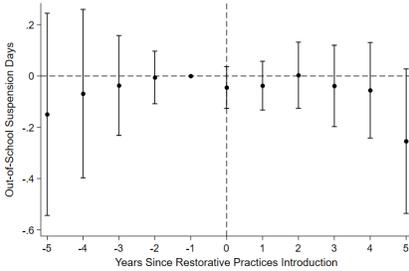


Notes: These figures show the event studies around the introduction of RP on out-of-school suspensions and test scores over time for Black students assigned to schools that are classified as having an above- or below-median share of Black students relative to the population of CPS schools based on SY13 enrollment and the median student's school-level share in SY13. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

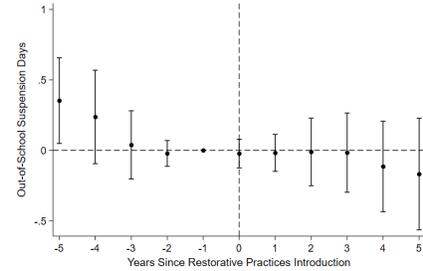
Figure A11: High School Event Studies: Treatment Heterogeneity by Race and School Share of Black Students: Latine Students

Panel A: Out-of-School Suspensions

(a) Below-Median Black Student Share

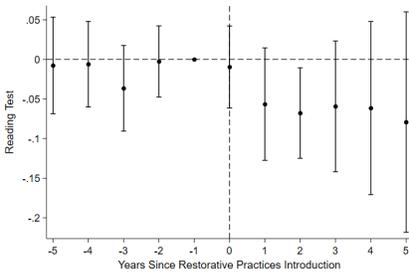


(b) Above-Median Black Student Share

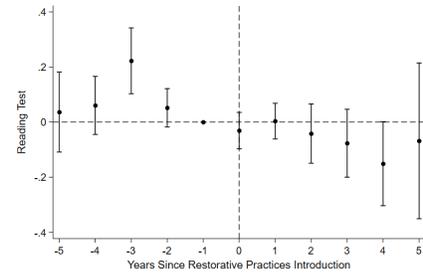


Panel B: Reading Scores

(c) Below-Median Black Student Share

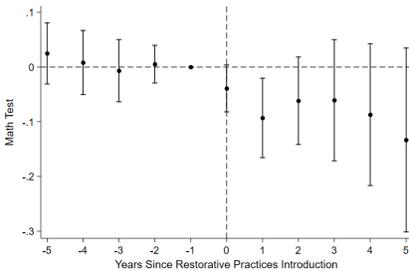


(d) Above-Median Black Student Share

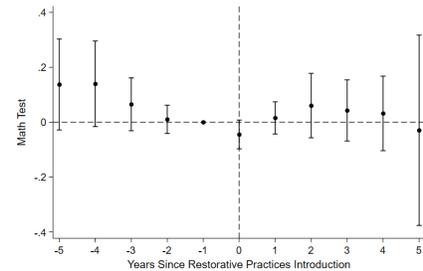


Panel C: Math Scores

(e) Below-Median Black Student Share



(f) Above-Median Black Student Share



Notes: These figures show the event studies around the introduction of RP on out-of-school suspensions and test scores over time for Latine students assigned to schools that are classified as having an above- or below-median share of Black students relative to the population of CPS schools based on SY13 enrollment and the median student's school-level share in SY13. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Bars represent 95% confidence intervals based on standard errors clustered by school.

B Appendix Tables

Table A1: Number of High Schools by Initial Restorative Practices Program Type

| RP Status | Total | District | Charter |
|--|-------|----------|---------|
| Schools whose first RP received was Coaching | 60 | 56 | 4 |
| First RP Received is Coaching in SY14 | 18 | 18 | 0 |
| First RP Received is Coaching in SY15 | 23 | 23 | 0 |
| First RP Received is Coaching in SY16 | 5 | 5 | 0 |
| First RP Received is Coaching in SY17 | 3 | 3 | 0 |
| First RP Received is Coaching in SY18 | 7 | 4 | 3 |
| First RP Received is Coaching in SY19 | 4 | 3 | 1 |
| Schools whose first RP received was Leadership | 4 | 4 | 0 |
| First RP Received is Leadership in SY17 | 1 | 1 | 0 |
| First RP Received is Leadership in SY18 | 3 | 3 | 0 |
| Schools whose first RP received was Peer Council | 8 | 8 | 0 |
| First RP Received is Peer Council in SY14 | 4 | 4 | 0 |
| First RP Received is Peer Council in SY19 | 4 | 4 | 0 |

Notes: This table presents summary statistics on the number of high schools by first RP type by school year. RP programming data covers SY14-SY19. If schools received multiple RP programs in the same initial year, this table only counts the most intensive of the RP programs in which they participated (with Coaching being the most intensive and Peer Council being the least).

Table A2: High School Student Baseline Characteristics by English Language Learner Status, Grade Level, and Disability Status

| | Out-of-School Suspension Days (1) | Absent Days (2) | Number of Arrests (3) | School Climate (4) | Reading Scores (5) | Math Scores (6) |
|--------------------------------|--|-----------------------|-----------------------------|--------------------------|--------------------------|-----------------------|
| <i>Native English Speakers</i> | | | | | | |
| Baseline Mean | 0.965 | 18.452 | 0.132 | 0.001 | 0.050 | 0.043 |
| Observations | 96,850 | 96,850 | 99,508 | 65,214 | 67,142 | 67,223 |
| <i>English Learners</i> | | | | | | |
| Baseline Mean | 0.561 | 17.603 | 0.058 | -0.020 | -0.792 | -0.680 |
| Observations | 6,287 | 6,287 | 6,298 | 4,037 | 4,204 | 4,236 |
| <i>Grades 9 and 10</i> | | | | | | |
| Baseline Mean | 1.140 | 16.807 | 0.153 | 0.000 | 0.000 | 0.000 |
| Observations | 56,654 | 56,654 | 58,988 | 39,663 | 49,264 | 49,347 |
| <i>Grades 11 and 12</i> | | | | | | |
| Baseline Mean | 0.696 | 20.343 | 0.096 | 0.000 | 0.000 | 0.000 |
| Observations | 46,483 | 46,483 | 46,818 | 29,588 | 22,082 | 22,112 |
| <i>Any Disability</i> | | | | | | |
| Baseline Mean | 1.049 | 21.976 | 0.195 | -0.019 | -0.514 | -0.649 |
| Observations | 16,655 | 16,655 | 17,181 | 9,930 | 10,920 | 10,964 |
| <i>504 Disability</i> | | | | | | |
| Baseline Mean | 0.728 | 19.982 | 0.065 | 0.036 | 0.446 | 0.357 |
| Observations | 2,827 | 2,827 | 2,901 | 1,960 | 2,195 | 2,197 |
| <i>Physical Disability</i> | | | | | | |
| Baseline Mean | 0.885 | 24.140 | 0.190 | 0.056 | -0.334 | -0.485 |
| Observations | 1,089 | 1,089 | 1,153 | 553 | 639 | 642 |
| <i>Cognitive Disability</i> | | | | | | |
| Baseline Mean | 1.134 | 22.234 | 0.224 | -0.040 | -0.789 | -0.934 |
| Observations | 12,739 | 12,739 | 13,127 | 7,417 | 8,086 | 8,125 |
| <i>No Disability</i> | | | | | | |
| Baseline Mean | 0.922 | 17.750 | 0.114 | 0.003 | 0.093 | 0.118 |
| Observations | 86,221 | 86,221 | 88,241 | 59,308 | 60,421 | 60,490 |

Notes: This table displays the mean for each outcome variable in SY13 for students at any grade level between 9 and 12, disaggregated by English Language Learner (ELL) status, grade grouping, and disability status. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. The absent days outcome is adjusted to equal total absent days minus out-of-school suspension days. Arrest data are collected by the Chicago Police Department. The arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. Math and reading scores are standardized by test, school year, and grade.

Table A3: High School Student Baseline Characteristics by Race and Gender

| | Out-of-School Suspension Days (1) | Absent Days (2) | Number of Arrests (3) | School Climate (4) | Reading Scores (5) | Math Scores (6) |
|----------------------|--|-----------------------|-----------------------------|--------------------------|--------------------------|-----------------------|
| <i>Black Female</i> | | | | | | |
| Baseline Mean | 1.322 | 21.814 | 0.085 | -0.107 | -0.081 | -0.233 |
| Observations | 23,998 | 23,998 | 24,550 | 15,754 | 16,200 | 16,232 |
| <i>White Female</i> | | | | | | |
| Baseline Mean | 0.180 | 15.650 | 0.013 | 0.163 | 0.941 | 0.752 |
| Observations | 4,899 | 4,899 | 5,098 | 3,573 | 3,516 | 3,515 |
| <i>Latina Female</i> | | | | | | |
| Baseline Mean | 0.304 | 17.084 | 0.015 | 0.030 | 0.027 | 0.005 |
| Observations | 21,687 | 21,687 | 22,053 | 15,668 | 15,749 | 15,758 |
| <i>Black Male</i> | | | | | | |
| Baseline Mean | 1.880 | 21.010 | 0.387 | -0.048 | -0.341 | -0.334 |
| Observations | 21,850 | 21,850 | 22,581 | 13,079 | 14,208 | 14,258 |
| <i>White Male</i> | | | | | | |
| Baseline Mean | 0.417 | 15.370 | 0.059 | 0.137 | 0.653 | 0.734 |
| Observations | 4,779 | 4,779 | 4,995 | 3,343 | 3,321 | 3,326 |
| <i>Latino Male</i> | | | | | | |
| Baseline Mean | 0.670 | 16.465 | 0.086 | 0.025 | -0.106 | 0.043 |
| Observations | 21,228 | 21,228 | 21,632 | 14,653 | 15,291 | 15,295 |

Notes: This table displays the mean for each outcome variable in SY13 for students at any grade level between 9 and 12, disaggregated by race and gender. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. The absent days outcome is adjusted to equal total absent days minus out-of-school suspension days. Arrest data are collected by the Chicago Police Department. The arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. Math and reading scores are standardized by test, school year, and grade.

Table A4: Baseline Characteristics: Chicago Public Schools Elementary School Students

| Variable | Treated (1) | Non-Treated (2) | Difference (3) |
|---|-----------------|--------------------|-------------------|
| Number of Students | 557.16 (308.12) | 541.06 (319.29) | 16.10 (27.18) |
| Out-of-School Suspension Days | 0.44 (1.96) | 0.38 (1.78) | 0.06 (0.05) |
| In-School Suspension Days | 0.06 (0.44) | 0.05 (0.39) | 0.01 (0.01) |
| Absent Days | 8.81 (9.88) | 8.32 (9.26) | 0.49* (0.24) |
| Number of Arrests | 0.02 (0.22) | 0.02 (0.22) | 0.00 (0.00) |
| Ever Arrested | 0.01 (0.10) | 0.01 (0.10) | 0.00 (0.00) |
| GPA | 2.94 (0.70) | 2.99 (0.69) | -0.06 (0.04) |
| Math Scores | -0.03 (1.00) | 0.02 (1.00) | -0.05 (0.04) |
| Reading Scores | -0.03 (1.00) | 0.02 (1.00) | -0.06 (0.04) |
| School Climate Index | -0.02 (0.65) | 0.01 (0.66) | -0.02 (0.02) |
| English Learner | 0.14 (0.35) | 0.14 (0.34) | 0.00 (0.01) |
| Students in Temporary Living Situations | 0.04 (0.20) | 0.04 (0.19) | 0.01 (0.01) |
| Individualized Education Plan | 0.15 (0.36) | 0.14 (0.35) | 0.01 (0.01) |
| Economically Disadvantaged | 0.85 (0.36) | 0.85 (0.36) | 0.00 (0.02) |
| Gender: Female | 0.49 (0.50) | 0.49 (0.50) | 0.00 (0.00) |
| Race: Black | 0.44 (0.50) | 0.42 (0.49) | 0.02 (0.04) |
| Race: White | 0.11 (0.31) | 0.10 (0.30) | 0.00 (0.02) |
| Race: Latine | 0.42 (0.49) | 0.44 (0.50) | -0.01 (0.04) |
| Disability: Cognitive | 0.15 (0.35) | 0.14 (0.35) | 0.01 (0.00) |
| Disability: None | 0.79 (0.40) | 0.80 (0.40) | -0.01 (0.01) |
| Disability: Physical | 0.02 (0.13) | 0.02 (0.12) | 0.00 (0.00) |
| Disability: 504 | 0.04 (0.20) | 0.05 (0.21) | 0.00 (0.00) |
| Observations | 67,348 | 100,037 | |

Notes: This table presents student-level means in subsequently treated elementary schools (column 1) and non-treated elementary schools (column 2), with means constructed in SY13 (prior to the introduction of RP). The associated differences (column 3) are derived from student-level regressions of the given outcome on a treatment indicator variable, with the standard errors clustered at the school level. Absent Days is defined as the total number of days absent, minus the total number of out-of-school suspension days that a student had in the school year, regardless of school. Arrest data are collected by the Chicago Police Department. GPA is calculated using semester final grades. Math and reading scores are standardized by test, school year, and grade. The School Climate Index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. See Data Appendix C for detailed variable definitions. Standard errors are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A5: Elementary School Restorative Practices: In-School Behavioral Outcomes

| | Out-of-School Suspension | | In-School Suspension | | Absent Days |
|---------------|--------------------------|---------------------|----------------------|------------------|-------------------|
| | Days (1) | Binary (2) | Days (3) | Binary (4) | (5) |
| RP | -0.050** (0.019) | -0.010** (0.004) | 0.007 (0.005) | 0.003 (0.002) | -0.077 (0.095) |
| Baseline Mean | 0.401 | 0.098 | 0.054 | 0.029 | 8.497 |
| Observations | 2,536,517 | 2,536,517 | 2,536,517 | 2,536,517 | 2,536,517 |

Notes: Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. In columns 1 and 3, the out-of-school suspension (OSS) days and in-school suspension (ISS) days outcomes are the total number of OSS or ISS days that the student received in the corresponding school year, regardless of the school. In columns 2 and 4, the OSS and ISS binary outcomes indicate whether a student ever received either of these types of suspensions in the corresponding school year, regardless of the school. Suspension data are collected by Chicago Public Schools. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. An in-school suspension is defined as the removal of a student from their regular educational schedule for more than 60 minutes of the school day to an alternative supervised setting inside the school building. In column 5, the absent days outcome is adjusted to equal total absent days minus out-of-school suspension days. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Regressions for the absent days outcome include student member days in the corresponding school year as a control. Estimates are based on the methodology developed in *de Chaisemartin and D'Haultfoeuille (2020)* and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A6: Elementary School Restorative Practices: Policing Outcomes

| | Number of Arrests Overall | Number of In-School Arrests | Number of Out-of-School Arrests | Number of Violent Arrests | Number of Non-Violent Arrests |
|---------------|---------------------------|-----------------------------|---------------------------------|---------------------------|-------------------------------|
| RP | -0.002 (0.001) | -0.001 (0.000) | -0.001 (0.001) | 0.000 (0.001) | -0.002+ (0.001) |
| Baseline Mean | 0.017 | 0.005 | 0.012 | 0.006 | 0.011 |
| Observations | 2,546,569 | 2,546,569 | 2,546,569 | 2,546,569 | 2,546,569 |

Notes: Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. Arrest data are collected by the Chicago Police Department. The arrest data includes information on the type (violent or non-violent), the location, and the time of arrest. The main arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. In-school arrests are defined as incidents that happened both inside the school location and during school hours, and out-of-school arrests are defined as incidents that happened either outside the school location or outside school hours. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A7: High School Restorative Practices: Policing Outcomes (Binary)

| | Arrests Overall (Binary) (1) | In-School Arrests (Binary) (2) | Out-of-School Arrests (Binary) (3) | Violent Arrests (Binary) (4) | Non-Violent Arrests (Binary) (5) |
|---------------|---------------------------------------|---|---|---------------------------------------|---|
| RP | -0.009** (0.003) | -0.006** (0.002) | -0.006* (0.002) | -0.003* (0.001) | -0.009** (0.002) |
| Baseline Mean | 0.071 | 0.022 | 0.056 | 0.023 | 0.058 |
| Observations | 1,380,959 | 1,380,959 | 1,380,959 | 1,380,959 | 1,380,959 |

Notes: Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Arrest data are collected by the Chicago Police Department. The arrest data includes information on the type (violent or non-violent), the location, and the time of arrest. The main arrest outcome is defined as whether a student was arrested in a given year, regardless of the type of arrest or the location of the arrest. In-school arrests are defined as incidents that happened both inside the school location and during school hours, and out-of-school arrests are defined as incidents that happened either outside the school location or outside school hours. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A8: High School MVMS Constructs

| | RP | Observations |
|---|----------------|--------------|
| <u>School Climate Index Components</u> | | |
| Emotional Health | 0.013 (0.016) | 742,979 |
| Student Classroom Behavior | 0.054*(0.022) | 744,023 |
| Academic Personalism | 0.028 (0.018) | 723,840 |
| Psychological Sense of School Membership | 0.050**(0.019) | 745,163 |
| Safety | 0.036+ (0.021) | 714,731 |
| School-Wide Future Orientation | 0.047*(0.023) | 714,173 |
| School Safety | 0.064*(0.026) | 750,261 |
| Student Teacher Trust | 0.034+ (0.021) | 738,723 |
| <u>Placebo Constructs</u> | | |
| Human and Social Resources in Community | 0.001 (0.019) | 719,136 |
| Parent Supportiveness | -0.009 (0.015) | 736,676 |

Notes: This table displays constructs included in the MVMS student survey. All constructs are corrected to fit a positively valenced metric, where higher scores are better. Observations are at the student-school year level, and we report the average effect of restorative practices over five periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY11 and SY18. Each MVMS construct is standardized by year and grade. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A9: Elementary School Restorative Practices: School Climate and Learning Outcomes

| | School Climate | GPA | Reading Scores | Math Scores |
|---------------|-------------------|-------------------|-------------------|-------------------|
| RP | -0.011 (0.015) | -0.002 (0.011) | -0.010 (0.011) | -0.016 (0.014) |
| Baseline Mean | -0.002 | 2.968 | 0.000 | 0.000 |
| Observations | 740,741 | 2,154,479 | 2,381,814 | 2,391,933 |

Notes: Observations are at the student-school year level, and we report the average effect of restorative practices over six periods (five periods for the school climate index). Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. GPA is calculated using semester final grades. Math and reading scores are standardized by test, school year, and grade. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A10: High School Behavioral and Policing Outcomes by Implementation Type

| | Out-of-School Suspension Days (1) | In-School Suspension Days (2) | Absent Days (3) | Number of Arrests (4) |
|------------------------|---|-------------------------------------|--------------------|-----------------------------|
| <i>RP Coaching</i> | -0.177* (0.076) | -0.019 (0.075) | -0.561 (0.483) | -0.025** (0.007) |
| Baseline Mean | 0.987 | 0.429 | 18.713 | 0.134 |
| Observations | 1,236,747 | 1,236,747 | 1,236,747 | 1,259,492 |
| <i>RP Leadership</i> | 0.024 (0.170) | -0.023 (0.110) | 0.970 (2.190) | 0.000 (0.058) |
| Baseline Mean | 0.831 | 0.364 | 15.428 | 0.106 |
| Observations | 251,770 | 251,770 | 251,770 | 256,423 |
| <i>RP Peer Council</i> | -0.114 (0.129) | -0.085 (0.176) | -0.777 (1.361) | -0.012 (0.018) |
| Baseline Mean | 0.904 | 0.370 | 16.412 | 0.114 |
| Observations | 331,002 | 331,002 | 331,002 | 336,292 |

Notes: This table shows results by the first type of restorative practices (RP) programming that was implemented in a student's high school. If schools received multiple RP programs in the same initial year, they are classified based on the most intensive of the RP programs in which they participated (with Coaching being the most intensive and Peer Council being the least). Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A11: High School Climate and Learning Outcomes by Implementation Type

| | School Climate (1) | GPA (2) | Reading Scores (3) | Math Scores (4) |
|------------------------|-----------------------|-------------------|-----------------------|--------------------|
| <i>RP Coaching</i> | 0.036+ (0.018) | -0.028 (0.024) | 0.002 (0.025) | 0.008 (0.024) |
| Baseline Mean | -0.016 | 2.416 | -0.042 | -0.042 |
| Observations | 671,492 | 783,423 | 763,506 | 756,681 |
| <i>RP Leadership</i> | 0.029 (0.041) | -0.033 (0.086) | 0.036 (0.098) | 0.000 (0.042) |
| Baseline Mean | 0.109 | 2.601 | 0.201 | 0.211 |
| Observations | 109,684 | 144,087 | 75,471 | 75,502 |
| <i>RP Peer Council</i> | 0.044 (0.048) | -0.069 (0.050) | -0.015 (0.039) | -0.096+ (0.052) |
| Baseline Mean | 0.071 | 2.550 | 0.060 | 0.088 |
| Observations | 171,884 | 199,365 | 203,201 | 200,919 |

Notes: This table shows results by the first type of restorative practices (RP) programming that was implemented in a student's high school. If schools received multiple RP programs in the same initial year, they are classified based on the most intensive of the RP programs in which they participated (with Coaching being the most intensive and Peer Council being the least). Observations are at the student-school year level, and we report the average effect of restorative practices over six periods (five periods for the school climate index). Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A12: Robustness: Behavioral Outcomes

| | Out-of-School Suspension Days (1) | In-School Suspension Days (2) | Absent Days (3) | Number of Arrests (4) |
|---|--|--|---------------------|-----------------------------|
| <i>Panel A: Difference-in-Differences</i> | | | | |
| RP | -0.154+ (0.087) | 0.001 (0.07) | -2.133** (0.772) | -0.012* (0.006) |
| Baseline Mean | 0.940 | 0.413 | 18.401 | 0.128 |
| Observations | 1,126,004 | 1,126,004 | 1,126,004 | 1,173,658 |
| <i>Panel B: Dropping Charter and Contract School Students</i> | | | | |
| RP | -0.155* (0.076) | -0.005 (0.091) | -0.213 (0.534) | -0.023** (0.008) |
| Baseline Mean | 0.946 | 0.469 | 19.904 | 0.129 |
| Observations | 873,906 | 873,906 | 873,906 | 893,546 |
| <i>Panel C: Only Age and Cohort Fixed Effects as Controls</i> | | | | |
| RP | -0.156* (0.068) | -0.001 (0.059) | -0.328 (0.448) | -0.025** (0.006) |
| Baseline Mean | 0.940 | 0.413 | 18.401 | 0.128 |
| Observations | 1,359,296 | 1,359,296 | 1,359,296 | 1,383,914 |

Notes: Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Estimates in panel A are based on standard difference-in-differences models with entry school and school year fixed effects. Estimates in panels B and C are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020), and we report the average effect of restorative practices over six periods. Panel B restricts the sample to students who remained in traditional district-run schools in each school year and so excludes all observations for students who ever attended a charter school or contract school in that year. See Data Appendix C for detailed variable definitions. Each specification in panels A and B includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A13: Robustness: School Climate and Learning

| | School Climate | GPA | Reading Scores | Math Scores |
|---|--------------------|--------------------|-------------------|-------------------|
| <i>Panel A: Difference-in-Differences</i> | | | | |
| RP | 0.054** (0.017) | -0.013 (0.021) | -0.011 (0.022) | -0.015 (0.022) |
| Baseline Mean | 0.000 | 2.457 | 0.000 | 0.000 |
| Observations | 578,657 | 893,520 | 722,012 | 718,745 |
| <i>Panel B: Dropping Charter and Contract School Students</i> | | | | |
| RP | 0.032 (0.019) | -0.024 (0.021) | -0.031 (0.024) | -0.030 (0.027) |
| Baseline Mean | -0.019 | 2.473 | 0.052 | 0.041 |
| Observations | 510,617 | 851,492 | 524,886 | 519,040 |
| <i>Panel C: Only Age and Cohort Fixed Effects as Controls</i> | | | | |
| RP | 0.039* (0.017) | -0.038+ (0.021) | -0.005 (0.027) | -0.006 (0.027) |
| Baseline Mean | 0.000 | 2.457 | 0.000 | 0.000 |
| Observations | 752,644 | 899,222 | 833,502 | 825,849 |

Notes: Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. Estimates in panel A are based on standard difference-in-differences models with entry school and school year fixed effects. Estimates in panels B and C are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020), and we report the average effect of restorative practices over six periods (five periods for the school climate index). Panel B restricts the sample to students who remained in traditional district-run schools in each school year and so excludes all observations for students who ever attended a charter school or contract school in that year. See Data Appendix C for detailed variable definitions. Each specification in panels A and B includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A14: High School Baseline Outcomes by Predicted Out-of-School Suspension Days Cell (Below- vs. Above-Median)

| Outcome | <i>Below-Median OSS Days</i> | | <i>Above-Median OSS Days</i> | | Difference (5) |
|-------------------------------|------------------------------|---------------------|------------------------------|---------------------|-------------------|
| | Baseline Mean (1) | Observations (2) | Baseline Mean (3) | Observations (4) | |
| Math Score | 0.319 (1.035) | 38,103 | -0.363 (0.819) | 33,272 | 0.681**(0.118) |
| Reading Score | 0.283 (1.060) | 38,080 | -0.324 (0.814) | 33,182 | 0.607**(0.126) |
| GPA | 2.643 (0.958) | 44,245 | 2.228 (0.963) | 35,735 | 0.415**(0.071) |
| School Climate Index | 0.106 (0.639) | 37,010 | -0.122 (0.625) | 32,195 | 0.228**(0.041) |
| Out-of-School Suspension Days | 0.439 (1.917) | 53,008 | 1.468 (3.795) | 50,009 | -1.029**(0.105) |
| In-School Suspension Days | 0.143 (0.870) | 53,008 | 0.697 (2.088) | 50,009 | -0.555**(0.076) |
| Number of Arrests | 0.049 (0.356) | 53,896 | 0.208 (0.802) | 51,708 | -0.160**(0.017) |

Notes: This table displays the mean for each outcome variable in SY13 for students at any grade level between 9 and 12, disaggregated by whether students belong to school-by-cohort cells that are above- versus below-median in predicted suspension days within a given cohort. Predictions for out-of-school suspension days for each student are constructed using a random forest algorithm as described in the text in Section VII. The associated differences (column 5) are derived from student-level regressions of the given outcome on an indicator for a student belonging to a below-median school-by-cohort cell, with the standard errors clustered at the school level. Math and reading scores are standardized by test, school year, and grade. GPA is calculated using semester final grades. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. An in-school suspension is defined as the removal of a student from their regular educational schedule for more than 60 minutes of the school day to an alternative supervised setting inside the school building. Arrest data are collected by the Chicago Police Department. The arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. See Data Appendix C for detailed variable definitions. Standard errors are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A15: High School Restorative Practices: Behavioral and Policing Outcomes by English Learner Status

| | Out-of-School Suspension Days (1) | In-School Suspension Days (2) | Absent Days (3) | Number of Arrests (4) |
|--------------------------------|--|--|-----------------------|-----------------------------|
| <i>Native English Speakers</i> | -0.183* (0.073) | -0.033 (0.065) | -0.575 (0.479) | -0.025** (0.007) |
| Baseline Mean | 0.965 | 0.424 | 18.452 | 0.132 |
| Observations | 1,263,729 | 1,263,729 | 1,263,729 | 1,288,130 |
| <i>English Learners</i> | 0.094 (0.099) | 0.198* (0.096) | -0.115 (0.728) | -0.021+ (0.012) |
| Baseline Mean | 0.561 | 0.236 | 17.603 | 0.058 |
| Observations | 91,658 | 91,658 | 91,658 | 91,692 |

Notes: This table shows results by English Language Learner status. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). The baseline mean for each outcome reflects the SY13 mean value. Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A16: High School Restorative Practices: School Climate and Learning Outcomes by English Learner Status

| | School Climate | GPA | Reading Scores | Math Scores |
|--------------------------------|--------------------|-------------------|-------------------|-------------------|
| <i>Native English Speakers</i> | 0.045** (0.017) | -0.025 (0.020) | -0.007 (0.022) | -0.012 (0.024) |
| Baseline Mean | 0.001 | 2.463 | 0.050 | 0.043 |
| Observations | 701,961 | 836,752 | 769,045 | 760,364 |
| <i>English Learners</i> | 0.012 (0.030) | -0.037 (0.038) | 0.049+ (0.026) | 0.038 (0.037) |
| Baseline Mean | -0.020 | 2.377 | -0.792 | -0.680 |
| Observations | 49,124 | 59,431 | 61,638 | 62,640 |

Notes: This table shows results by English Language Learner status. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods (five periods for the school climate index). Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). The baseline mean for each outcome reflects the SY13 mean value. Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A17: High School Restorative Practices: Behavioral and Policing Outcomes by Grade Levels

| | Out-of-School Suspension Days (1) | In-School Suspension Days (2) | Absent Days (3) | Number of Arrests (4) |
|-------------------------|--|--|-----------------------|-----------------------------|
| <i>Grades 9 and 10</i> | -0.219* (0.087) | -0.027 (0.075) | -0.759+ (0.400) | -0.032** (0.009) |
| Baseline Mean | 1.140 | 0.490 | 16.807 | 0.153 |
| Observations | 724,191 | 724,191 | 724,191 | 744,536 |
| <i>Grades 11 and 12</i> | -0.121* (0.055) | 0.021 (0.047) | -0.264 (0.679) | -0.009+ (0.005) |
| Baseline Mean | 0.696 | 0.319 | 20.343 | 0.096 |
| Observations | 600,608 | 600,608 | 600,608 | 604,371 |

Notes: This table shows results by grade level. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). The baseline mean for each outcome reflects the SY13 mean value. Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A18: High School Restorative Practices: School Climate and Learning Outcomes by Grade Levels

| | School Climate | GPA | Reading Scores | Math Scores |
|-------------------------|-------------------|--------------------|-------------------|-------------------|
| <i>Grades 9 and 10</i> | 0.034 (0.020) | -0.013 (0.025) | 0.001 (0.026) | 0.001 (0.026) |
| Baseline Mean | 0.000 | 2.393 | 0.000 | 0.000 |
| Observations | 409,076 | 469,662 | 529,028 | 521,261 |
| <i>Grades 11 and 12</i> | 0.048* (0.021) | -0.052* (0.026) | -0.033 (0.021) | -0.011 (0.024) |
| Baseline Mean | 0.000 | 2.531 | 0.000 | 0.000 |
| Observations | 325,810 | 413,172 | 285,728 | 285,944 |

Notes: This table shows results by grade level. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods (five periods for the school climate index). Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). The baseline mean for each outcome reflects the SY13 mean value. Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A19: High School Restorative Practices: Behavioral and Policing Outcomes by Disability Status

| | Out-of-School Suspension Days (1) | In-School Suspension Days (2) | Absent Days (3) | Number of Arrests (4) |
|-----------------------------|--|--|-----------------------|-----------------------------|
| <i>Any Disability</i> | -0.152* (0.071) | -0.016 (0.093) | -0.699 (0.466) | -0.050** (0.012) |
| Baseline Mean | 1.049 | 0.628 | 21.976 | 0.195 |
| Observations | 249,606 | 249,606 | 249,606 | 256,787 |
| <i>504 Disability</i> | -0.106 (0.089) | -0.021 (0.080) | -0.749 (0.888) | -0.025+ (0.014) |
| Baseline Mean | 0.728 | 0.298 | 19.982 | 0.065 |
| Observations | 59,912 | 59,912 | 59,912 | 60,764 |
| <i>Physical Disability</i> | -0.124 (0.199) | 0.007 (0.164) | -1.199 (1.828) | -0.053 (0.052) |
| Baseline Mean | 0.885 | 0.389 | 24.140 | 0.190 |
| Observations | 18,518 | 18,518 | 18,518 | 19,450 |
| <i>Cognitive Disability</i> | -0.127 (0.083) | -0.003 (0.113) | -1.211* (0.544) | -0.058** (0.014) |
| Baseline Mean | 1.134 | 0.722 | 22.234 | 0.224 |
| Observations | 170,356 | 170,356 | 170,356 | 175,747 |
| <i>No Disability</i> | -0.170* (0.071) | -0.037 (0.064) | -0.483 (0.490) | -0.018** (0.006) |
| Baseline Mean | 0.922 | 0.372 | 17.750 | 0.114 |
| Observations | 1,105,898 | 1,105,898 | 1,105,898 | 1,123,012 |

Notes: This table shows results by disability status. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). The baseline mean for each outcome reflects the SY13 mean value. Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A20: High School Restorative Practices: School Climate and Learning Outcomes by Disability Status

| | School Climate | GPA | Reading Scores | Math Scores |
|-----------------------------|--------------------|--------------------|-------------------|-------------------|
| <i>Any Disability</i> | 0.025 (0.019) | -0.047+ (0.028) | 0.038 (0.026) | 0.035 (0.026) |
| Baseline Mean | -0.019 | 2.259 | -0.514 | -0.649 |
| Observations | 115,114 | 153,265 | 143,852 | 142,836 |
| <i>504 Disability</i> | 0.058+ (0.032) | -0.010 (0.049) | -0.048 (0.049) | -0.020 (0.040) |
| Baseline Mean | 0.036 | 2.492 | 0.446 | 0.357 |
| Observations | 32,986 | 42,328 | 37,906 | 37,324 |
| <i>Physical Disability</i> | 0.072 (0.077) | -0.158* (0.069) | -0.006 (0.073) | -0.049 (0.073) |
| Baseline Mean | 0.056 | 2.417 | -0.334 | -0.485 |
| Observations | 7,182 | 11,765 | 9,336 | 9,259 |
| <i>Cognitive Disability</i> | 0.023 (0.025) | -0.048 (0.031) | 0.048* (0.021) | 0.043+ (0.024) |
| Baseline Mean | -0.040 | 2.190 | -0.789 | -0.934 |
| Observations | 81,161 | 97,767 | 95,281 | 94,885 |
| <i>No Disability</i> | 0.042** (0.016) | -0.025 (0.022) | -0.009 (0.021) | -0.014 (0.023) |
| Baseline Mean | 0.003 | 2.494 | 0.093 | 0.118 |
| Observations | 628,888 | 743,177 | 687,473 | 680,868 |

Notes: This table shows results by disability status. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods (five periods for the school climate index). Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). The baseline mean for each outcome reflects the SY13 mean value. Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A21: High School Restorative Practices: Imputed Test Scores by Race and Gender

| | Reading Scores (Imputed) (1) | Math Scores (Imputed) (2) |
|----------------------|------------------------------------|---------------------------------|
| <i>Black Female</i> | 0.017 | 0.024 |
| Standard Error | (0.021) | (0.020) |
| Baseline Mean | -0.108 | -0.250 |
| Observations | 200,156 | 200,023 |
| <i>White Female</i> | 0.001 | 0.007 |
| Standard Error | (0.039) | (0.027) |
| Baseline Mean | 0.855 | 0.680 |
| Observations | 45,880 | 45,850 |
| <i>Latina Female</i> | -0.029 | -0.027 |
| Standard Error | (0.036) | (0.037) |
| Baseline Mean | 0.010 | -0.015 |
| Observations | 213,420 | 213,289 |
| <i>Black Male</i> | 0.011 | 0.038* |
| Standard Error | (0.020) | (0.017) |
| Baseline Mean | -0.351 | -0.347 |
| Observations | 181,817 | 181,704 |
| <i>White Male</i> | 0.043 | 0.039 |
| Standard Error | (0.040) | (0.039) |
| Baseline Mean | 0.564 | 0.643 |
| Observations | 43,866 | 43,850 |
| <i>Latino Male</i> | -0.031 | -0.036 |
| Standard Error | (0.026) | (0.035) |
| Baseline Mean | -0.127 | 0.009 |
| Observations | 212,140 | 211,999 |

Notes: This table shows the average effect of restorative practices by student race and gender. Math and reading scores are imputed based on the methodology described in the text in Section VI. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A22: High School Baseline Outcomes by School Share of Black Students

| Outcome | <i>Below-Median Share</i> | | <i>Above-Median Share</i> | | Difference (5) |
|-------------------------------|---------------------------|---------------------|---------------------------|---------------------|-------------------|
| | Baseline Mean (1) | Observations (2) | Baseline Mean (3) | Observations (4) | |
| Math Score | 0.122 (1.008) | 35,086 | -0.117 (0.978) | 36,373 | 0.239 (0.15) |
| Reading Score | 0.076 (1.029) | 35,066 | -0.073 (0.965) | 36,280 | 0.149 (0.157) |
| GPA | 2.556 (0.982) | 42,385 | 2.346 (0.971) | 37,686 | 0.21* (0.092) |
| School Climate Index | 0.052 (0.629) | 34,406 | -0.052 (0.652) | 34,845 | 0.104* (0.05) |
| Out-of-School Suspension Days | 0.503 (2.071) | 49,101 | 1.338 (3.641) | 54,036 | -0.835** (0.122) |
| In-School Suspension Days | 0.169 (0.931) | 49,101 | 0.635 (2.012) | 54,036 | -0.466** (0.085) |
| Number of Arrests | 0.061 (0.411) | 50,108 | 0.188 (0.76) | 55,698 | -0.128** (0.019) |

Notes: This table displays the mean for each outcome variable in SY13 for students at any grade level between 9 and 12, disaggregated by whether students attend schools classified as having an above- or below-median share of Black students relative to the population of CPS schools based on SY13 enrollment and the median student’s school-level share in SY13. The share of Black students for schools not observed in SY13 is calculated using the first year that the school appears in the sample. The associated differences (column 5) are derived from student-level regressions of the given outcome on an indicator for a student’s school having a below-median share of Black students, with the standard errors clustered at the school level. Math and reading scores are standardized by test, school year, and grade. GPA is calculated using semester final grades. The school climate index measures student socioemotional wellbeing levels and perceptions regarding the supportiveness of school environments based on constructs from the My Voice My School (MVMS) survey. An out-of-school suspension is defined as the removal of a student from class attendance or school attendance. An in-school suspension is defined as the removal of a student from their regular educational schedule for more than 60 minutes of the school day to an alternative supervised setting inside the school building. Arrest data are collected by the Chicago Police Department. The arrest outcome is defined as the number of arrests experienced by students in a given year, regardless of the type of arrest or the location of the arrest. See Data Appendix C for detailed variable definitions. Standard errors are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A23: High School Restorative Practices: Imputed Test Scores by Predicted Peer Group Suspension Days

| | <i>All Students</i> | | <i>Low Predicted OSS Days Students</i> | |
|--------------------------------------|------------------------------------|---------------------------------|--|---------------------------------|
| | Reading Scores (Imputed) (1) | Math Scores (Imputed) (2) | Reading Scores (Imputed) (3) | Math Scores (Imputed) (4) |
| <i>Below-Median Predicted OSS</i> | -0.025 (0.036) | -0.038 (0.039) | -0.009 (0.035) | -0.031 (0.037) |
| Observations | 404,347 | 404,158 | 275,833 | 275,717 |
| <i>Above-Median Predicted OSS</i> | 0.012 (0.021) | 0.022 (0.018) | -0.018 (0.029) | 0.005 (0.024) |
| Observations | 433,459 | 433,321 | 133,900 | 133,943 |
| Control for Own Predicted Suspension | | | ✓ | ✓ |

Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. We present results for students belonging to school-by-cohort cells that are above- versus below-median in predicted suspension days within a given cohort. Predictions for out-of-school suspension days for each student are constructed using a random forest algorithm as described in Section VII. Math and reading score imputation is based on the methodology described in the text in Section VI. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

Table A24: High School Restorative Practices: Imputed Test Scores by Race and School Share of Black Students

| | Reading Scores (Imputed) | Math Scores (Imputed) |
|---|-----------------------------|--------------------------|
| <i>Panel A: All Students</i> | | |
| <i>Below-Median Black Student Share</i> | -0.010 (0.029) | -0.023 (0.035) |
| Observations | 324,283 | 324,194 |
| <i>Above-Median Black Student Share</i> | -0.008 (0.022) | 0.024 (0.019) |
| Observations | 502,178 | 501,870 |
| <i>Panel B: Black Students</i> | | |
| <i>Below-Median Black Student Share</i> | -0.013 (0.037) | -0.026 (0.038) |
| Observations | 33,730 | 33,720 |
| <i>Above-Median Black Student Share</i> | 0.015 (0.023) | 0.037+ (0.019) |
| Observations | 337,414 | 337,183 |
| <i>Panel C: Latine Students</i> | | |
| <i>Below-Median Black Student Share</i> | -0.034 (0.032) | -0.050 (0.042) |
| Observations | 224,002 | 223,932 |
| <i>Above-Median Black Student Share</i> | -0.042 (0.034) | 0.011 (0.036) |
| Observations | 118,930 | 118,885 |

Notes: A school is classified as having an above- or below-median share of Black students based on SY13 enrollment and the median student's school-level share in SY13. The share of Black students for schools not observed in SY13 is calculated using the first year that the school appears in the sample. Observations are at the student-school year level, and we report the average effect of restorative practices over six periods. Math and reading score imputation is based on the methodology described in the text in Section VI. Each specification includes the following covariates: student age fixed effects, student cohort fixed effects (based on grade and school year of entry), ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates are based on the methodology developed in de Chaisemartin and D'Haultfoeuille (2020) and described in the text. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1 percent level, * at the 5 percent level, and + at the 10 percent level.

C Data Appendix

C.A Variable Definitions

Suspensions (OSS and ISS). Suspension data are collected by Chicago Public Schools. The out-of-school suspension (OSS) days and in-school suspension (ISS) days outcomes are the total number of ISS or OSS days that the student received in the corresponding school year, regardless of the school.

An in-school suspension (ISS) is the removal of a student from their regular educational schedule for more than 60 minutes of the school day to an alternative supervised setting inside the school building. The student is meant to engage in structured activities that develop their academic, social, emotional, and/or behavioral skills. A student may receive an in-school suspension if it is an available consequence for the behavior they exhibited as categorized in the Student Code of Conduct (SCC). CPS categorizes student misconduct behaviors into six categories, ranging from Group 1 (Inappropriate Behaviors) to Group 6 (Illegal and Most Seriously Disruptive Behaviors). ISS is an available consequence for Group 2 (Disruptive Behaviors) - Group 6 (Illegal and Most Seriously Disruptive Behaviors) SCC violations.

An out-of-school suspension (OSS) is the removal of a student from class attendance or school attendance. By law, out-of-school suspensions may only be issued as a last resort consequence if all other appropriate and available responses have been exhausted, and suspensions longer than 3 days must receive approval from a designated CPS district employee. An out-of-school suspension may be issued if it is an available consequence for the behavior exhibited as categorized in the SCC. OSS is an available consequence for Group 3 (Seriously Disruptive Behaviors) - Group 6 (Illegal and Most Seriously Disruptive Behaviors) SCC violations.

If a student does not have any recorded in-school (out-of-school) suspensions in a given school year, we code in-school (out-of-school) suspension days as zero.

Absent Days. The absent days outcome is the total number of days absent minus the total number of OSS days that the student received in the corresponding year, regardless of school. Regressions that employ absent days as an outcome always include yearly total member days (defined by CPS as the number of days a student is enrolled in any CPS school) as a control.

School Climate Index. The school climate index measures student socioemotional well-being levels and perceptions regarding the supportiveness of school environments. The estimates are drawn from the My Voice, My School (MVMS) survey. MVMS was developed by the University of Chicago Consortium on School Research and was administered to all students in grades 6-12 beginning in SY11 (data were unavailable for SY19 at the time of analysis). The available data contains one Rasch score per survey construct (21 constructs in total) per student in each school year. The school climate index is created as the average

of the following eight constructs (each standardized by grade and school year) that may be directly affected by the introduction of RP: Emotional Health, Student Classroom Behavior, Academic Personalism, Psychological Sense of School Membership, Safety, School-Wide Future Orientation, School Safety, Student-Teacher Trust.

GPA. Student grades can be recorded as: progress grades, semester final grades, or yearly final grades of semester classes. The yearly final grades do not always match the semester grades because they reflect a holistic assessment of the student’s performance over the entire year, not just over a semester. However, not all courses have associated yearly final grades, and many students do not receive any yearly final grades for multiple academic years. The GPA outcome, therefore, is calculated using only semester final grades. There are two kinds of courses: “for-credit” courses which are assigned a letter grade (A through F) and pass/fail courses (which a student either passes or fails). The GPA outcome is the mean of the numeric grades (equivalent to the letter grades) registered in the data for all for-credit courses. Numeric grades are calculated as follows: A is equivalent to 4; B, to 3; C, to 2; D, to 1; and F, to 0. The data do not differentiate an F grade from a pass/fail versus a for-credit course. In our analyses, all F grades are counted as if they were grades of for-credit courses and included in our GPA calculation. It is important to note that the GPA used in our analyses may not reflect the same GPA the students see on their transcripts due to differing procedures used to calculate GPAs within schools.

Standardized Tests. Standardized test scores are available for students in grades 3 to 11; tests are typically administered in April of the given school year. The CPS assessments used for elementary and high school grades changed during our study period from SY09-SY19. Below is a timeline and brief description of each test instrument used in our analysis. Math and reading scores are standardized by subject, school year, and grade within the study sample, where we center and scale the scores to have zero mean and unit variance. Note that only a single test instrument was used for each grade-level-by-school-year cell.

Applicable Timeline for Grades 3-8: Illinois Standards Achievement Test (ISAT) was used in SY09-SY14, Northwest Evaluation Association (NWEA) MAP exam was used in SY15-SY19. ISAT was administered to CPS students in grades 3-8 as a mandatory assessment. Students were assessed on Reading and Math in grades 3-8, and additionally in Science in grades 4 and 7. The ISAT was replaced by the NWEA MAP exam after SY14, which was administered in grades 2-8. The NWEA was a district-required test during the SY15-SY19 period.

Applicable Timeline for Grade 9: EXPLORE was used in SY09-SY14, Partnership for Assessment of Readiness for College and Careers (PARCC) was used in SY15-SY16, and the PSAT was used in SY17-SY19. EXPLORE was administered to CPS students in grade 9 as a mandatory assessment with valid scores available through SY14. In SY15, the ACT company (which provides testing services) made errors in the administration of EXPLORE leading scores to be invalidated. For SY15-SY16, the only available test scores for grade 9 students are derived from the PARCC exam, which is intended to assess students in English

Language Arts/Literacy and Mathematics. PARCC was given to high school students based on their enrollment in specific courses and has a higher rate of missingness (30-40%) as compared to the other mandatory assessments administered in prior and subsequent years. Beginning in SY17, the PSAT became a required assessment for students in grades 9 and 10. The test is designed specifically for students in grades 9-10.

Applicable Timeline for Grade 10: PLAN was used in SY09-SY14 and the PSAT was used in SY17-SY19. PLAN was administered to CPS students in grade 10 as a mandatory assessment with valid scores available through SY14. In SY15, the ACT company (which provides testing services) made errors in the administration of PLAN leading scores to be invalidated. For SY15-SY16, there are no available test scores for grade 10 students. Beginning in SY17, the PSAT became a required assessment for students in grades 9 and 10. The test is designed specifically for students in grades 9-10.

Applicable Timeline for Grade 11: Prairie State Achievement Examination (PSAE) was used in SY09-SY14, ACT was used in SY15-SY16, and the SAT was used in SY17-SY19. PSAE, which includes the ACT, was administered to CPS students in grade 11 as a mandatory assessment through SY14. In SY15-SY16, only the ACT component was administered to grade 11 students. Beginning in SY17, the SAT replaced the ACT as the test instrument administered to all grade 11 students in CPS.

Arrests. The arrest data are derived from Chicago Police Department (CPD) arrest files, which include detailed information about each unique arrest that occurred between July 1, 2008 and June 30, 2019. The arrest outcome used in benchmark analyses is the number of times a student was arrested in a given year, regardless of the type of arrest or the location of the arrest. In-school arrests are classified as incidents happening both inside the school location and during school hours while out-of-school arrests are incidents happening either outside the school location or outside school hours. Arrest records with the following FBI National Incident-Based Reporting System (NIBRS) codes are categorized as violent arrests, while all other arrests are categorized as non-violent: 01A (Homicide, 1st 2nd Degree), 02 (Criminal Sexual Assault), 03 (Robbery), 04A (Aggravated Assault), 04B (Aggravated Battery), 08A (Simple Assault), 08B (Simple Battery). In additional specifications, we examine the likelihood of arrest by creating an indicator variable for whether or not an individual was arrested in a given academic year.

Demographics. CPS provided demographic information on each student for each school year from SY09-SY19, including information about student gender, race, disability status, English Language Learner status, whether students are classified as needing an individualized education plan (IEP) or a 504 plan, whether they are eligible for the free- or reduced-price lunch program, and whether they are classified as unhoused.

Variables related to disabilities. IEP plans and 504 plans are distinct categories, but both can offer formal assistance for students who need additional help in school at no cost to families (Understood, 2023). IEP plans serve as a blueprint for a student's education experience at school and provide individualized special education and related services to meet a student's needs. There are two criteria for being given an IEP plan: (1) the student must have at least one of 13 disabilities listed in the Individuals with Disabilities Education Act

(IDEA) (e.g. learning disabilities, physical disabilities), and (2) the disability must affect the child’s educational performance or ability to learn and benefit from the curriculum such that the student could not make progress in school without specialized instruction. A 504 plan defines disability more broadly than IDEA so can be applied to students whose disabilities or barriers are not covered by an IEP. It is called “504” because it comes from Section 504 of the Rehabilitation Act of 1973, a federal law aimed to protect people with disabilities. In order to qualify for a 504 plan, students must have (1) any disability (which is less restrictive than an IEP), and (2) this disability must interfere with a student’s ability to learn. For students with IEP plans (also referred to as special education plans), we observe the specific primary disability that a student is classified as having (with the primary disability determined by the student). In contrast, for students with 504 plans, we do not see information on specific disability classifications.

In benchmark models, we include indicators for physical disability (classified as blind, having a visual impairment, deaf/blind, deaf, hearing impaired, hard of hearing, having “[an]other health impairment,” physical handicap, partial sight, severe/profound disability, or traumatic brain injury), cognitive disability (classified as autistic, having a behavior disorder, developmental delay, emotional and behavioral disorder, having emotional disturbance, having an educable mental disability, a profound intellectual disability, having a learning disability, having a mental disability, a moderate learning disability, a severe learning disability, speech/language challenges, or a trainable mental disability), or having a 504 plan.

C.B Study Sample

Student treatment assignment for high school (elementary school) analyses is determined by the first high school (elementary school) a student had been enrolled in since SY09. The sample covers students in grades 9 to 12 (3 to 8) between SY09 and SY19. Observations are at the student-school year level, and we include a student-school year in the sample if the student was enrolled in any high school (elementary school) under the purview of CPS for at least one day in a given school year. To construct the sample and assign treatment, we rely on enrollment history records provided by CPS. The enrollment history data are prepared by CPS using data reported by each school and reflect the most accurate available information on a student’s journey through CPS, including enrollment start and termination dates at each school in CPS. Though rare, there are occasional data errors related to the construction of unique student identifiers, which are used to track students across schools and school years. Our study sample excludes students who are classified as having progressed to grade levels not offered by their initial schools, students past their expected school exit year, and any observations beyond our event study window for students assigned to treatment schools. Specifically, we include students who appear in the data between five years before and five years after treatment for all outcomes other than the school climate index. Due to more limited data availability, specifications with the school climate index outcome are restricted to including students who appear in the data between three years before and four years after treatment.