

Supplementary Appendix

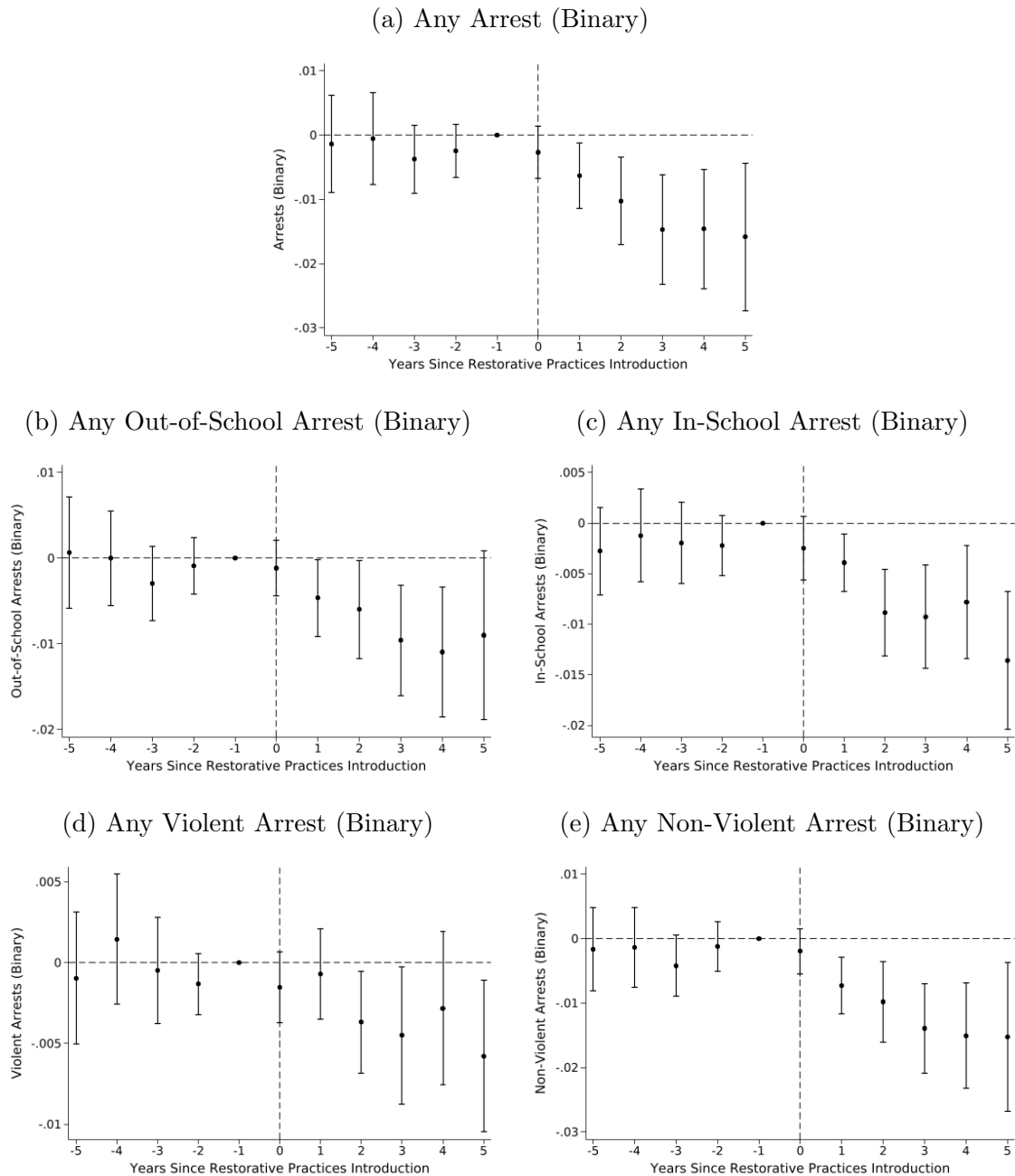
**“From Retributive to Restorative:
An Alternative Approach to Justice in Schools”**

Anjali Adukia, Benjamin Feigenberg, Fatemeh Momeni

Appendices

A Appendix Figures

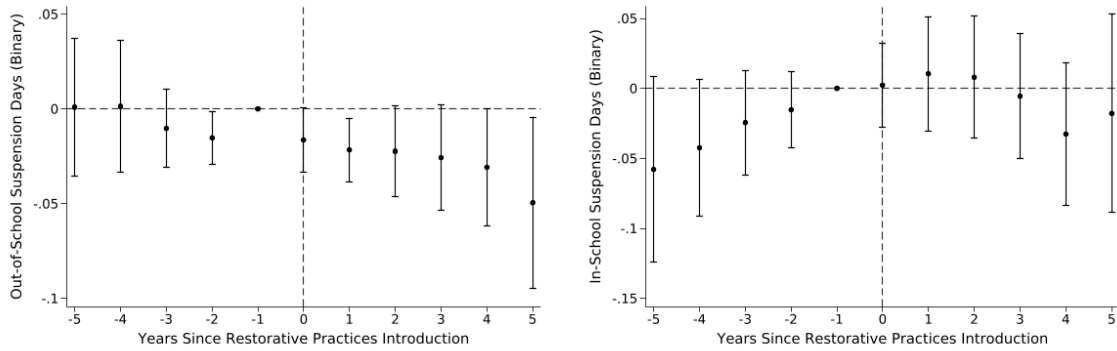
Figure A1: High School Event Studies: Binary Policing Outcomes



Notes: These figures show the event studies around the introduction of RP on students' binary arrest outcomes 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, and the sample covers students in grades 9 to 12 between SY09 and SY19. Arrest data are collected by the Chicago Police Department. 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 (outside of 7am-6:59pm on school 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/ethnicity 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.

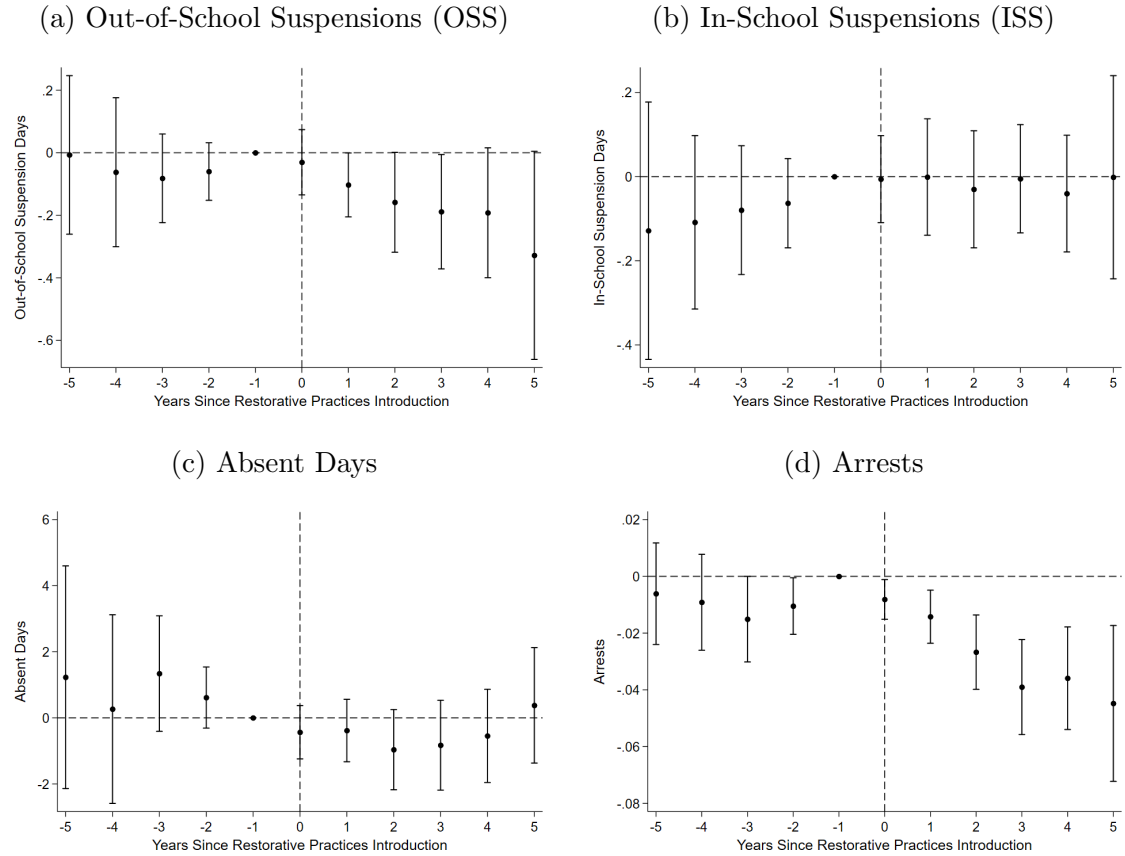
Figure A2: High School Event Studies: Binary Suspension Outcomes

(a) Any Out-of-School Suspension (Binary) (b) Any In-School Suspension (Binary)



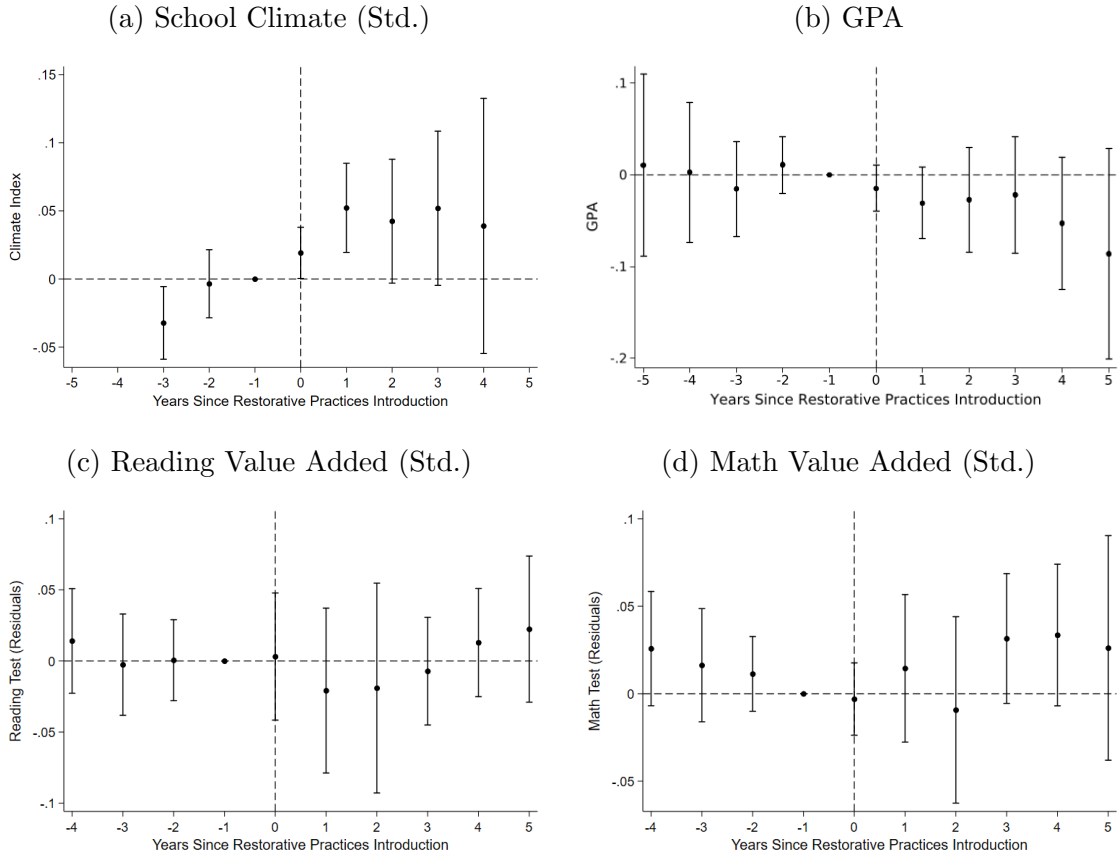
Notes: These figures show the event studies around the introduction of RP on students' binary suspension outcomes (any out-of-school suspension, any in-school suspension) over time. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. 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. 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/ethnicity 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 Excluding Controls: Behavioral Outcomes



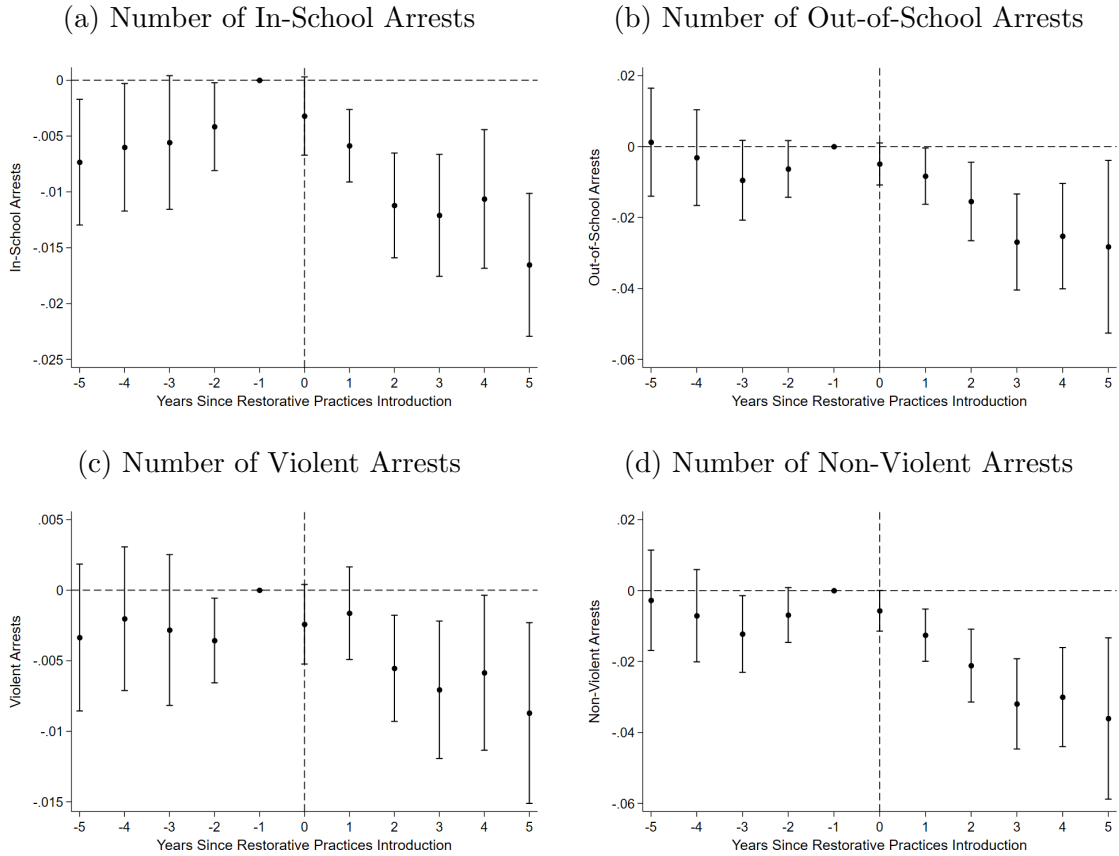
Notes: These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions, in-school suspensions, and absent days) and policing outcomes (overall arrests) over time in high schools. Observations are at the student-school year level. Student 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. 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 Event Studies Excluding Controls: School Climate and Learning



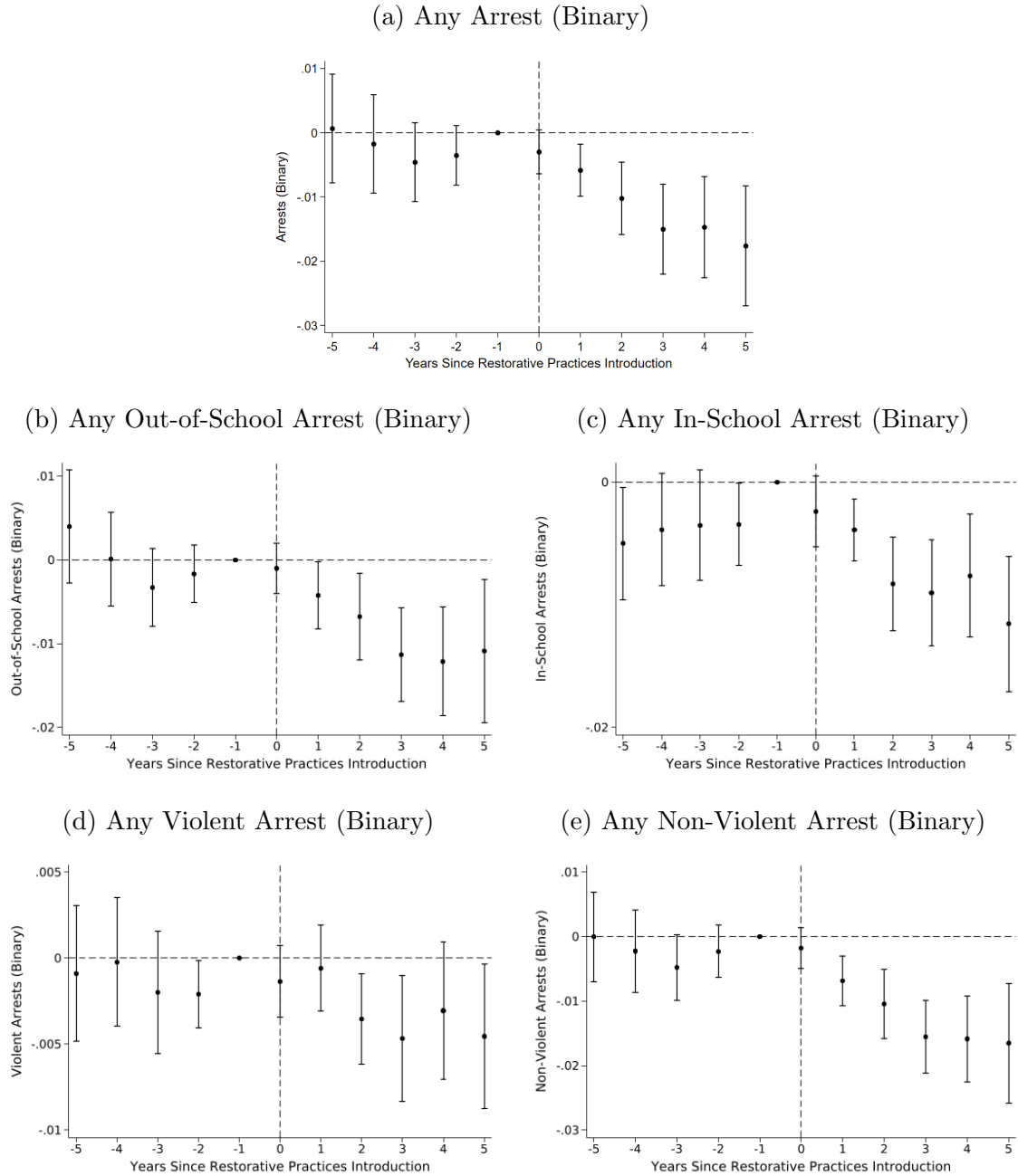
Notes: These figures show the event studies around the introduction of RP on students’ perceptions of school climate and academic outcomes (GPA, reading value added, and math value added) over time in high schools. Observations are at the student-school year level. Student treatment assignment is determined 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. The school climate index is standardized by school year and grade. 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; value added is then constructed based on the methodology described in the text in Section IV.A. Value added cannot be constructed for SY09; value-added graphs therefore reflect one fewer estimated placebo effect. See Data Appendix C for detailed variable definitions. 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 A5: High School Event Studies Excluding Controls: Policing Outcomes



Notes: These figures show the event studies around the introduction of RP on students' arrest outcomes (out-of-school vs. in-school, and violent vs. non-violent) 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, 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. 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 (outside of 7am-6:59pm on school days). See Data Appendix C for detailed variable definitions. 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.

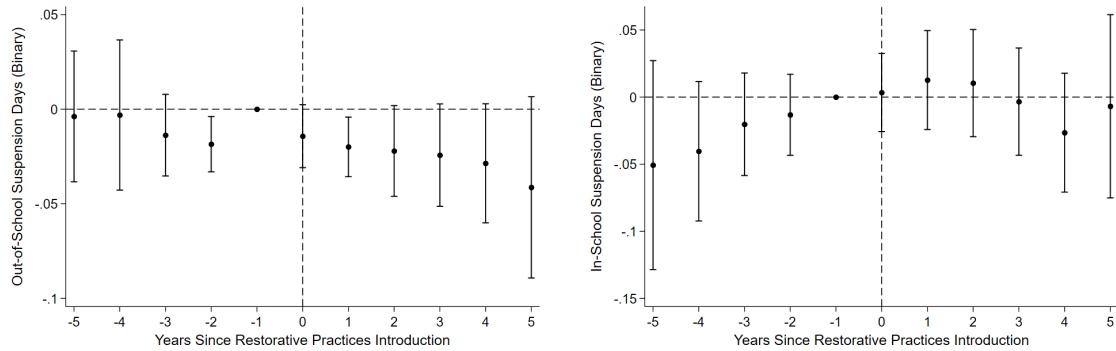
Figure A6: High School Event Studies Excluding Controls: Binary Policing Outcomes



Notes: These figures show the event studies around the introduction of RP on students' binary arrest outcomes (any arrest, any out-of-school vs. in-school, and any violent vs. non-violent) 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, 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. 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 (outside of 7am-6:59pm on school days). See Data Appendix C for detailed variable definitions. 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.

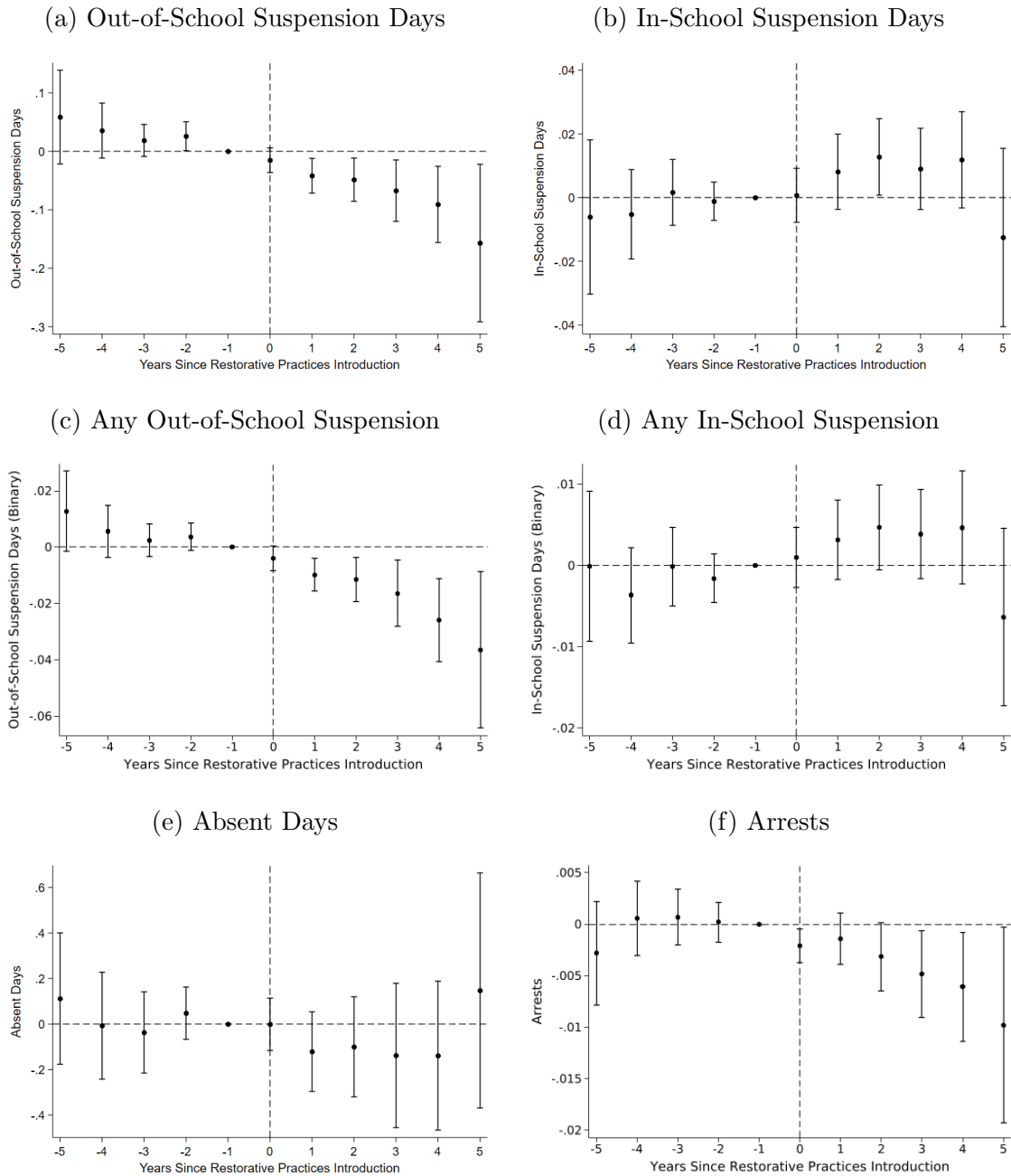
Figure A7: High School Event Studies Excluding Controls: Binary Suspension Outcomes

(a) Any Out-of-School Suspension (Binary) (b) Any In-School Suspension (Binary)



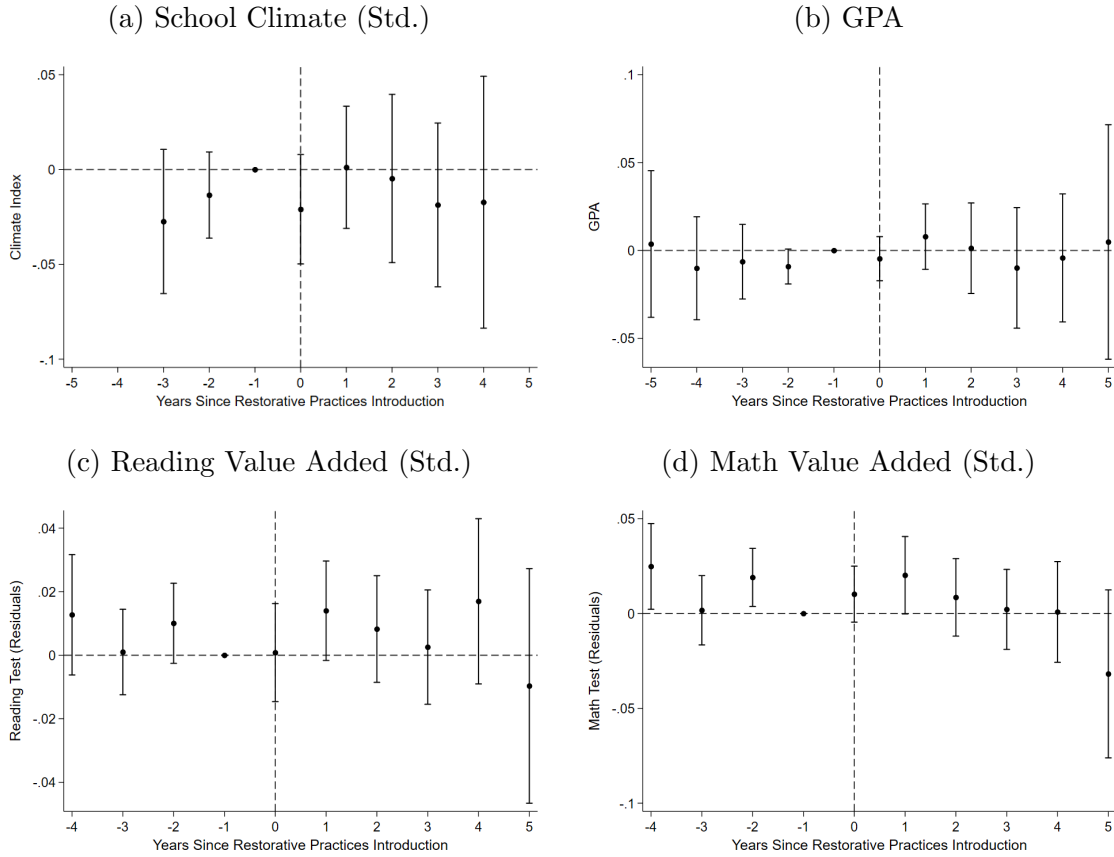
Notes: These figures show the event studies around the introduction of RP on students' binary suspension outcomes (any out-of-school suspension, any in-school suspension) over time. Observations are at the student-school year level. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09, and the sample covers students in grades 9 to 12 between SY09 and SY19. 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. See Data Appendix C for detailed variable definitions. 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: Elementary School Event Studies: Behavioral Outcomes



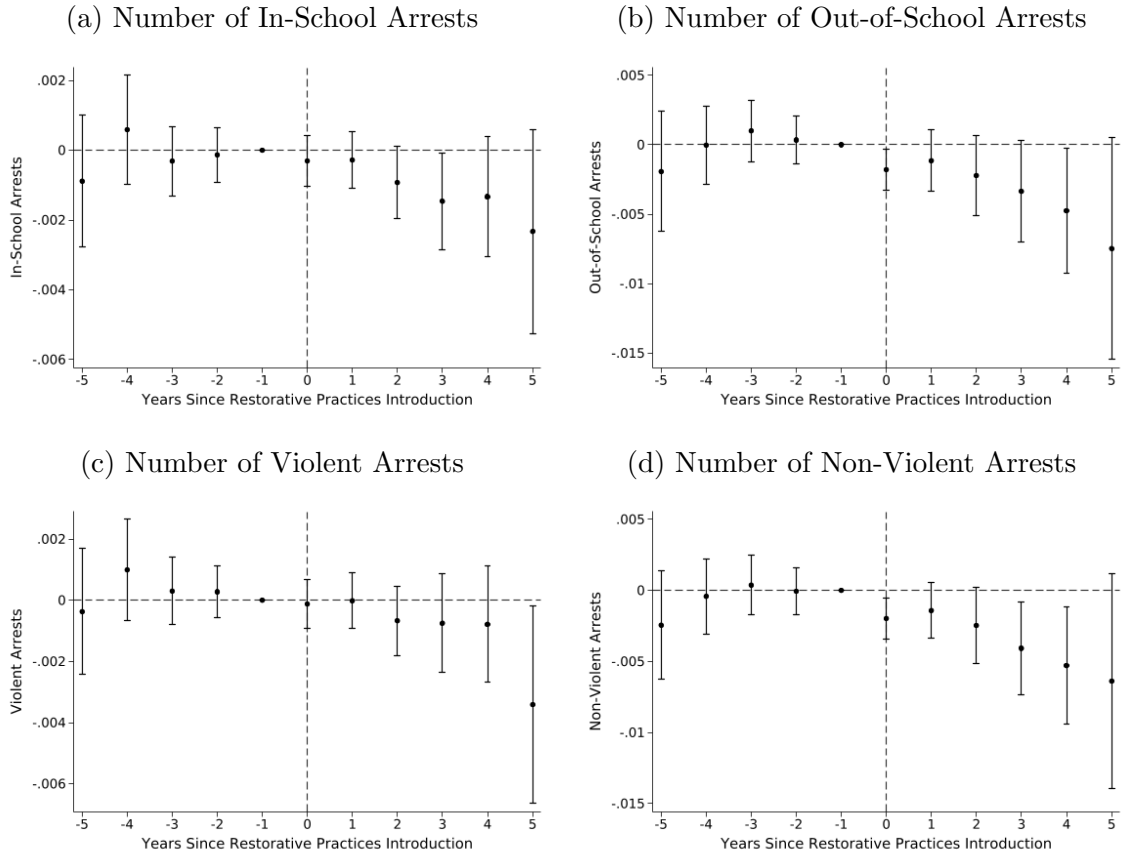
Notes: These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions, in-school suspensions, and absent days) and policing outcomes (overall arrests) over time in elementary schools. Observations are at the student-school year level. Student 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/ethnicity 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 A9: 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 value added, and math value added) 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. The school climate index is standardized by school year and grade. 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; value added is then constructed based on the methodology described in the text in Section IV.A. Value added cannot be constructed for SY09; value-added graphs therefore reflect one fewer estimated placebo effect. 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/ethnicity 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.

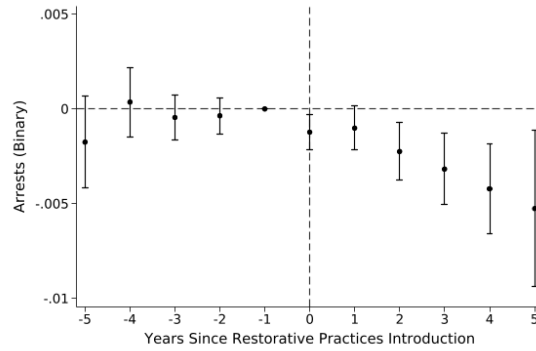
Figure A10: Elementary School Event Studies: Policing Outcomes



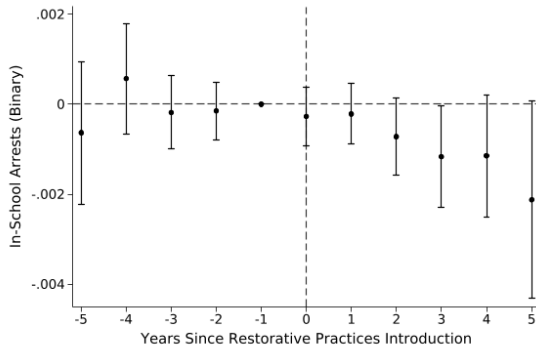
Notes: These figures show the event studies around the introduction of RP on students' arrest outcomes (out-of-school vs. in-school, and violent vs. non-violent) over time. Observations are at the student-school year level. Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. 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. 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 (outside of 7am-6:59pm on school 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/ethnicity 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: Elementary School Event Studies: Binary Policing Outcomes

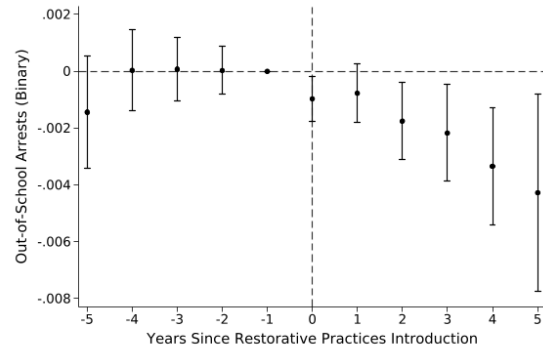
(a) Any Arrest (Binary)



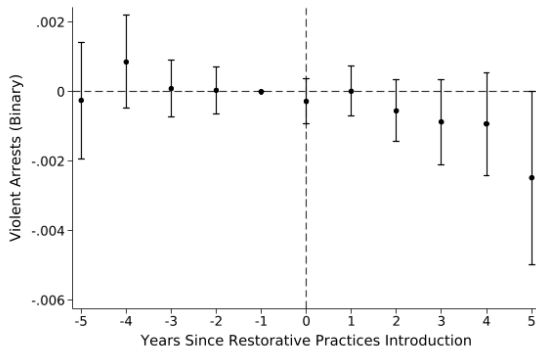
(b) Any In-School Arrest (Binary)



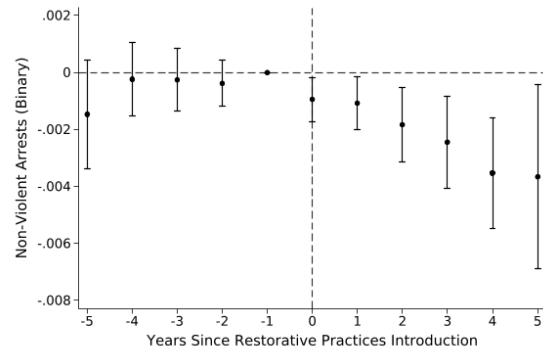
(c) Any Out-of-School Arrest (Binary)



(d) Any Violent Arrest (Binary)

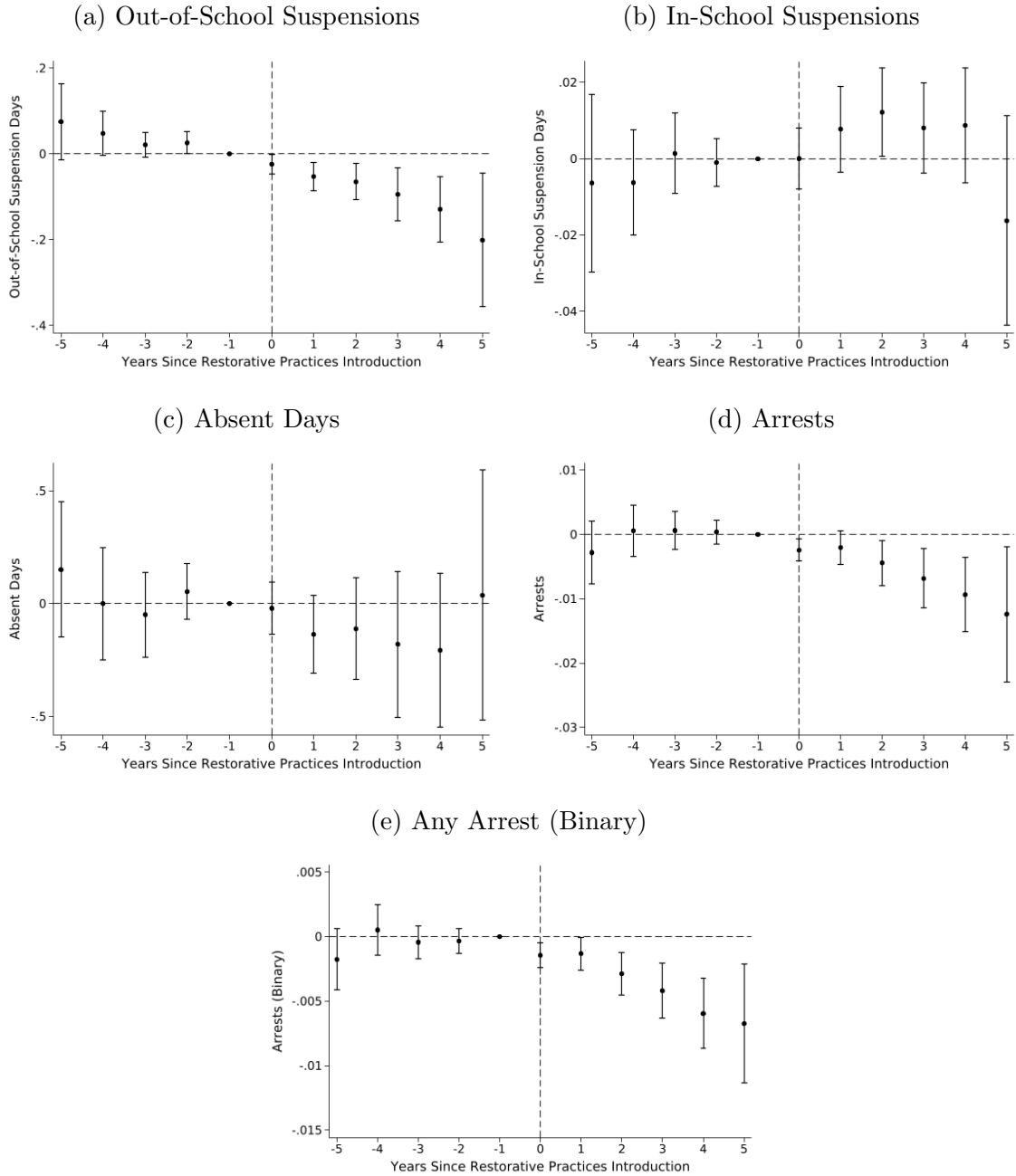


(e) Any Non-Violent Arrest (Binary)



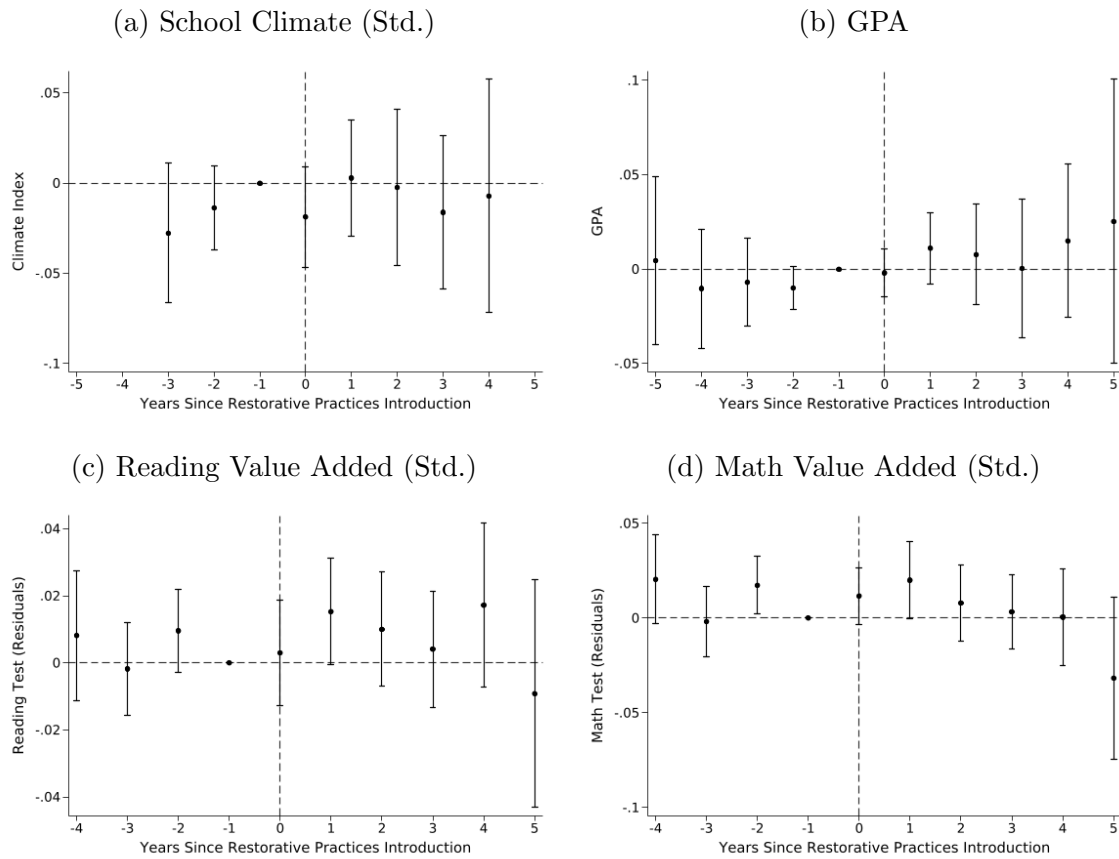
Notes: These figures show the event studies around the introduction of RP on students' binary arrest outcomes (any out-of-school vs. in-school, and any violent vs. non-violent) over time. Observations are at the student-school year level. Student treatment assignment is determined by the first elementary school a student had been enrolled in since SY09, and the sample covers students in grades 3 to 8 between SY09 and SY19. 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. 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 (outside of 7am-6:59pm on school 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/ethnicity 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 A12: Elementary School Event Studies Excluding Controls: 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 (in (d)) or an indicator for any arrest (in (e)) 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. 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.

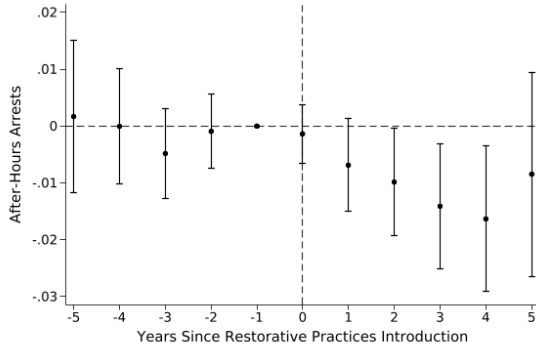
Figure A13: Elementary School Event Studies Excluding Controls: School Climate and Learning



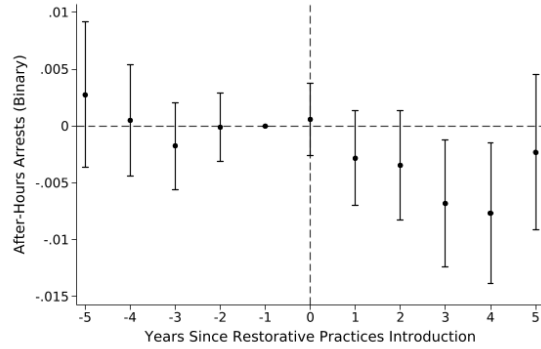
Notes: These figures show the event studies around the introduction of RP on students’ perceptions of school climate and academic outcomes (GPA, reading value added, and math value added) over time in elementary schools. Observations are at the student-school year level. Student treatment assignment 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. The school climate index is standardized by school year and grade. 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; value added is then constructed based on the methodology described in the text in Section IV.A. Value added cannot be constructed for SY09; value-added graphs therefore reflect one fewer estimated placebo effect. See Data Appendix C for detailed variable definitions. 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 A14: High School Event Studies: Alternative Out-of-School Arrest Outcomes

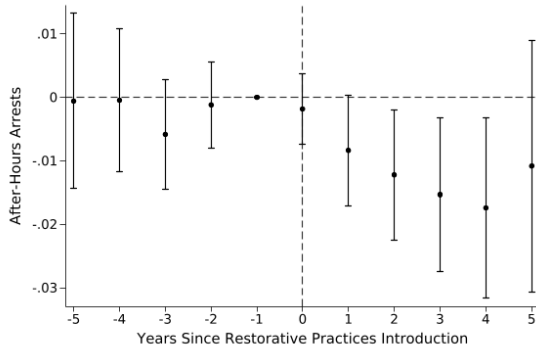
(a) Number of Arrests, Excluding 7am - 6:59pm on School Days



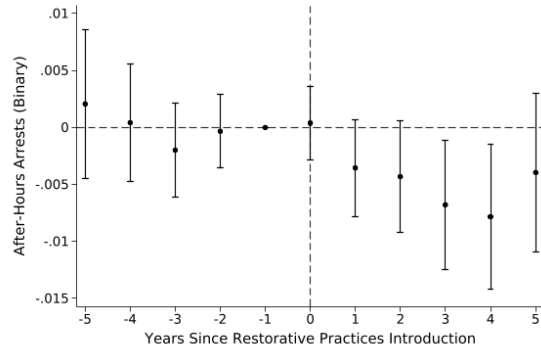
(b) Any Arrest, Excluding 7am - 6:59pm on School Days (Binary)



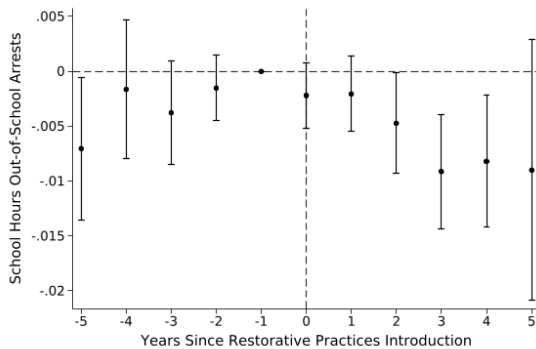
(c) Number of Arrests, Excluding 7am - 4:59pm on School Days



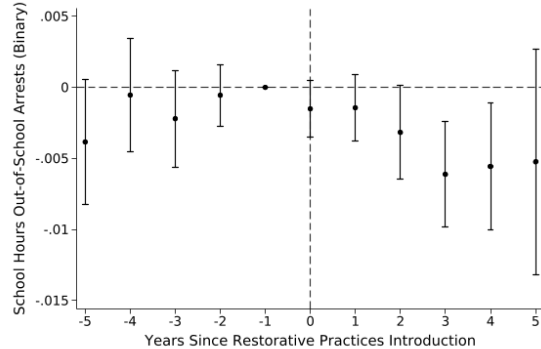
(d) Any Arrest, Excluding 7am - 4:59pm on School Days (Binary)



(e) Number of Arrests Outside School, 7am - 4:59pm on School Days

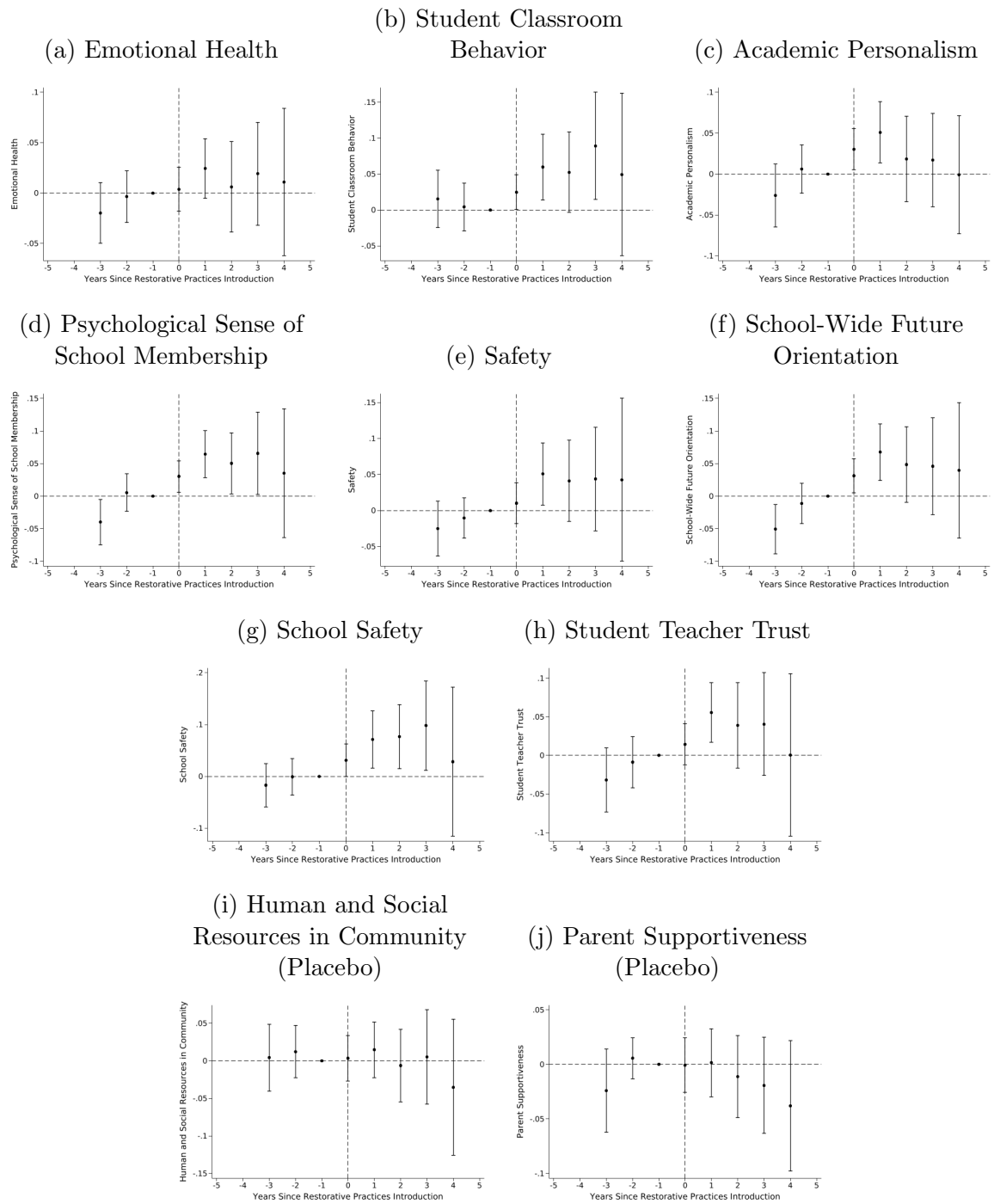


(f) Any Arrest Outside School, 7am - 4:59pm on School Days (Binary)



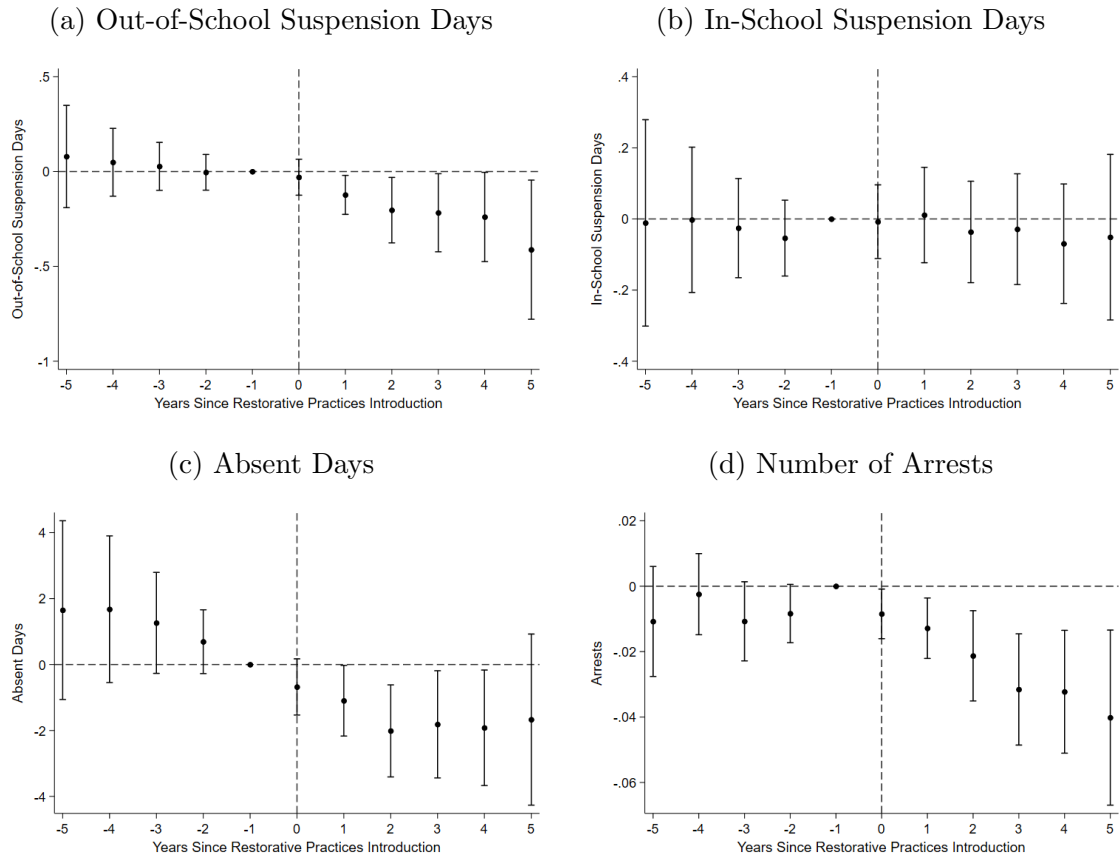
Notes: These figures show the event studies around the introduction of RP on alternative out-of-school arrest outcomes. In (a)-(d), these outcomes include all arrests taking place on non-school days as well as arrests on school days outside of the referenced time window. Observations are at the student-school 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. 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. 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/ethnicity 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 A15: High School Event Studies: Student Perception Survey (“My Voice, My School”) Constructs (Std.)



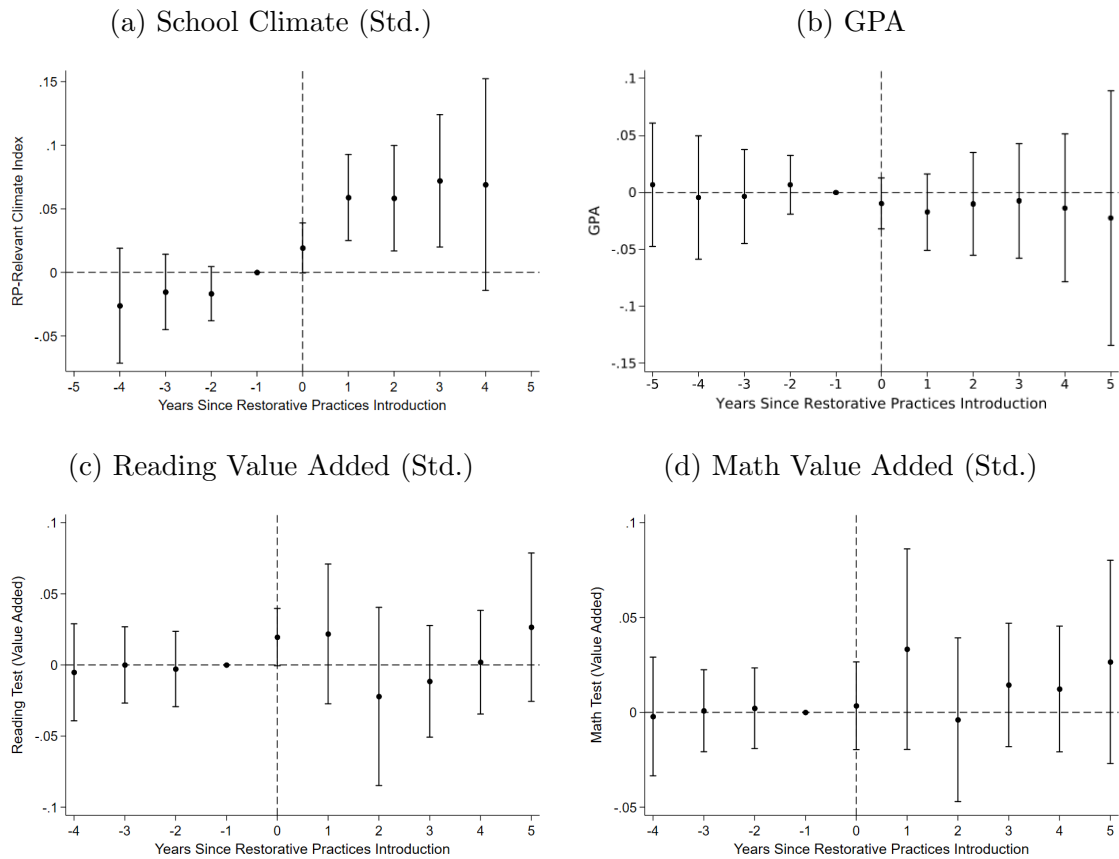
Notes: These figures show the event studies around the introduction of RP on constructs included in the student perception survey, which included measures of perceptions related to school climate. All constructs are corrected to fit a positively valenced metric, where higher scores are better. Observations 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 SY11 and SY18. Constructs are standardized by year and grade.

Figure A16: High School Event Studies Using a Standard Difference-in-Differences (Two-Way Fixed Effects) Design: Behavioral and Policing Outcomes



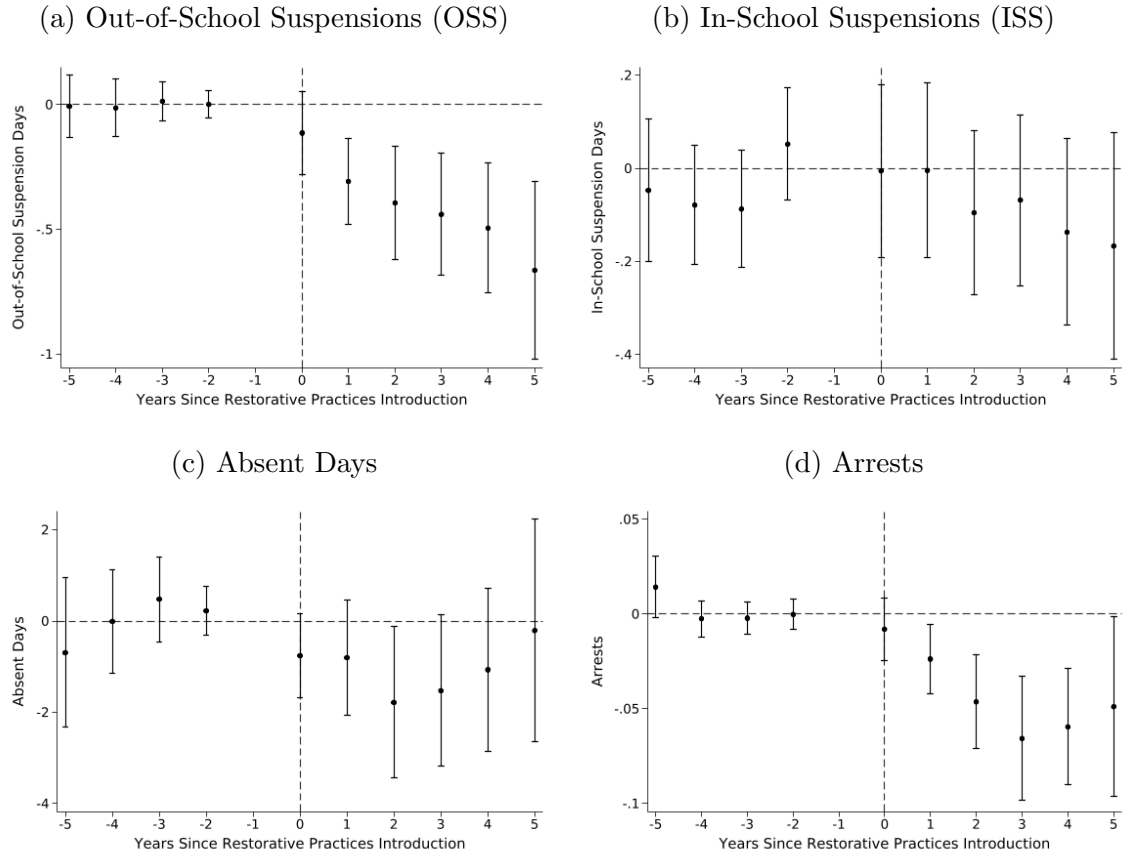
Notes: These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions, in-school suspensions, and absent days) and policing outcomes (overall arrests) over time in high schools. Observations are at the student-school year level. Student 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/ethnicity 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 A17: High School Event Studies Using a Standard Difference-in-Differences (Two-Way Fixed Effects) Design: School Climate and Learning



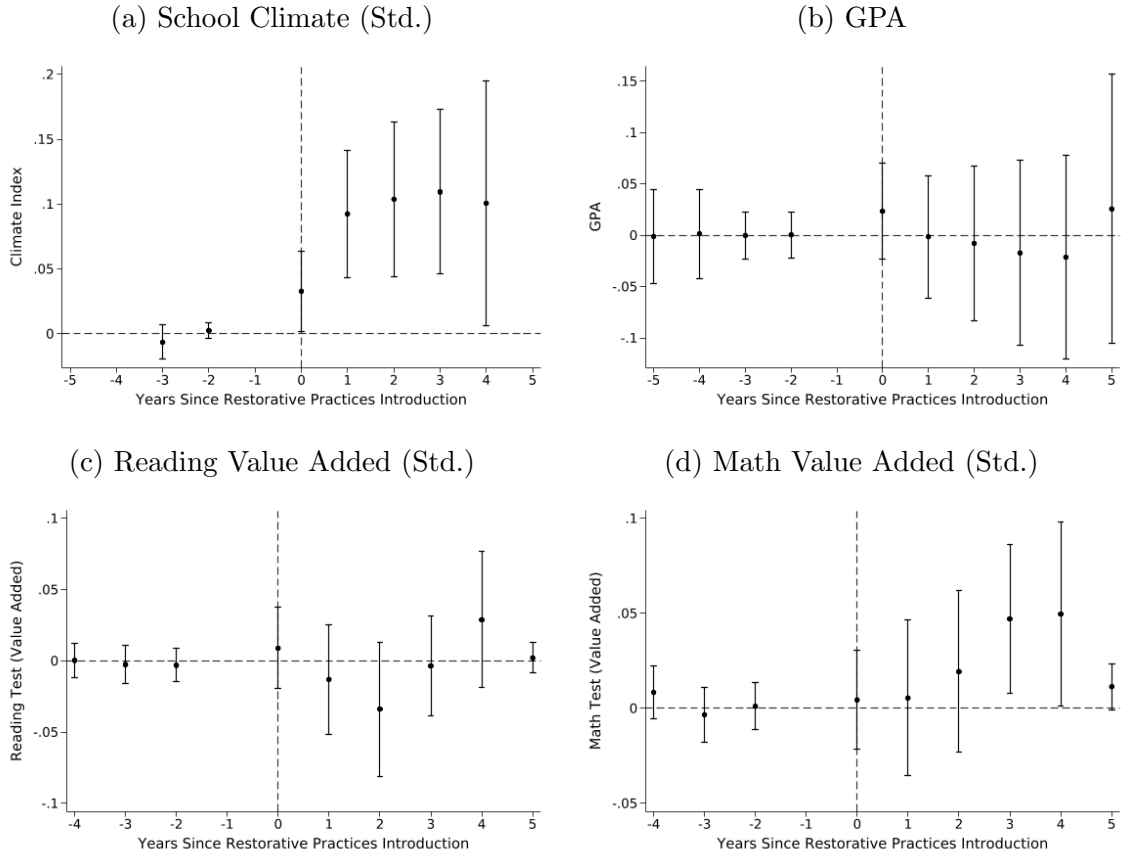
Notes: These figures show the event studies around the introduction of RP on students’ perceptions of school climate and academic outcomes (GPA, reading value added, and math value added) over time in high schools. Observations are at the student-school year level. Student treatment assignment is determined 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. The school climate index is standardized by school year and grade. 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; value added is then constructed based on the methodology described in the text in Section IV.A. Value added cannot be constructed for SY09; value-added graphs therefore reflect one fewer estimated placebo effect. 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/ethnicity 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 A18: Synthetic Difference-in-Differences: Behavioral Outcomes



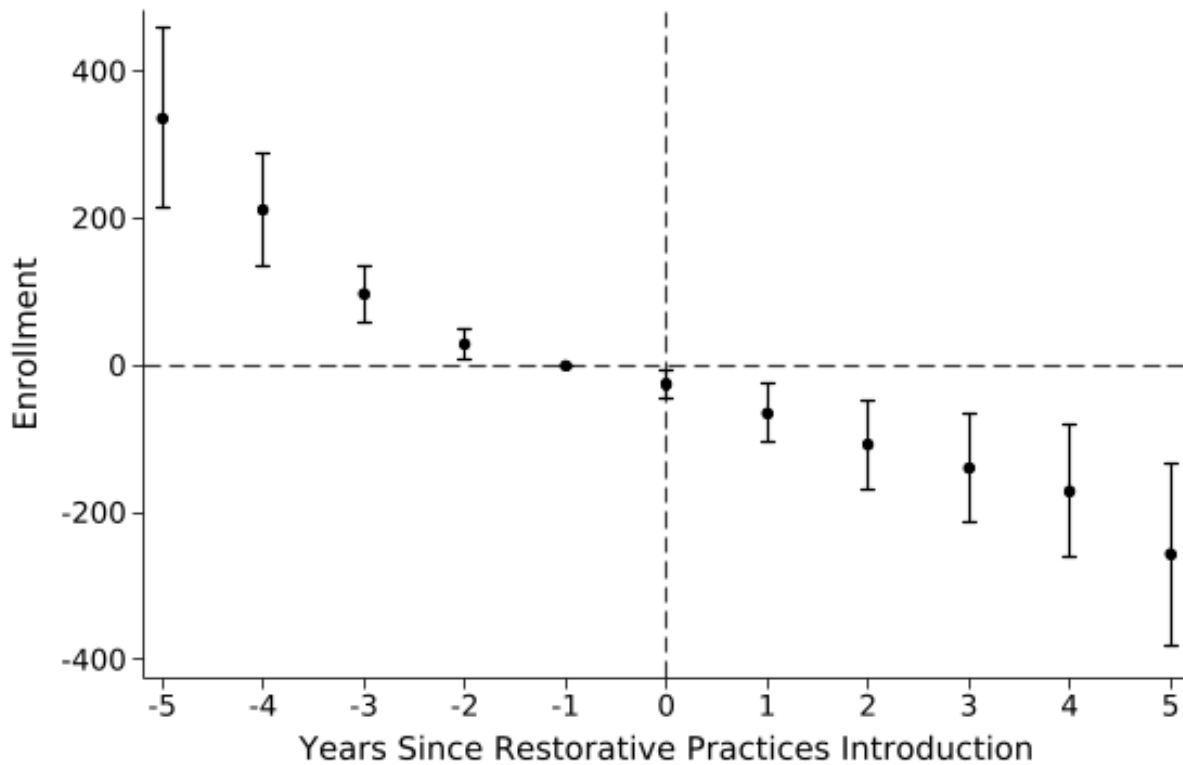
Notes: These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions, in-school suspensions, and absent days) and policing outcomes (overall arrests) over time in high schools. Observations are at the school-school year level and include only schools with non-missing observations in each school year between SY09 and SY19. Before data are collapsed to the school-school year level, students are assigned to the first high school they had been enrolled in since SY09. 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. Estimates are based on the methodology developed in Arkhangelsky et al. (2021). Bars represent 95% confidence intervals.

Figure A19: Synthetic Difference-in-Differences: School Climate and Learning



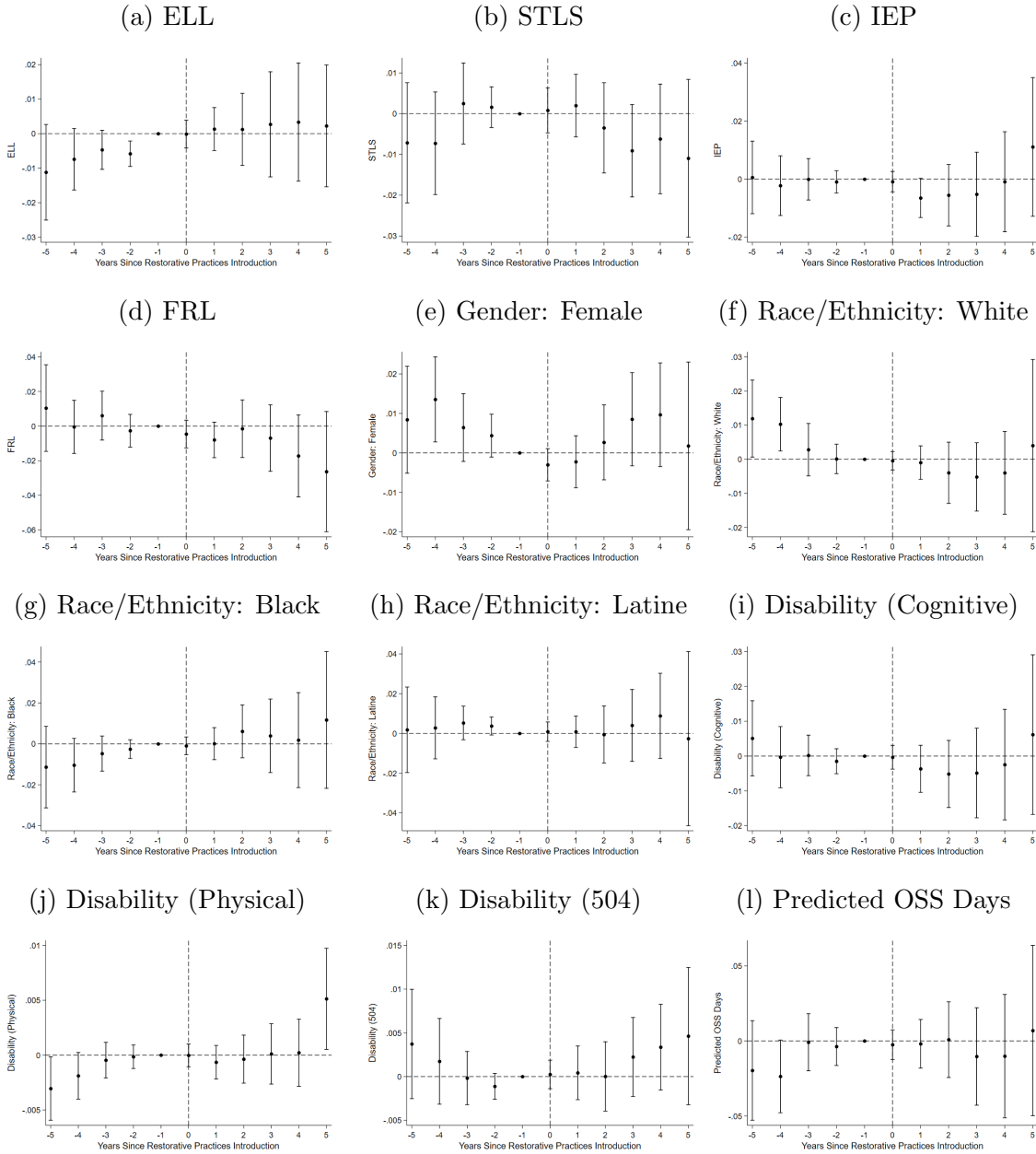
Notes: These figures show the event studies around the introduction of RP on students’ perceptions of school climate and academic outcomes (GPA, reading value added, and math value added) over time in high schools. Observations are at the school-school year level and include only schools with non-missing observations in each school year between SY09 and SY19 (between SY11 and SY18 for school climate). Before data are collapsed to the school-school year level, students are assigned to the first high school they had been enrolled in since SY09. 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. The school climate index is standardized by school year and grade. 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; value added is then constructed based on the methodology described in the text in Section IV.A. Value added cannot be constructed for SY09; value-added graphs therefore reflect one fewer estimated placebo effect. See Data Appendix C for detailed variable definitions. Estimates are based on the methodology developed in Arkhangelsky et al. (2021). Bars represent 95% confidence intervals.

Figure A20: High School Event Studies: Enrollment



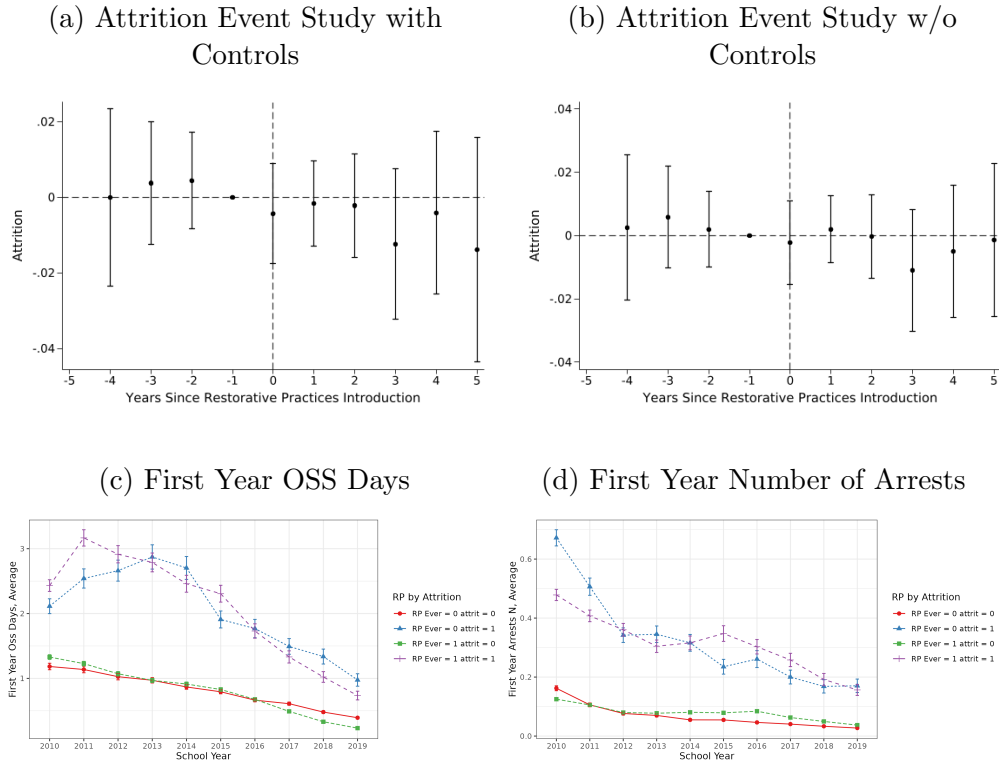
This figure shows the event study around the introduction of RP on annual high school (grades 9-12) enrollment. Observations are at the school-school year level. The sample covers all schools serving high school students for at least one year between SY09 and SY19. Estimates are based on the methodology 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 A21: High School Event Studies: Characteristics as Outcomes



These figures show the event studies around the introduction of RP where the outcomes are student characteristics (in Panels (a)-(k)) or predicted out-of-school suspension days (in Panel (l)). 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. ELL is an indicator for English Language Learner, STLS is an indicator for Students in Temporary Living Situations, IEP is an indicator for Individualized Education Plan, and FRL is an indicator for Free of Reduced-Price Lunch. See Data Appendix C for detailed variable definitions. 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.

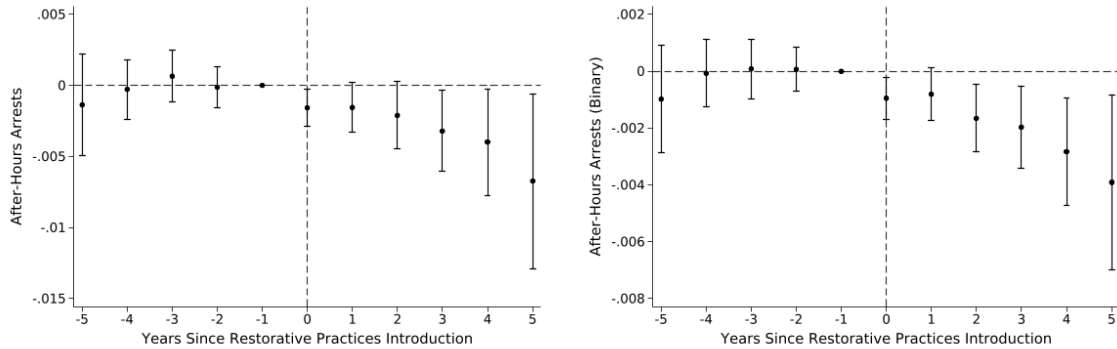
Figure A22: High School: Attrition Analysis



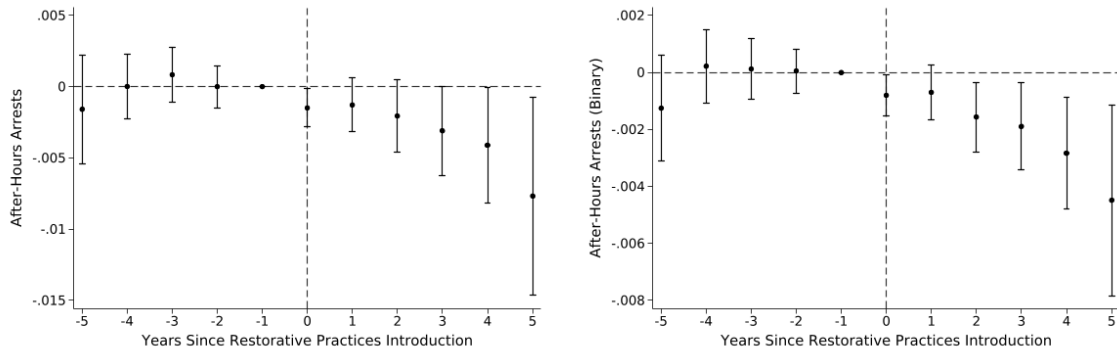
Notes: Panels A and B present the event studies for student attrition with and without controls. The sample covers students in grades 9 to 12 between SY09 and SY19. To produce these plots, an artificial panel was first constructed that includes one observation for each student in each grade level between nine and twelve under the assumption that students progressed one grade level each year. For the first student-grade observation that does not appear in our study sample due to student attrition, we code an attrition indicator variable equal to one. The attrition variable is set to missing in all subsequent years for that student. Panel A includes the following covariates: student cohort fixed effects (based on grade and school year of entry), gender fixed effects, race/ethnicity fixed effects, and disability status indicators (having a 504 plan, physical disability, or cognitive disability). Estimates in Panels A and B are based on the methodology developed in de Chaisemartin and D’Haultfoeuille (2020) and described in the text. In Panels C and D, 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. Panels C and D present the average out-of-school suspension (OSS) days and number of arrests, respectively, in a student’s first high school year by school treatment status and attrition status. In each panel, bars represent 95% confidence intervals based on standard errors clustered by school.

Figure A23: Elementary School Event Studies: Alternative Out-of-School Arrest Outcomes

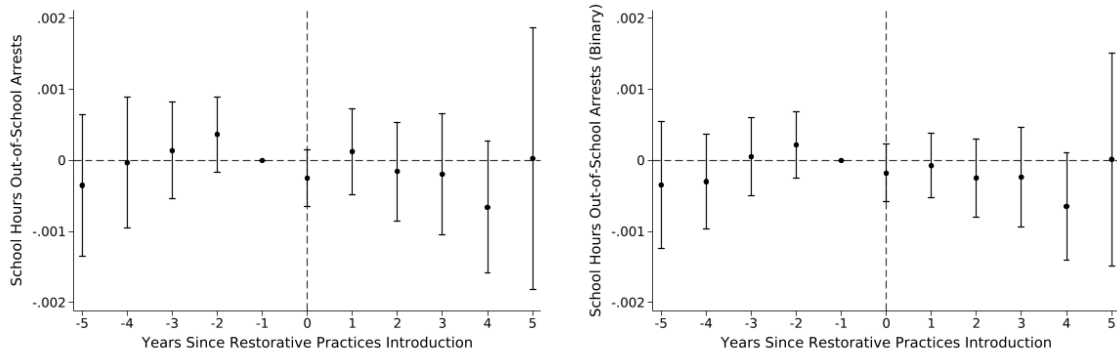
(a) Number of Arrests, Excluding 7am - 6:59pm on School Days (b) Any Arrest, Excluding 7am - 6:59pm on School Days (Binary)



(c) Number of Arrests, Excluding 7am - 4:59pm on School Days (d) Any Arrest, Excluding 7am - 4:59pm on School Days (Binary)



(e) Number of Arrests Outside School, 7am - 4:59pm on School Days (f) Any Arrest Outside School, 7am - 4:59pm on School Days (Binary)

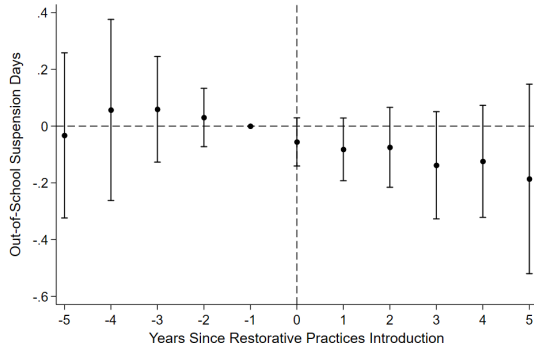


Notes: These figures show the event studies around the introduction of RP on alternative out-of-school arrest outcomes. In (a)-(d), these outcomes include all arrests taking place on non-school days as well as arrests on school days outside of the referenced time window. Observations are at the student-school 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. 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. 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/ethnicity 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.

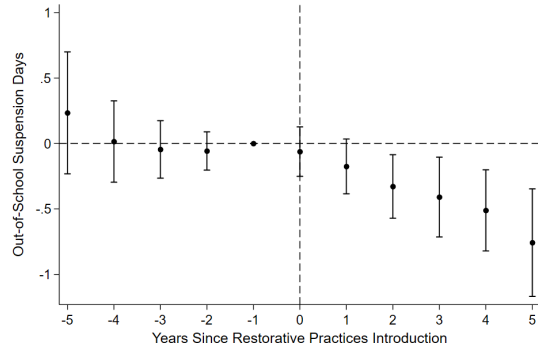
Figure A24: High School OSS Event Studies: Above- and Below-Median Predicted OSS

Panel A. All Students

(a) School*Cohort Predicted to have Below-Median OSS

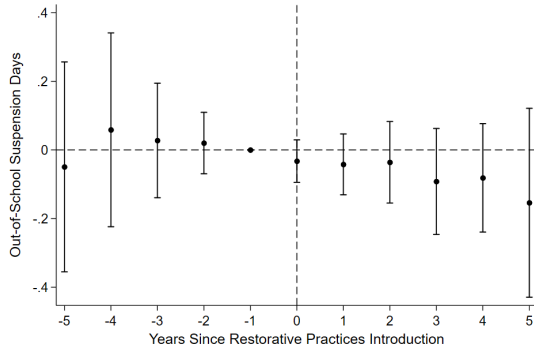


(b) School*Cohort Predicted to have Above-Median OSS

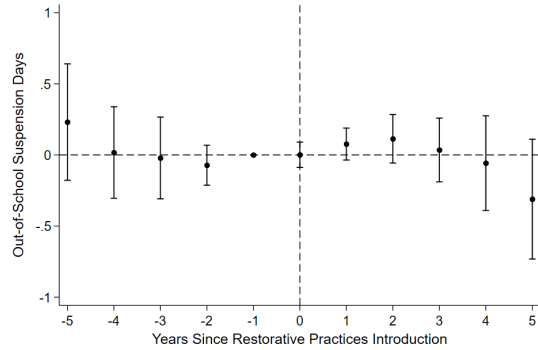


Panel B. Students Predicted to have Low-Suspension Propensity

(c) School*Cohort Predicted to have Below-Median OSS



(d) School*Cohort Predicted to have Above-Median OSS

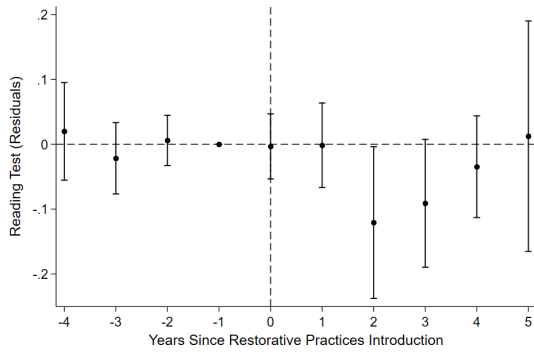


Notes: These figures show the event studies around the introduction of RP on out-of-school suspensions over time for all students and students with below-median predicted suspension days, by whether they belong to school-by-cohort cells that are above- versus below-median in predicted suspension days 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 VI. 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/ethnicity 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.

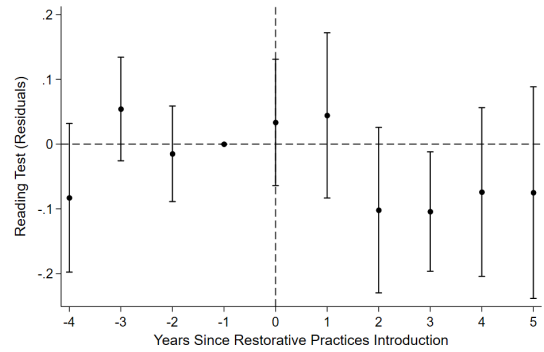
Figure A25: High School Event Studies for Learning Outcomes: Low Suspension Propensity Students

Panel A: Reading Value Added (Std.)

(a) School*Cohort Predicted to have Below-Median OSS

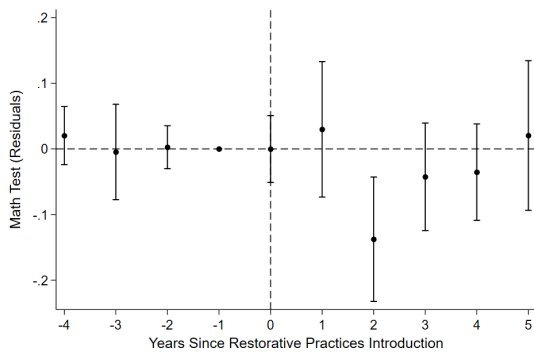


(b) School*Cohort Predicted to have Above-Median OSS

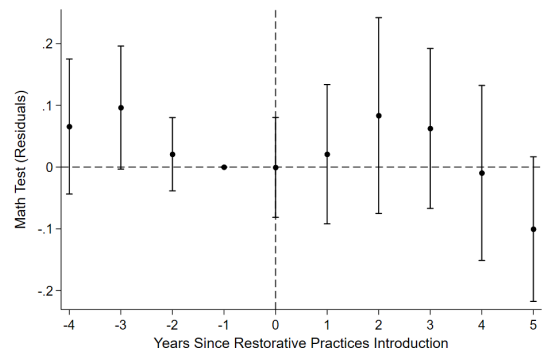


Panel B: Math Value Added (Std.)

(c) School*Cohort Predicted to have Below-Median OSS

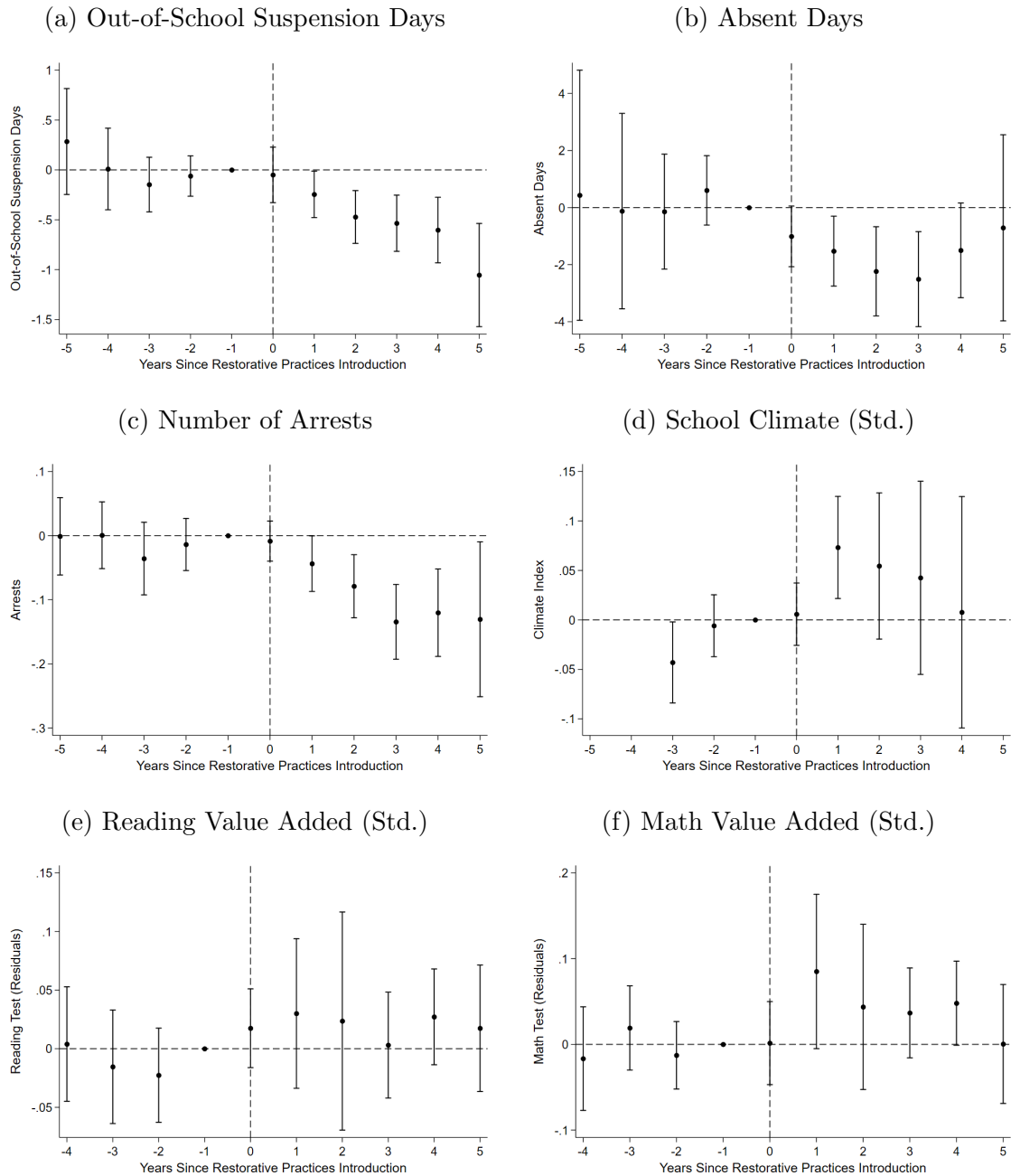


(d) School*Cohort Predicted to have Above-Median OSS



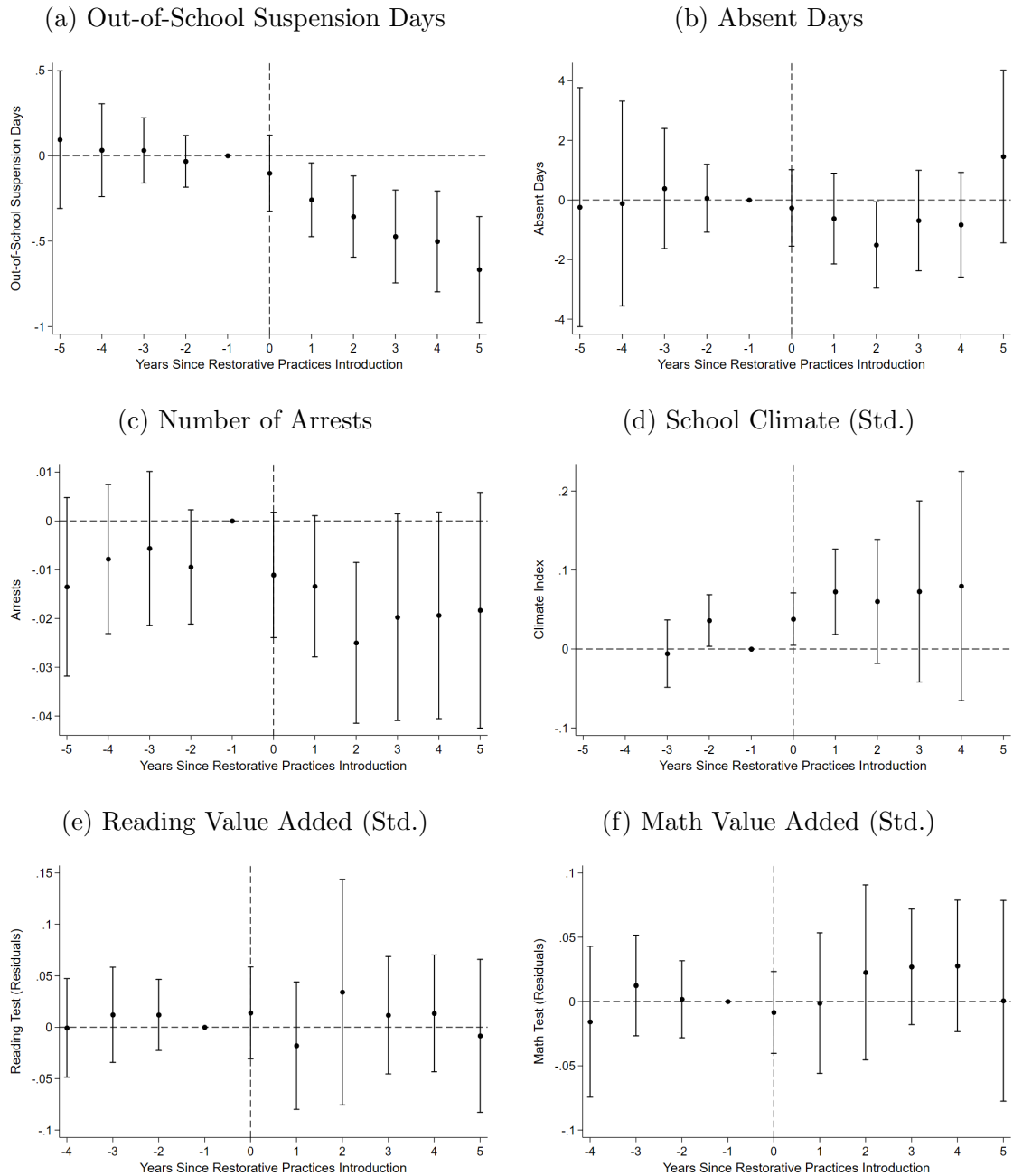
Notes: These figures show the event studies around the introduction of RP on reading value added and math value added over time for students with below-median predicted suspension days belonging to school-by-cohort cells that are above- versus below-median in predicted suspension days within a given 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 VI. 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. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. 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/ethnicity 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 A26: High School Event Studies: Black Males



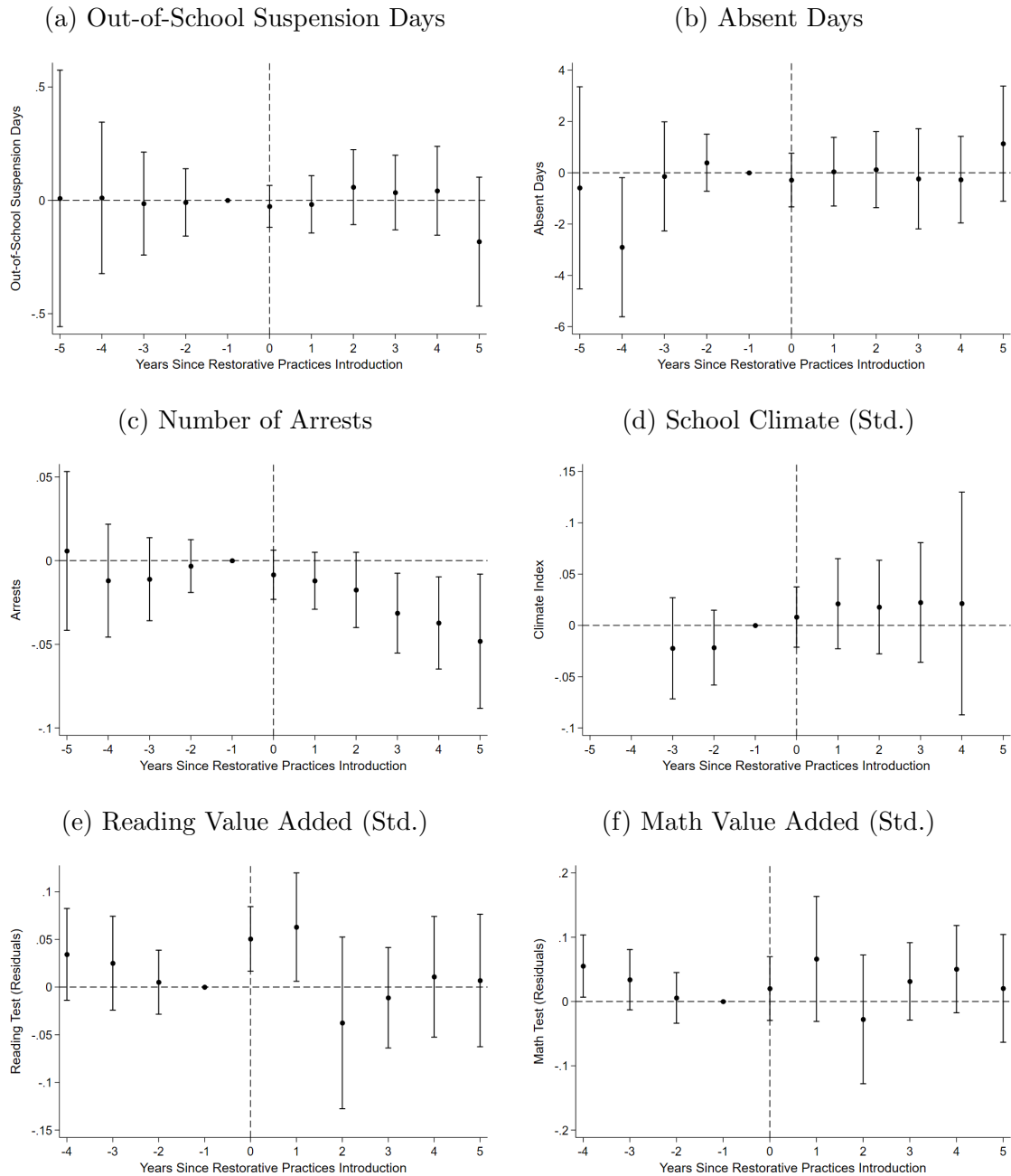
These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions and absent days), policing outcomes (overall arrests), perceptions of school climate, and academic outcomes (reading value added and math value added) among Black male students 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, 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, 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 A27: High School Event Studies: Black Females



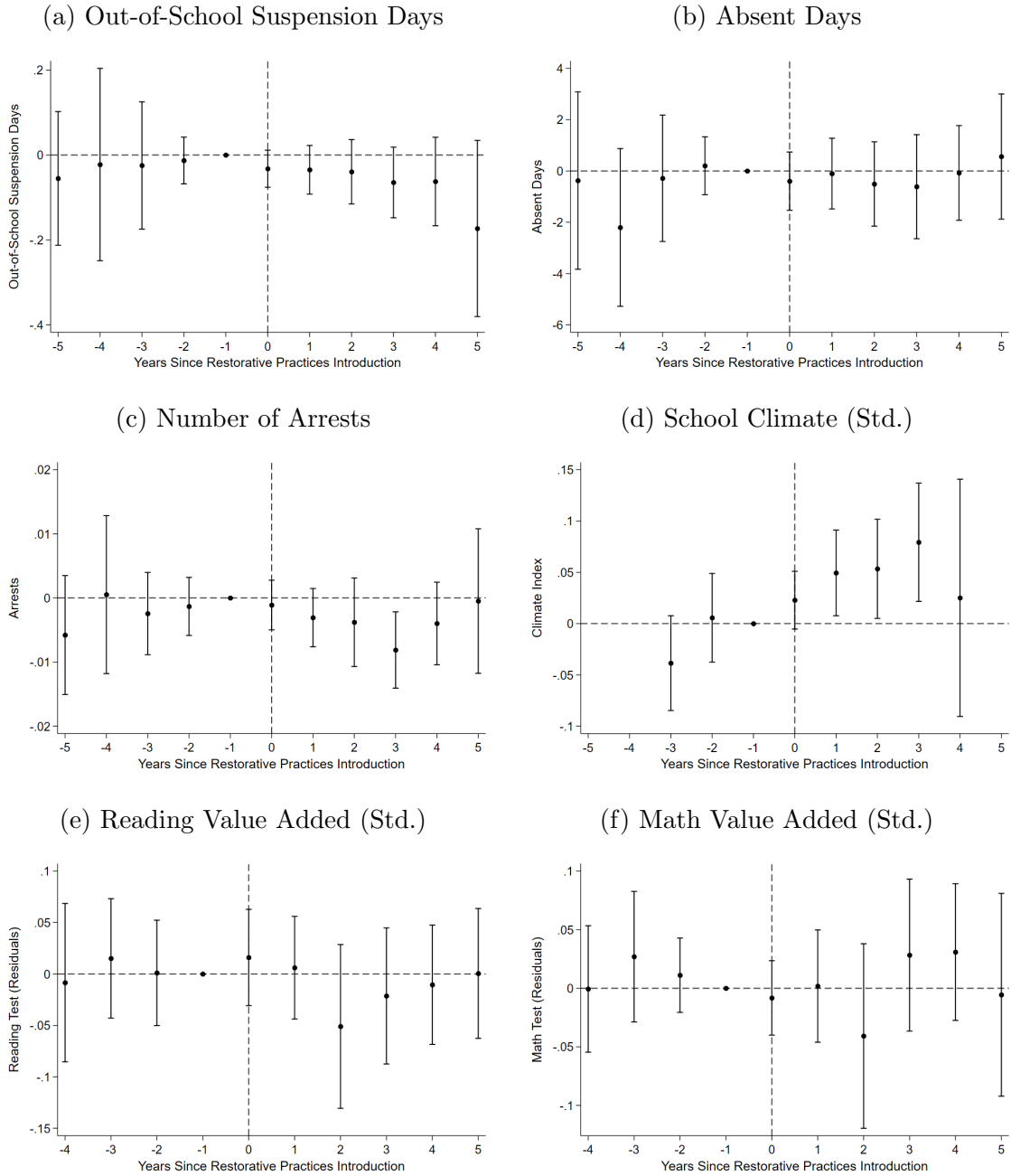
These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions and absent days), policing outcomes (overall arrests), perceptions of school climate, and academic outcomes (reading value added and math value added) among Black female students 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, 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, 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 A28: High School Event Studies: Latine Males



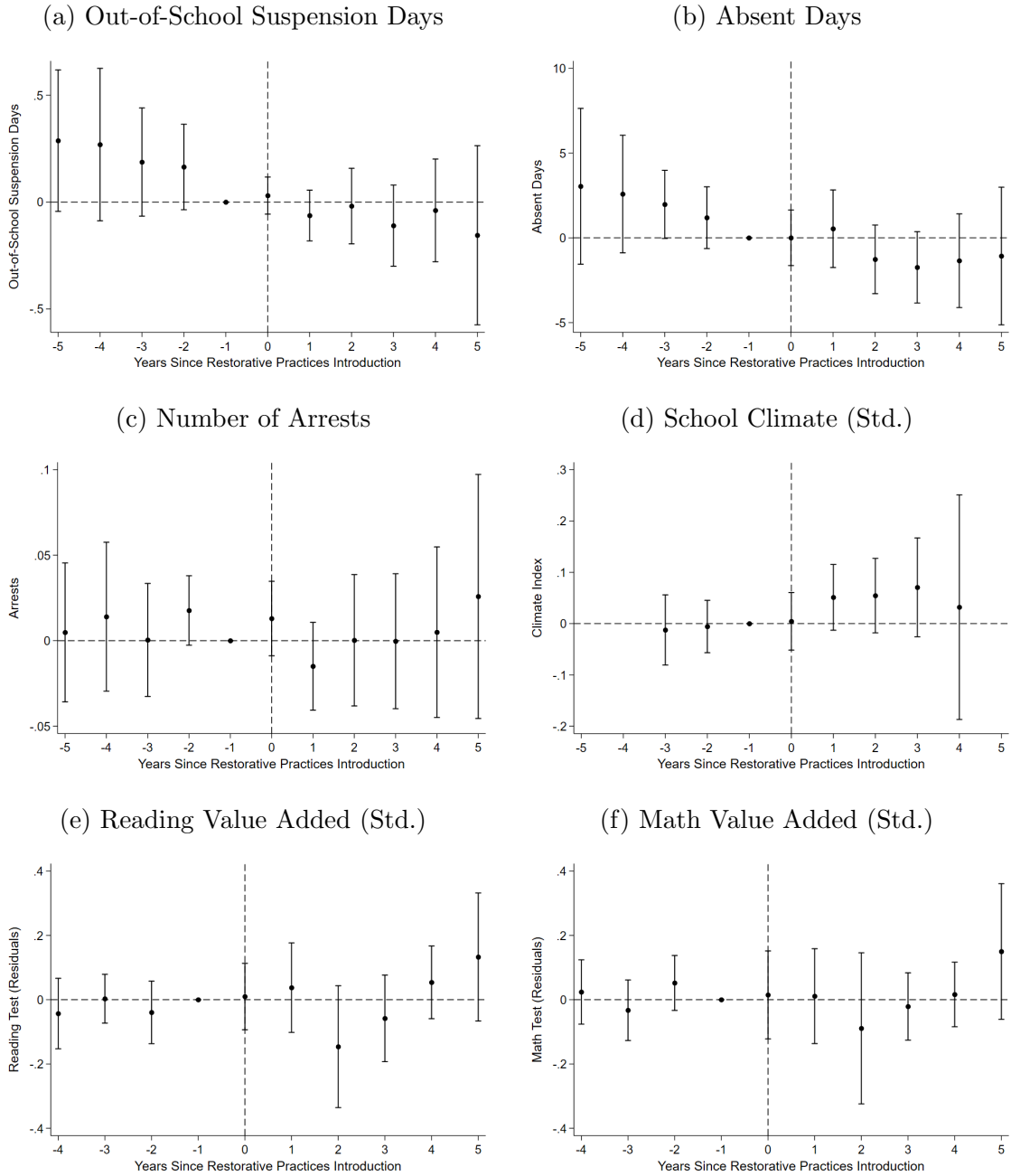
These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions and absent days), policing outcomes (overall arrests), perceptions of school climate, and academic outcomes (reading value added and math value added) among Latine male students 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, 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, 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 A29: High School Event Studies: Latine Females



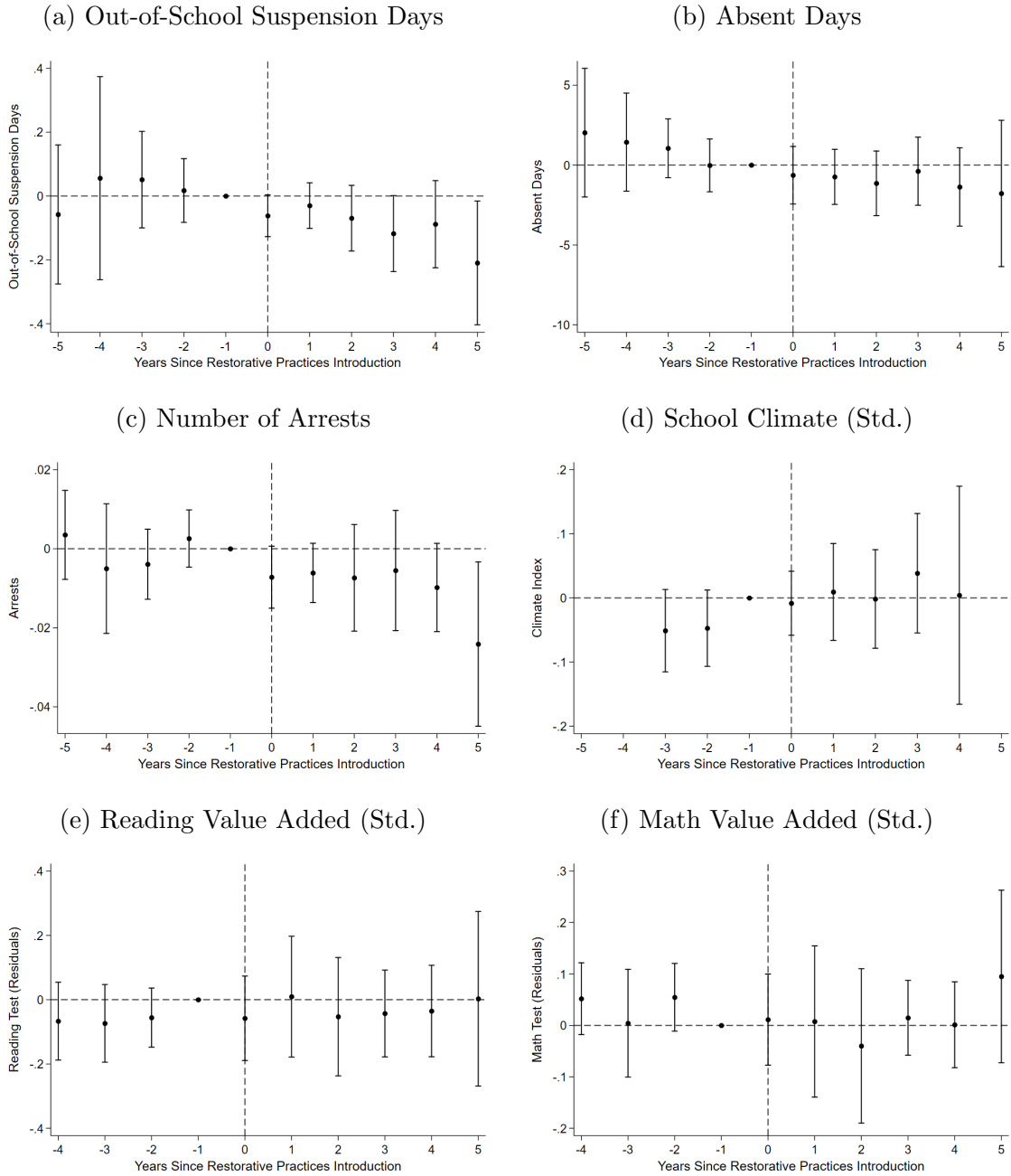
These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions and absent days), policing outcomes (overall arrests), perceptions of school climate, and academic outcomes (reading value added and math value added) among Latine female students 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, 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, 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 A30: High School Event Studies: White Males



These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions and absent days), policing outcomes (overall arrests), perceptions of school climate, and academic outcomes (reading value added and math value added) among White male students 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, 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, 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.

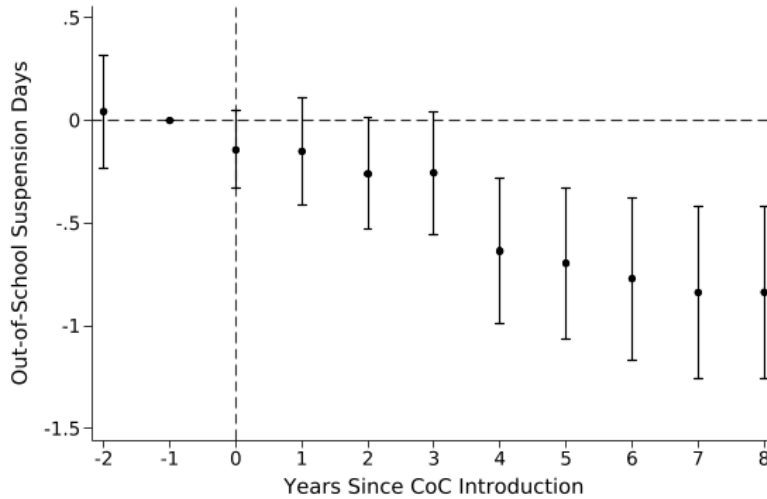
Figure A31: High School Event Studies: White Females



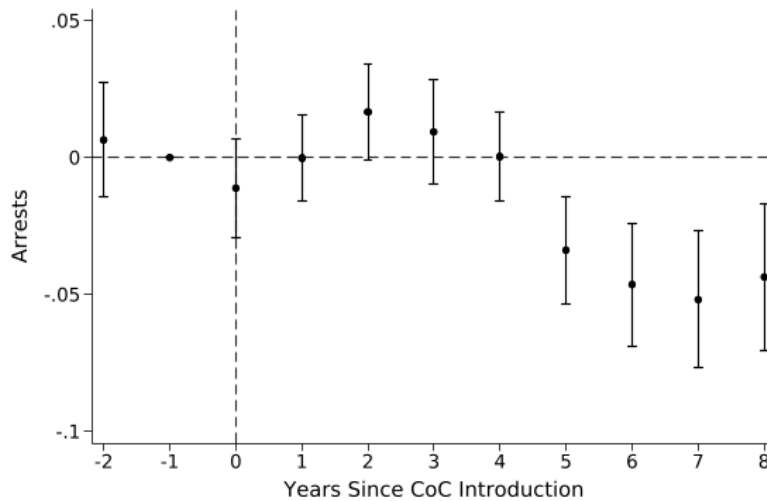
These figures show the event studies around the introduction of RP on in-school behavioral outcomes (out-of-school suspensions and absent days), policing outcomes (overall arrests), perceptions of school climate, and academic outcomes (reading value added and math value added) among White female students 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, 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, 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 A32: Culture of Calm (CoC) Event Studies: OSS and Arrests

(a) Out-of-School Suspension Days



(b) Arrests



These figures show the event studies around the introduction of Culture of Calm (CoC) on out-of-school suspension (OSS) days and number of overall arrests 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, 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/ethnicity 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.

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 was Coaching in SY14	18	18	0
First RP Received was Coaching in SY15	23	23	0
First RP Received was Coaching in SY16	5	5	0
First RP Received was Coaching in SY17	3	3	0
First RP Received was Coaching in SY18	7	4	3
First RP Received was Coaching in SY19	4	3	1
Schools whose first RP received was Leadership	4	4	0
First RP Received was Leadership in SY17	1	1	0
First RP Received was Leadership in SY18	3	3	0
Schools whose first RP received was Peer Council	8	8	0
First RP Received was Peer Council in SY14	4	4	0
First RP Received was 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 (Std.) (4)	Reading Value Added (Std.) (5)	Math Value Added (Std.) (6)
<i>Native English (NE)</i>						
Baseline Mean	0.965	18.452	0.132	0.001	0.003	0.006
Observations	96,850	96,850	99,508	65,214	57,810	57,868
<i>English Learners (EL)</i>						
Baseline Mean	0.561	17.603	0.060	-0.020	-0.068	-0.113
Observations	6,287	6,287	6,298	4,037	2,907	2,912
<i>Test (NE=EL): p-value</i>	0.000	0.001	0.000	0.032	0.000	0.000
<i>Grades 9 and 10 (G1)</i>						
Baseline Mean	1.140	16.807	0.154	0.000	0.000	0.000
Observations	56,654	56,654	58,988	39,663	41,586	41,646
<i>Grades 11 and 12 (G2)</i>						
Baseline Mean	0.696	20.343	0.095	0.000	0.000	0.000
Observations	46,483	46,483	46,818	29,588	19,131	19,134
<i>Test (G1=G2): p-value</i>	0.000	0.000	0.000	-	-	-
<i>Any Disability (AD)</i>						
Baseline Mean	1.049	21.976	0.191	-0.019	0.006	-0.025
Observations	16,655	16,655	17,181	9,930	9,196	9,226
<i>504 Disability</i>						
Baseline Mean	0.728	19.982	0.062	0.036	0.106	0.069
Observations	2,827	2,827	2,901	1,960	1,864	1,865
<i>Physical Disability</i>						
Baseline Mean	0.885	24.140	0.180	0.056	0.119	0.070
Observations	1,089	1,089	1,153	553	523	526
<i>Cognitive Disability</i>						
Baseline Mean	1.134	22.234	0.220	-0.040	-0.030	-0.058
Observations	12,739	12,739	13,127	7,417	6,809	6,835
<i>No Disability (ND)</i>						
Baseline Mean	0.922	17.750	0.116	0.003	-0.001	0.004
Observations	86,221	86,221	88,241	59,308	51,521	51,554
<i>Test (AD=ND): p-value</i>	0.000	0.000	0.000	0.002	0.321	0.000

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. The school climate index is standardized by school year and grade. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. See Data Appendix C for additional variable definitions. For each source of heterogeneity, we present *p*-values from the test of the null hypothesis that the mean for one group equals the mean for its complement (i.e., Test (NE=EL) is the test of the null hypothesis that the average variable value for native English speakers equals the average value for English learners).

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 (Std.) (4)	Reading Value Added (Std.) (5)	Math Value Added (Std.) (6)
<i>Male</i>						
Baseline Mean	1.149	18.045	0.212	0.011	0.004	-0.005
Observations	50,148	50,148	51,548	32,647	28,922	28,955
<i>Female</i>						
Baseline Mean	0.746	18.805	0.047	-0.011	-0.004	0.005
Observations	52,728	52,728	53,874	36,591	31,795	31,825
<i>Test (M = F): p-value</i>	0.000	0.000	0.000	0.000	0.123	0.033
<i>Black Male</i>						
Baseline Mean	1.880	21.010	0.384	-0.048	-0.024	-0.050
Observations	21,850	21,850	22,581	13,079	11,733	11,768
<i>Test (BM = NBM): p-value</i>	0.000	0.000	0.000	0.000	0.000	0.000
<i>Black Female</i>						
Baseline Mean	1.322	21.814	0.086	-0.107	-0.001	-0.007
Observations	23,998	23,998	24,550	15,754	13,865	13,891
<i>Test (BF = NBF): p-value</i>	0.000	0.000	0.000	0.000	0.424	0.001
<i>Latine Male</i>						
Baseline Mean	0.670	16.465	0.089	0.025	-0.012	0.004
Observations	21,228	21,228	21,632	14,653	13,002	12,999
<i>Test (LM = NLM): p-value</i>	0.000	0.000	0.000	0.001	0.000	0.019
<i>Latine Female</i>						
Baseline Mean	0.304	17.084	0.015	0.030	-0.048	-0.005
Observations	21,687	21,687	22,053	15,668	13,658	13,660
<i>Test (LF = NLF): p-value</i>	0.000	0.000	0.000	0.000	0.000	0.005
<i>White Male</i>						
Baseline Mean	0.417	15.370	0.060	0.137	0.113	0.084
Observations	4,779	4,779	4,995	3,343	2,882	2,883
<i>Test (WM = NWM): p-value</i>	0.000	0.000	0.000	0.000	0.000	0.000
<i>White Female</i>						
Baseline Mean	0.180	15.650	0.014	0.163	0.135	0.070
Observations	4,899	4,899	5,098	3,573	3,059	3,058
<i>Test (WF = NWF): p-value</i>	0.000	0.000	0.000	0.000	0.000	0.000

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/ethnicity and gender. The school climate index is standardized by school year and grade. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. See Data Appendix C for additional variable definitions. For each source of heterogeneity, we present *p*-values from the test of the null hypothesis that the mean for one group equals the mean for its complement (i.e., Test (BF=NBF) is the test of the null hypothesis that the average variable value for Black females equals the average value for non-Black females).

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

Variable	Treated (1)	Non-Treated (2)	Difference (3)
Total Enrollment	557 (308)	541 (319)	16 (27)
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 (0.24)	0.02 (0.24)	0.00 (0.00)
Ever Arrested	0.01 (0.11)	0.01 (0.10)	0.00 (0.00)
GPA	2.94 (0.70)	3.00 (0.69)	-0.06 (0.04)
Math Scores (Std.)	-0.03 (1.00)	0.02 (1.00)	-0.05 (0.04)
Reading Scores (Std.)	-0.03 (1.00)	0.02 (1.00)	-0.06 (0.04)
Math Value Added	0.005 (0.50)	-0.003 (0.51)	0.008 (0.01)
Reading Value Added	-0.004 (0.54)	0.002 (0.54)	-0.006 (0.01)
School Climate Index (Std.)	-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/Ethnicity: Black	0.44 (0.50)	0.42 (0.49)	0.02 (0.04)
Race/Ethnicity: White	0.11 (0.31)	0.10 (0.30)	0.00 (0.02)
Race/Ethnicity: 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) for the sample of students in grades 3-8. 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; value added is then constructed based on the methodology described in the text in Section IV.A. 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: High School Restorative Practices: Alternative Out-of-School Arrest Outcomes

	Number of Arrests, Excluding 7am - 6:59pm (1)	Any Arrest, Excluding 7am - 6:59pm (Binary) (2)	Number of Arrests, Excluding 7am - 4:59pm (3)	Any Arrest, Excluding 7am - 4:59pm (Binary) (4)	Number of Arrests Not at School, Between 7am - 4:59pm (5)	Any Arrest Not at School, Between 7am - 4:59pm (Binary) (6)
RP	-0.0089* (0.0039)	-0.0035+ (0.0020)	-0.0102* (0.0042)	-0.0040* (0.0020)	-0.0051* (0.0020)	-0.0034* (0.0014)
Baseline Mean	0.0641	0.0409	0.0729	0.0454	0.0295	0.0223
Observations	1,380,959	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. In columns 1-4, out-of-school arrest outcomes include all arrests taking place on non-school days as well as arrests on school days outside of the referenced time window. In columns 5-6, the arrest outcome includes arrests on school days and during school hours (7am-4:59pm) but outside of the school location. 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. 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A6: High School Student Perception Survey (“My Voice, My School”) Constructs

	RP	Observations
School Climate Index Components (Std.)		
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
Placebo Constructs (Std.)		
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 student perception survey, “My Voice, My School.” 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A7: Robustness: Behavioral Outcomes (High School)

	Out-of-School Suspension Days (1)	In-School Suspension Days (2)	Absent Days (3)	Number of Arrests (4)
<i>Panel A: Benchmark Estimates</i>				
RP	-0.167* (0.068)	-0.028 (0.068)	-0.540 (0.484)	-0.024** (0.007)
Baseline Mean	0.940	0.413	18.401	0.128
Observations	1,356,512	1,356,512	1,356,512	1,380,959
<i>Panel B: 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 C: Dropping Charter and Contract School Students</i>				
RP	-0.155* (0.076)	-0.005 (0.091)	-0.213 (0.534)	-0.022** (0.008)
Baseline Mean	0.946	0.469	19.904	0.128
Observations	873,906	873,906	873,906	893,546

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 B are based on standard difference-in-differences models with entry school and school year fixed effects. Estimates in Panels A 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 C 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 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A8: Robustness: School Climate and Learning (High School)

	School Climate (Std.)	GPA	Reading Value Added (Std.)	Math Value Added (Std.)
<i>Panel A: Benchmark Estimates</i>				
RP	0.042*	-0.024	-0.002	0.016
	(0.017)	(0.021)	(0.016)	(0.015)
Baseline Mean	0.000	2.473	0.000	0.000
Observations	751,792	851,492	421,783	421,864
<i>Panel B: Difference-in-Differences</i>				
RP	0.054**	-0.012	0.009	0.011
	(0.017)	(0.022)	(0.012)	(0.011)
Baseline Mean	0.000	2.473	0.000	0.000
Observations	578,657	865,184	465,351	465,501
<i>Panel C: Dropping Charter and Contract School Students</i>				
RP	0.032	-	-0.020	0.001
	(0.019)	-	(0.018)	(0.016)
Baseline Mean	-0.019	-	-0.005	-0.024
Observations	510,617	-	269,818	269,787

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 B are based on standard difference-in-differences models with entry school and school year fixed effects. Estimates in Panels A 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 C 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. We omit the result with GPA as the outcome from Panel C as charter and contract school students are already excluded from the benchmark GPA regression sample. 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A9: High School Restorative Practices: In-School Behavioral Outcomes (No Controls)

	Out-of-School Suspension		In-School Suspension		Absent Days
	Days (1)	Binary (2)	Days (3)	Binary (4)	(5)
RP	-0.136*	-0.022*	-0.014	0.000	-0.559
	(0.067)	(0.010)	(0.055)	(0.017)	(0.478)
Baseline Mean	0.940	0.177	0.413	0.132	18.401
Observations	1,362,075	1,362,075	1,362,075	1,362,075	1,362,075

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. 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. 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.

Regressions for the absent days outcome include student member days in the corresponding school year as a control. Data were collected by Chicago Public Schools. 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% level, * at the 5% level, and + at the 10% level.

Table A10: High School Restorative Practices: Policing Outcomes (No Controls)

	Arrests (Overall) (1)	In-School Arrests (2)	Out-of-School Arrests (3)	Violent Arrests (4)	Non-Violent Arrests (5)
<i>Panel A: Arrest Outcomes (Counts)</i>					
RP	-0.024** (0.006)	-0.009** (0.002)	-0.016** (0.005)	-0.004** (0.002)	-0.020** (0.004)
Baseline Mean	0.128	0.026	0.102	0.027	0.101
Observations	1,388,033	1,388,033	1,388,033	1,388,033	1,388,033
<i>Panel B: Binary Arrest Outcomes</i>					
RP	-0.010** (0.002)	-0.006** (0.002)	-0.007** (0.002)	-0.003* (0.001)	-0.010** (0.002)
Baseline Mean	0.071	0.022	0.057	0.023	0.058
Observations	1,388,033	1,388,033	1,388,033	1,388,033	1,388,033

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 the number of arrests (in Panel A) or an indicator for any arrest (in Panel B) 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 (outside of 7am-6:59pm on school days). See Data Appendix C for detailed variable definitions. 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% level, * at the 5% level, and + at the 10% level.

Table A11: High School Restorative Practices: School Climate and Learning Outcomes (No Controls)

	School Climate (Std.) (1)	GPA (2)	Reading Value Added (Std.) (3)	Math Value Added (Std.) (4)
RP	0.040* (0.017)	-0.031 (0.022)	0.002 (0.015)	0.018 (0.013)
Baseline Mean	0.000	2.473	0.000	0.000
Observations	752,767	853,583	421,783	421,864

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 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. The school climate index is standardized by school year and grade. GPA is calculated using semester final grades. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. See Data Appendix C for detailed variable definitions. 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% level, * at the 5% level, and + at the 10% level.

Table A12: Elementary School Restorative Practices: In-School Behavioral, Learning, and Policing Outcomes (No Controls)

	OSS Days (1)	ISS Days (2)	Absent Days (3)	GPA (4)	Reading Value Added (Std.) (5)	Math Value Added (Std.) (6)	Arrests (Count) (7)	Arrests (Binary) (8)
RP	-0.068** (0.022)	0.006 (0.005)	-0.106 (0.098)	0.006 (0.011)	0.008 (0.007)	0.009 (0.007)	-0.005** (0.002)	-0.003** (0.001)
Baseline Mean	0.401	0.054	8.497	2.970	0.000	0.000	0.018	0.011
Observations	2,540,376	2,540,376	2,540,376	2,129,867	1,807,421	1,808,004	2,551,398	2,551,398

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 2, 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. 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 3, the absent days outcome is adjusted to equal total absent days minus out-of-school suspension days. GPA is calculated using semester final grades. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. The column 7 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 column 8 arrest outcome is an indicator for any arrest experienced by students in a given year, regardless of the type of arrest or the location of the arrest. Arrest data are collected by the Chicago Police Department. See Data Appendix C for detailed variable definitions. 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% level, * at the 5% level, and + at the 10% level.

Table A13: High School Restorative Practices: Treatment Heterogeneity by Predicted Peer Group Suspension Days (No Controls)

		<i>All Students</i>			<i>Low Predicted OSS Days Students</i>		
		Out-of-School Suspension Days	Reading Value Added (Std.)	Math Value Added (Std.)	Out-of-School Suspension Days	Reading Value Added (Std.)	Math Value Added (Std.)
		(1)	(2)	(3)	(4)	(5)	(6)
	<i>School*Cohort Predicted to Have Below-Median OSS</i>	-0.073 (0.045)	-0.013 (0.018)	-0.005 (0.022)	-0.054 (0.036)	-0.001 (0.019)	-0.005 (0.023)
	Observations	569,846	182,971	182,920	478,882	155,331	155,281
A x	<i>School*Cohort Predicted to Have Above-Median OSS</i>	-0.271** (0.098)	0.016 (0.018)	0.035* (0.016)	0.009 (0.061)	-0.032 (0.027)	0.030 (0.038)
	Observations	661,448	192,518	192,532	171,399	51,722	51,725
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. See Data Appendix C for detailed variable definitions. We present results for students belonging to school-by-cohort cells that are above- versus below-median in predicted out-of-school suspension (OSS) days within a given cohort. Low predicted OSS days students are those with below-median 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 VI. 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% level, * at the 5% level, and + at the 10% level.

Table A14: High School Restorative Practices: Race-by-Gender Treatment Heterogeneity (No Controls)

	Out-of-School Suspension Days (1)	Absent Days (2)	Number of Arrests (3)	School Climate (Std.) (4)	Reading Value Added (Std.) (5)	Math Value Added (Std.) (6)
Overall Male	-0.132 (0.085)	-0.658 (0.474)	-0.039** (0.011)	0.033* (0.017)	0.011 (0.014)	0.029+ (0.016)
Overall Female	-0.126* (0.060)	-0.519 (0.447)	-0.009** (0.003)	0.048** (0.018)	-0.002 (0.017)	0.012 (0.013)
Black Male	-0.346** (0.110)	-1.715** (0.576)	-0.077** (0.021)	0.052* (0.026)	0.012 (0.015)	0.033* (0.016)
Black Female	-0.293** (0.109)	-0.849 (0.607)	-0.018** (0.007)	0.068* (0.028)	0.021 (0.019)	0.029* (0.013)
Latine Male	-0.006 (0.066)	-0.096 (0.638)	-0.019* (0.008)	0.022 (0.019)	0.013 (0.017)	0.029 (0.020)
Latine Female	-0.028 (0.033)	-0.434 (0.657)	-0.005* (0.002)	0.053* (0.021)	-0.003 (0.019)	0.010 (0.019)
White Male	-0.098+ (0.052)	-0.784 (0.740)	-0.013 (0.013)	0.018 (0.029)	0.003 (0.032)	-0.009 (0.038)
White Female	-0.081* (0.034)	-0.965 (0.787)	-0.007+ (0.004)	0.005 (0.034)	-0.034 (0.039)	0.008 (0.029)

Notes: This table shows results by student race/ethnicity and gender. 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. 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% level, * at the 5% level, and + at the 10% level.

Table A15: Synthetic Difference-in-Differences Estimates: Behavioral Outcomes

	Out-of-School Suspension Days	In-School Suspension Days	Absent Days	Number of Arrests
RP	-0.362** (0.098)	-0.066 (0.075)	-1.087+ (0.659)	-0.039** (0.011)
Baseline Mean	1.013	0.435	18.109	0.149
Observations	1,144	1,144	1,144	1,155

Notes: Observations are at the school-school year level and include only schools with non-missing observations in each school year between SY09 and SY19. We report the average effect of restorative practices over six periods. Before data are collapsed to the school-school year level, students are assigned to the first high school they had been enrolled in since SY09. See Data Appendix C for detailed variable definitions. Estimates are based on the methodology developed in Arkhangelsky et al. (2021). Standard errors are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A16: Synthetic Difference-in-Differences Estimates: School Climate and Learning

	School Climate (Std.)	GPA	Reading Value Added (Std.)	Math Value Added (Std.)
RP	0.084** (0.023)	-0.001 (0.034)	-0.006 (0.015)	0.020 (0.015)
Baseline Mean	-0.009	2.430	-0.022	-0.011
Observations	832	869	882	882

Notes: Observations are at the school-school year level and include only schools with non-missing observations in each school year between SY09 and SY19 (between SY11 and SY18 for school climate). We report the average effect of restorative practices over six periods. Before data are collapsed to the school-school year level, students are assigned to the first high school they had been enrolled in since SY09. See Data Appendix C for detailed variable definitions. Estimates are based on the methodology developed in Arkhangelsky et al. (2021). Standard errors are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A17: Student Characteristics as Outcomes (High School)

<i>Panel A</i>	STLS (Unhoused) (1)	Individualized Education Plan (2)	Free or Reduced-Price Lunch (3)	Disability (504 Plan) (4)	Disability (Cognitive) (5)	Disability (Physical) (6)
RP	-0.0031 (0.0043)	-0.003 (0.0048)	-0.0083 (0.0062)	0.0013 (0.0016)	-0.0026 (0.0045)	0.0001 (0.0009)
Baseline Mean	0.06	0.13	0.80	0.03	0.12	0.01
Observations	1,388,033	1,388,033	1,388,033	1,383,914	1,383,914	1,383,914
<i>Panel B</i>	English Learner (7)	Gender: Female (8)	Race/ Ethnicity: White (9)	Race/ Ethnicity: Black (10)	Race/ Ethnicity: Latine (11)	Predicted OSS Days (12)
RP	0.0016 (0.0046)	0.0023 (0.0039)	-0.0023 (0.0037)	0.0025 (0.0061)	0.0022 (0.0064)	-0.0038 (0.0114)
Baseline Mean	0.06	0.51	0.10	0.45	0.42	1.18
Observations	1,388,033	1,383,914	1,380,959	1,380,959	1,380,959	1,388,033

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. STLS is an indicator for Students in Temporary Living Situations. 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% level, * at the 5% level, and + at the 10% level.

Table A18: Controls for Lagged Outcomes (High School)

	Out-of-School Suspension Days (1)	In-School Suspension Days (2)	Absent Days (3)	Number of Arrests (4)	School Climate (Std.) (5)	GPA (6)
RP	-0.132+ (0.072)	-0.011 (0.067)	-0.554 (0.483)	-0.019** (0.007)	0.040** (0.016)	-0.042* (0.020)
Baseline Mean	0.940	0.413	18.401	0.128	0.000	2.473
Observations	1,201,794	1,201,794	1,201,794	1,222,942	561,474	725,845

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 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 controls for the lagged value of the dependent variable as well as 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A19: High School-by-Cohort Grouping Models

	OSS Days (1)	ISS Days (2)	Absent Days (3)	Arrests (4)	School Climate (Std.) (5)	GPA (6)
RP	-0.069+ (0.041) [0.066]	-0.064 (0.050) [0.071]	0.098 (0.319) [0.419]	-0.025** (0.0059) [0.0062]	0.040** (0.013) [0.015]	0.020 (0.021) [0.018]
Baseline Mean	0.940	0.413	18.401	0.128	0.000	2.473
Observations	486,947	486,947	486,947	489,612	298,499	304,558

Notes: Observations are at the student-school year level, and we report the average effect of restorative practices over three periods. Student treatment assignment is determined by the first high school a student had been enrolled in since SY09. The sample covers students in grades 9 to 12 between SY09 and SY19 and includes only students who appeared in grade 9 CPS enrollment records. See Data Appendix C for detailed variable definitions. Each specification includes the following covariates: student age fixed effects, ELL indicator, unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race/ethnicity 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) with school-by-cohort (as opposed to school) as the grouping variable. Robust standard errors clustered by school-cohort are reported in parentheses with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level. We also report standard errors clustered by school in brackets.

Table A20: Elementary School Restorative Practices: Additional In-School Behavioral, School Climate and Policing Outcomes

	Out-of-School Suspensions (Binary) (1)	In-School Suspensions (Binary) (2)	School Climate (Std.) (3)	In-School Arrests (Count) (4)	In-School Arrests (Binary) (5)	Non-Violent Arrests (Count) (6)	Non-Violent Arrests (Binary) (7)	Violent Arrests (Count) (8)	Violent Arrests (Binary) (9)
RP	-0.010** (0.004)	0.003 (0.002)	-0.011 (0.015)	-0.0008 (0.0005)	-0.0006 (0.0004)	-0.0028* (0.0011)	-0.0018** (0.0005)	-0.0005 (0.0005)	-0.0005 (0.0004)
Baseline Mean	0.098	0.029	-0.002	0.005	0.004	0.012	0.007	0.006	0.005
Observations	2,536,517	2,536,517	740,741	2,546,569	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 (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. In columns 1 and 2, the out-of-school suspension (OSS) and in-school suspension (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 3, 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. The school climate index is standardized by school year and grade. 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. In-school arrests are defined as incidents that happened both inside the school location and during school hours (7am-6:59pm). 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A21: Elementary School Restorative Practices: Alternative Out-of-School Arrest Outcomes

	Number of Out-of- School Arrests, Benchmark (1)	Any Out-of- School Arrest, Benchmark (Binary) (2)	Number of Arrests, Excluding 7am-6:59pm (3)	Any Arrest, Excluding 7am-6:59pm (Binary) (4)	Number of Arrests, Excluding 7am-4:59pm (5)	Any Arrest, Excluding 7am-4:59pm (Binary) (6)	Number of Arrests Not at School, Between 7am-4:59pm (7)	Any Arrest Not at School, Between 7am-4:59pm (Binary) (8)
RP	-0.0025* (0.0012)	-0.0016** (0.0006)	-0.0024* (0.0010)	-0.0015** (0.0005)	-0.0023* (0.0011)	-0.0015** (0.0005)	-0.0002 (0.0003)	-0.0002 (0.0002)
Baseline Mean	0.012	0.0077	0.0096	0.0062	0.0110	0.0068	0.0028	0.0021
Observations	2,546,569	2,546,569	2,546,569	2,546,569	2,546,569	2,546,569	2,546,569	2,546,569

111

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. In the benchmark definition, out-of-school arrests are defined as incidents that happened either outside the school location or outside school hours (outside of 7am-6:59pm on school days). In columns 3-6, out-of-school arrest outcomes include all arrests taking place on non-school days as well as arrests on school days outside of the referenced time window. In columns 7-8, the arrest outcome includes arrests on school days and during school hours but outside of the school location. 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A22: High School Baseline Outcomes by Predicted Out-of-School Suspension Days Cell (Below- vs. Above-Median)

Outcome	<i>School*Cohort Predicted Below-Median OSS Days</i>		<i>School*Cohort Predicted Above-Median OSS Days</i>		Difference (5)
	Baseline Mean (1)	Observations (2)	Baseline Mean (3)	Observations (4)	
Math Score (Std.)	0.228 (1.059)	37,467	-0.250 (0.863)	33,908	0.478** (0.132)
Reading Score (Std.)	0.185 (1.076)	37,451	-0.204 (0.863)	33,811	0.389** (0.141)
GPA	2.594 (0.979)	45,171	2.281 (0.957)	34,809	0.312** (0.079)
School Climate (Std.)	0.070 (0.634)	37,499	-0.084 (0.643)	31,706	0.154** (0.045)
Out-of-School Suspension Days	0.461 (1.954)	52,517	1.435 (3.768)	50,500	-0.974** (0.108)
In-School Suspension Days	0.179 (0.986)	52,517	0.654 (2.035)	50,500	-0.475** (0.083)
Number of Arrests	0.054 (0.383)	53,586	0.202 (0.789)	52,018	-0.147** (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 VI. 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. The school climate index is standardized by school year and grade. 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% level, * at the 5% level, and + at the 10% level.

Table A23: Within-school, Cross-group Outcome Gaps

	Out-of-School Suspension Days (1)	Number of Arrests (2)	Math Value Added (Std.) (3)	Reading Value Added (Std.) (4)
<i>Panel A: Black-White Gap</i>				
RP	-0.513* (0.224)	-0.094* (0.039)	-0.022 (0.077)	0.027 (0.073)
Baseline Mean	1.153	0.143	0.023	-0.029
Observations	466,462	477,895	121,418	121,413
<i>Panel B: Latine-White Gap</i>				
RP	0.024 (0.074)	0.010 (0.015)	-0.029 (0.053)	0.059 (0.055)
Baseline Mean	0.040	-0.001	0.023	0.012
Observations	710,647	721,767	219,321	219,356

Notes: This table displays results from weighted school-level regressions where the dependent variable is the difference between the average outcome of Black (or Latine) students and the average outcome of White students within a given school-by-year cell (weights correspond to the number of students in the given school-by-year cell and from the given racial/ethnic groups with non-missing outcomes). Each model controls for the cross-group difference in average values for 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A24: 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 (NE)</i>	-0.183* (0.073)	-0.033 (0.065)	-0.575 (0.479)	-0.026** (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 (EL)</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	0.060
Observations	91,658	91,658	91,658	91,692
Test (NE=EL): <i>p</i> -value	0.012	0.006	0.496	0.752

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), unhoused indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for native English speakers equals the estimate for English language learners. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A25: High School Restorative Practices: School Climate and Learning Outcomes by English Learner Status

	School Climate (Std.)	GPA	Reading Value Added (Std.)	Math Value Added (Std.)
<i>Native English (NE)</i>	0.045** (0.017)	-0.021 (0.022)	-0.002 (0.016)	0.013 (0.014)
Baseline Mean	0.001	2.480	0.003	0.006
Observations	701,961	793,254	392,849	392,805
<i>English Learners (EL)</i>	0.012 (0.030)	-0.053 (0.042)	0.032 (0.023)	0.055 (0.037)
Baseline Mean	-0.020	2.380	-0.068	-0.113
Observations	49,124	57,335	27,887	27,894
Test (NE=EL): p -value	0.309	0.780	0.234	0.309

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. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. 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), unboxed indicator, IEP indicator, free or reduced-price lunch indicator, gender fixed effects, race/ethnicity 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. We present p -values from tests of the null hypothesis that the estimate for native English speakers equals the estimate for English language learners. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A26: 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 (G1)</i>	-0.219* (0.087)	-0.027 (0.075)	-0.759+ (0.400)	-0.031** (0.009)
Baseline Mean	1.140	0.490	16.807	0.154
Observations	724,191	724,191	724,191	744,536
<i>Grades 11 and 12 (G2)</i>	-0.121* (0.055)	0.021 (0.047)	-0.264 (0.679)	-0.010* (0.005)
Baseline Mean	0.696	0.319	20.343	0.095
Observations	600,608	600,608	600,608	604,371
Test (G1=G2): <i>p</i> -value	0.062	0.373	0.374	0.011

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/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for students in grades 9-10 equals the estimate for students in grades 11-12. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A27: High School Restorative Practices: School Climate and Learning Outcomes by Grade Levels

	School Climate (Std.)	GPA	Reading Value Added (Std.)	Math Value Added (Std.)
<i>Grades 9 and 10 (G1)</i>	0.034 (0.020)	-0.007 (0.027)	-0.003 (0.019)	0.003 (0.018)
Baseline Mean	0.000	2.409	0.000	0.000
Observations	409,076	441,901	228,007	227,959
<i>Grades 11 and 12 (G2)</i>	0.048* (0.021)	-0.046+ (0.023)	0.007 (0.018)	0.034 (0.022)
Baseline Mean	0.000	2.546	0.000	0.000
Observations	325,810	398,334	126,565	126,576
Test (G1=G2): <i>p</i> -value	0.402	0.181	0.722	0.221

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. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. 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/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for students in grades 9-10 equals the estimate for students in grades 11-12. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A28: 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 (AD)</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	0.191
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)	-0.013 (0.014)
Baseline Mean	0.728	0.298	19.982	0.062
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)	-0.048 (0.054)
Baseline Mean	0.885	0.389	24.140	0.180
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)	-0.061** (0.014)
Baseline Mean	1.134	0.722	22.234	0.220
Observations	170,356	170,356	170,356	175,747
<i>No Disability (ND)</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	0.116
Observations	1,105,898	1,105,898	1,105,898	1,123,012
Test (AD=ND): <i>p</i> -value	0.740	0.779	0.632	0.004

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/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for students with any disability equals the estimate for students with no disability. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A29: High School Restorative Practices: School Climate and Learning Outcomes by Disability Status

	School Climate (Std.)	GPA	Reading Value Added (Std.)	Math Value Added (Std.)
<i>Any Disability (AD)</i>	0.026 (0.020)	-0.051+ (0.030)	0.039+ (0.020)	0.018 (0.023)
Baseline Mean	-0.019	2.274	0.006	-0.025
Observations	122,473	142,795	70,778	70,798
<i>504 Disability</i>	0.058+ (0.032)	-0.013 (0.047)	-0.047 (0.035)	0.012 (0.033)
Baseline Mean	0.036	2.509	0.106	0.069
Observations	32,986	40,639	18,594	18,591
<i>Physical Disability</i>	0.072 (0.077)	-0.149* (0.075)	0.013 (0.068)	-0.074 (0.087)
Baseline Mean	0.056	2.430	0.119	0.070
Observations	7,182	11,162	4,097	4,098
<i>Cognitive Disability</i>	0.023 (0.025)	-0.054+ (0.032)	0.065** (0.020)	0.018 (0.022)
Baseline Mean	-0.040	2.204	-0.030	-0.058
Observations	81,161	89,795	46,402	46,424
<i>No Disability (ND)</i>	0.042** (0.016)	-0.018 (0.020)	-0.003 (0.020)	0.022 (0.015)
Baseline Mean	0.003	2.510	-0.001	0.004
Observations	628,888	708,099	351,018	350,961
Test (AD=ND): <i>p</i> -value	0.408	0.415	0.075	0.857

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. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. 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/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for students with any disability equals the estimate for students with no disability. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A30: 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)	-0.026** (0.007)
Baseline Mean	0.987	0.429	18.713	0.135
Observations	1,236,747	1,236,747	1,236,747	1,259,492
Test (Coaching=Other): <i>p</i> -value	0.251	0.671	0.714	0.298
<i>RP Leadership</i>	0.024 (0.170)	-0.023 (0.110)	0.970 (2.190)	0.000 (0.048)
Baseline Mean	0.831	0.364	15.428	0.106
Observations	251,770	251,770	251,770	256,423
Test (Leadership=Other): <i>p</i> -value	0.140	0.970	0.569	0.464
Test (Coaching=Leadership): <i>p</i> -value	0.131	0.964	0.532	0.441
<i>RP Peer Council</i>	-0.114 (0.129)	-0.085 (0.176)	-0.777 (1.361)	-0.015 (0.016)
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/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for one RP program type equals the estimated RP impact excluding that program type (or, alternatively, that the estimate for RP Coaching equals the estimate for RP Leadership). Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A31: High School Climate and Learning Outcomes by Implementation Type

	School Climate (Std.) (1)	GPA (2)	Reading Value Added (Std.) (3)	Math Value Added (Std.) (4)
<i>RP Coaching</i>	0.036+ (0.018)	-0.023 (0.026)	0.008 (0.017)	0.019 (0.018)
Baseline Mean	-0.016	2.432	-0.007	-0.002
Observations	671,492	738,499	390,170	390,130
Test (Coaching=Other): <i>p</i> -value	0.907	0.509	–	–
<i>RP Leadership</i>	0.029 (0.041)	-0.034 (0.050)	–	–
Baseline Mean	0.109	2.638	–	–
Observations	109,684	135,468	–	–
Test (Leadership=Other): <i>p</i> -value	0.848	0.884	–	–
Test (Coaching=Leadership): <i>p</i> -value	0.897	0.922	–	–
<i>RP Peer Council</i>	0.044 (0.048)	-0.069 (0.067)	–	–
Baseline Mean	0.071	2.584	–	–
Observations	171,884	186,064	–	–

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. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. 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/ethnicity 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. We present *p*-values from tests of the null hypothesis that the estimate for one RP program type equals the estimated RP impact excluding that program type (or, alternatively, that the estimate for RP Coaching equals the estimate for RP Leadership). Value-added estimates are only available for RP Coaching; estimates cannot be produced for RP Leadership or RP Peer Council due to limited adoption years for those programs in combination with missing value-added measures in 2016. Robust standard errors clustered by school are reported with ** denoting statistical significance at the 1% level, * at the 5% level, and + at the 10% level.

Table A32: Heterogeneity by Culture of Calm (CoC): Behavioral Outcomes

	Out-of-School Suspension Days (1)	In-School Suspension Days (2)	Absent Days (3)	Number of Arrests (4)
<i>Panel A: Non-CoC Schools</i>				
RP	-0.026 (0.062)	-0.044 (0.081)	-0.140 (0.645)	-0.015* (0.007)
Baseline Mean	0.692	0.243	15.574	0.081
Observations	1,104,790	1,104,790	1,104,790	1,121,762
<i>Panel B: CoC Schools</i>				
RP	-0.102 (0.225)	0.566* (0.276)	0.138 (1.945)	-0.014 (0.021)
Baseline Mean	1.439	0.754	24.084	0.221
Observations	240,085	240,085	240,085	247,300
Test (Non-CoC = CoC): <i>p</i> -value	0.740	0.061	0.882	0.948

Notes: This table shows results by whether the Culture of Calm (CoC) Initiative was implemented in a student's high school. 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. 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

Table A33: Heterogeneity by Culture of Calm (CoC): School Climate and Learning

	School Climate	GPA	Reading Value Added (Std.)	Math Value Added (Std.)
<i>Panel A: Non-CoC Schools</i>				
RP	0.033 (0.022)	-0.033 (0.025)	0.005 (0.024)	0.015 (0.022)
Baseline Mean	0.076	2.607	0.045	0.044
Observations	610,350	641,001	349,281	349,332
<i>Panel B: CoC Schools</i>				
RP	-0.005 (0.061)	0.042 (0.050)	-0.011 (0.033)	0.047 (0.037)
Baseline Mean	-0.166	2.248	-0.112	-0.108
Observations	136,887	200,388	65,977	65,977
Test (Non-CoC = CoC): <i>p</i> -value	0.609	0.178	0.839	0.646

Notes: This table shows results by whether the Culture of Calm (CoC) Initiative was implemented in a student’s high school. 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. Math and reading scores are standardized by test, school year, and grade; value added is then constructed based on the methodology described in the text in Section IV.A. 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/ethnicity 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% level, * at the 5% level, and + at the 10% level.

C Data Appendix

Data were made available through data-sharing agreements between the University of Chicago Education Lab and the Chicago Public Schools (education data) and the Chicago Police Department (policing data). In these data sharing agreements, there are strict confidentiality restrictions for how the researchers may and may not use the data, including the prohibition to share any identified or confidential information outside the research team. In addition, only the minimum necessary members of lab staff have access to the confidential juvenile arrest records. Finally, in keeping with the Illinois Juvenile Court Act, the use of juvenile arrest records was approved by the presiding judge of the Cook County Juvenile Court, for the use of juvenile records for this purpose.

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. For grades three through twelve, out-of-school suspensions are now only permitted if a student's attendance endangers others, causes chronic/extreme interruption to others' participation in school, and prior interventions have been used. For students in kindergarten through grade two, central administration approval is required for any suspension.

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.

The canned command (`did_multiplygt`) used to produce estimates based on the methodology outlined in de Chaisemartin and D'Haultfoeuille (2020) returns an error when we attempt to estimate value-added treatment effects at the student level. This results from

the lack of value-added scores available in SY16 (see discussion of available test scores below). To circumvent this issue, we manually collapse the data to the school-by-year level and incorporate enrollment weights. This procedure is consistent with the collapsing undertaken in the first stage of the canned procedure. We confirm that we are able to reproduce the results for our main behavioral and academic outcomes when we follow the same collapsing and weighting procedure (in the instances in which estimates differ based on the canned versus manual procedure, differences never appear before the third decimal place).

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 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. We do not include PARCC scores in our analyses because PARCC was given to high school students based on their enrollment in specific courses and because PARCC 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 September 2, 2008 and September 2, 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, whether they are classified as unhoused. We measure student age as the student’s age by June 20 of the last calendar year of the school year (the last possible end date for a school year). The race/ethnicity categories include Asian, Black, Latine, White, or other races. CPS administrative files classify students as belonging to only one racial/ethnic group (e.g. Black, Latine).

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).

der, 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.