Trade-offs from Race-Neutral Affirmative Action: Evidence from Academically Selective High Schools*

Rene A. Crespin† Tyler Powell‡

September 2025

Abstract

This paper provides the first national empirical evidence on the impacts of race-neutral affirmative action (AA) admissions policies in academically selective public high schools (ASHSs). We compile a novel dataset of ASHS admissions policies from 2010 onward, linked to national school-cohort-level demographic and achievement records. Event study and difference-in-differences estimates show that the adoption of race-neutral AA policies increases low-income enrollment shares by 14 percentage points (40%) and underrepresented minority (URM) enrollment shares by 6.5 percentage points (28%). These effects are driven primarily by declines in White and "other race" student enrollment, with little effect on Asian enrollment. We identify four distinct policy types—low-SES quotas, geographic rank, class rank, and submarginal conditional acceptance—and find that low-SES quotas and geographic rank policies generate the largest increases in socioeconomic diversity. Across policy types, average test score declines are modest: treated schools experience a reduction of only 0.1σ in ACT/SAT scores. 2SLS estimates imply that a 10 percentage point increase in free lunch or URM enrollment reduces average test scores by about 0.2σ, closely mirroring predictions from simulation-based studies. We apply our results to estimate potential academic impacts from closing low-SES and URM representation gaps across the country. Our findings suggest that race-neutral AA reforms can meaningfully increase socioeconomic and racial diversity in ASHSs at the cost of relatively small reductions in average measured academic achievement, providing the first real-world evidence to benchmark against the available simulation literature.

*We thank Stacy Dickert-Conlin, Todd Elder, Maria Fitzpatrick, Scott Imberman, Michael Lovenheim, and Evan Riehl for helpful comments and discussions. Madeleine Smith provided excellent research assistance. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B200009 to Michigan State University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

†Department of Economics, Michigan State University; crespin@msu.edu.

‡Department of Economics, Michigan State University; powel282@msu.edu

Introduction

Affirmative action (AA) in elite school admissions remains one of America's most divisive policy debates. Selective educational institutions offer students access to competitive, well-resourced learning environments that are populated with high-achieving peers (Chetty et al., 2023; Dobbie & Fryer, 2014; Finn & Hockett, 2012). Persistent disparities in access to these institutions raise important questions about how to balance rigorous selectivity with racial and socioeconomic diversity, all while minimizing academic mismatch (Arcidiacono & Lovenheim, 2016). Although recent national discourse has focused on college admissions, K–12 districts with academically selective public schools have had similarly contentious debates over the past several decades on the topic of diversity.[1]

Selective K–12 schools are often regarded by their districts as a source of pride and a means to attract and maintain engagement with middle- and upper-class families in the public school system, offering advanced courses and extracurriculars typically not available in other public schools (Finn & Hockett, 2012; Hoxby, 2003). The rigorous academic preparation often leads elite colleges to view academically selective high schools (ASHSs) as reliable "feeders" (Hoxby & Avery, 2013). Moreover, these schools frequently play a role in families' educational decisions, with over 60% of eighth graders applying for admission in some districts (Feerick Center for Social Justice, 2019; Perez, 2018). Consequently, administrators are incentivized to uphold their schools' strong academic reputations while simultaneously addressing concerns about their racial and socioeconomic diversity (Cestau et al., 2017; Ellison & Pathak, 2021). In response to criticism of the lack of diversity at selective schools and limitations on the use of race for admissions, an increasing number of districts have begun to use race-neutral AA policies over the past decade. Despite the growing interest in reforming admissions, research on the potential impacts on school diversity and academic performance has been limited to simulation-based studies without use a real world policy variation.

In this paper, we provide the first empirical evidence on the effects of race-neutral AA admissions policies on the socioeconomic/racial diversity and academic performance in ASHSs across the United States, relative to status quo ASHS admissions. While most selective public schools do not change their admissions policies frequently, if at all, we leverage geographic and temporal variation in observed admissions policy changes to analyze their impacts.

We create a novel national dataset of selective public school admissions policies and their changes over time. We first gather data on the availability of ASHSs since 2010. Next,

[1] In this paper, we use the general term academically selective public schools to refer to any public high school that uses competitive admissions. In practice, this type of public school takes various names across districts. In Boston, public schools with academically-competitive admissions are Exam Schools; in Chicago, they are Selective Enrollment Schools; and in New York City they are Screened or Specialized Schools.

we use public records to search for changes in admissions policies at each of these schools. Finally, we link these data with national school-cohort-level records that capture shifts in demographic and academic performance over time.

Using this newly compiled dataset, we document geographic variation in the availability of ASHSs and patterns in their representativeness by socioeconomics and race. These selective schools are located across many parts of the US, usually in large urban school districts that serve a significant fraction of the US student population. We show that low-income, Black, and Latino students are underrepresented at ASHSs compared to their district composition. Meanwhile, ASHSs consistently overrepresent higher-income, Asian, and White students.

Next, we leverage the staggered rollout of race-neutral admissions reforms in our national dataset to estimate their demographic effects on ASHSs. About 30% of schools in our analytic sample implemented such policies. Policy implementation timing is likely not exogenous in our context because selective schools may change their admissions policies in response to new social pressures within their districts, or because of stagnation or deterioration in their diversity. Therefore, our preferred empirical approach accounts for school-specific trends prior to the policy change. Event study estimates show sharp shifts in demographic composition following policy adoption, implying that our estimates represent policy effects rather than the continuation of pre-existing trends.

Using a difference-in-differences (DID) approach that accounts for school-specific pre-trends, we find that, on average, race-neutral AA policies lead to a 14 percentage point (40%) increase in low-SES enrollment at selective high schools, compared to status quo admissions policies at control schools. We show that these policies also lead to a 6.5 percentage point (28%) increase in URM enrollment shares, which is driven by a 4 percentage point (12%) decline in the enrollment share of White students and a 2.4 percentage point decline in the enrollment shares of other race students. On average, these policy changes have little to no impact on Asian enrollment shares at selective schools.

Race-neutral AA policies that encourage diversity can vary in their actual implementation. In our data, we identify four distinct race-neutral AA policies: Low-Income Quotas require a fraction of admitted/enrolled students be from low-SES backgrounds; class rank-based policies guarantee admissions to a portion of the best-performing students from each middle school within the district; geographic rank-based admissions target admissions benefits based on a student's residential location; and Submarginal Conditional Acceptance (SCA) admits low-SES students who score marginally below admissions cutoffs and successfully attend a bridge program between 8th and 9th grade.

We find suggestive evidence of differential impacts between these different versions of race-neutral AA admissions. All four admissions policy types lead to increases in free lunch and

URM enrollment shares. However, Low-Income Quotas and Geographic Ranking admissions policies lead to significantly larger impacts on FL enrollment shares than class ranking and SCA policies. This difference may be due to Class Rank and SCA policies being academically stricter by targeting only students at the top of their eighth grade class or scoring slightly below the cutoff, while the other two policies essentially create new admissions cutoffs for students meeting non-academic criteria.

Next we provide evidence that race-neutral admissions policies have minimal negative effects on schools' average academic performance, with treated schools only experiencing a decline of 0.1σ on their average ACT/SAT scores.[2] We find heterogeneous impacts on academic achievement by type of SES-based admissions policy. Specifically, we find that SCA and Class Ranking admissions policies have smaller impacts on academic achievement than Low-Income Quotas or Geographic Ranking admissions policies. This may be due to SCA and Class Ranking policies being more directly beneficial to applicants based on quantitative academic measures than the other two policies.

In order to have comparable estimates across contexts, we next focus on the trade-offs between ASHSs' demographic diversity and average academic achievement. Using a 2SLS/IV approach, we estimate that a 10 percentage point increase in FL or URM enrollment shares leads schools' average ACT/SAT scores to decline by 0.2σ. To put this into context, our results are similar to predictions from simulation-based studies. Furthermore, we estimate similar trade-offs when we estimate a linear regression from the simulation-based predictions as data points of the relationship. Lastly, our estimates of the impact of changes in FL or URM enrollment shares on test scores allow us to consider how ASHSs' average test scores would change if they implemented admissions policies that led to their entering cohorts perfectly matching their districts' SES or racial composition.

Because our estimated test score impacts align with simulations based on incoming average achievement, our results suggest that ASHSs do not improve academic growth enough to offset declines in incoming student performance from race-neutral AA policies. This conclusion is suggestive, as our data do not directly measure changes in incoming achievement. Further research should examine how SES-based AA admissions policies influence the academic outcomes of students affected by these reforms.

Relationship to the literature. We make several important contributions to the literature. First, we contribute to a growing strand of research that examines the potential impacts of race-based and race-neutral admissions policies on the socioeconomic and racial diversity

[2]In this study, we rely on ACT and SAT scores as our primary indicator of students' academic performance. However, it is important to acknowledge that these standardized assessments may not fully reflect academic ability, as students from more affluent backgrounds often benefit from greater access to test preparation services and other supports that can boost performance (e.g., Goodman et al. (2020).)

of academically selective pre-collegiate schools. This question is increasingly important as educational institutions around the country are shifting away from purely achievement-based or race-based admissions to more race-neutral policies that prioritize low-SES students.

Despite this trend, research into the potential impacts of different admissions policies on selective K–12 school diversity is limited to theoretical and simulation-based studies. These studies hypothesize that race-neutral AA admissions policies can have a small, near zero, impact or nearly double FL enrollment shares, but can have slightly larger impacts on URM enrollment shares (Cestau et al., 2017; Corcoran & Baker-Smith, 2018; Ellison & Pathak, 2021; Idoux, 2022; Rokkanen, 2015; Rucinski & Goodman, 2022).[3] While all six studies provide important theoretical insights, they must rely on assumptions about families' preferences and admissions dynamics. We test their hypotheses by employing real-world policy variation.

Our results show that while simulations generally predict race-neutral AA impacts on low-income enrollment shares, they overestimate impacts on URM enrollment shares. This contrast in findings highlights the importance of accounting for differences in preferences and constraints in families' responsiveness or social network constraints when navigating school choice systems (Barrow, 2002; Lawrence & Mollborn, 2017; Posey-Maddox et al., 2025).[4]

Second, we contribute to a growing literature examining the distortionary impacts of admissions policies on student achievement at ASHSs. Cestau et al. (2017), Corcoran and Baker-Smith (2018), Ellison and Pathak (2021), and Rokkanen (2015) provide simulation-based predictions of the trade-offs between FL or URM enrollment changes on incoming students' academic achievement. We contribute to this work by comparing the literature's simulation-based predictions to estimates of the trade-offs based on real-world policy variation. Furthermore, we standardize and summarize existing estimates from the simulation literature to enable direct comparisons across studies and policy types.

Third, studies that examine the effects of AA policies on the demographic composition or academic achievement of ASHSs focus on ASHSs within individual school districts. More specifically, they focus on two of Chicago's Selective Enrollment High Schools (Ellison & Pathak, 2021), New York City's Specialized High Schools (Corcoran & Baker-Smith, 2018), and Boston's Exam High Schools (Idoux, 2022; Rokkanen, 2015; Rucinski & Goodman, 2022). We provide the first study of admissions policies at the national level. Among ASHSs

[3]These comparisons are not straightforward due to the varied contexts and different baseline starting points.

[4]Our study also complements a large and growing literature that examines access-oriented admissions policies on US universities' economic and racial composition. Those studies focus on the impacts of Top Percent and Holistic Review admissions on lower-income and URM undergraduate enrollment in California and Texas (Black et al., 2015; Bleemer, 2023; Cortes, 2010; Kapor, 2020; Klasik & Cortes, 2022).

in 47 school districts and 23 states, we observe 60 control ASHSs in 42 school districts and 28 treated schools within six districts.[5] This enables us to examine differential impacts from alternative AA admissions policies, allowing for variation in school and district-level implementation.

Roadmap. This paper is organized as follows: Section 1 discusses the concept of selective schools and the institutional background of SES-based affirmative action admissions policy changes. Section 2 describes the data collection process and selection of the analysis sample. Section 3 describes the empirical approach, while Sections 4 and 5 present the main results. Section 6 examines the trade-offs between ASHSs' low-income or URM enrollment and average academic achievement. Finally, Section 7 concludes.

1 Institutional Background

1.1 Where and what are academically selective public high schools?

ASHSs, a subset of magnet/choice schools, typically offer advanced courses and extracurriculars that are not available in conventional high schools (Finn & Hockett, 2012; Hoxby, 2003). Their rigorous preparation often leads elite colleges to view ASHSs as dependable "feeders", making ASHS admissions a potentially important first step for elite college admissions (Hoxby & Avery, 2013). Moreover, these schools frequently play a significant role in shaping families' school choices, with over 60% of eighth graders applying for admissions in some districts (Feerick Center for Social Justice, 2019; Perez, 2018).

ASHSs are viewed by local politicians and school district administrators as a source of pride and a way to attract and maintain engagement from middle- and upper-middle-class families with urban public school systems (Finn & Hockett, 2012; Hoxby, 2003). Figure 1 shows that these schools exist in many regions of the United States. However, consistent with the idea that they are intended to generate and maintain public school enrollment from higher-income families in lower-income areas, ASHSs are mostly concentrated within large, urban districts.

While all schools included in this study are academically selective in some way, their methods for admitting incoming cohorts vary widely. This is exemplified by two examples from different sides of the selection spectrum. Stuyvesant High School in Manhattan is a famous example of a school that takes the exam school moniker literally in its admissions system—basing admissions solely on an exam score (Stuyvesant High School, 2023). Stuyvesant is part of a subset of New York City public schools that are called Specialized

[5]For this statistic, we count all NYC school districts as a single district.

High Schools.[6] The New York City Specialized High Schools admission test (SHSAT) is offered to 8th and 9th grade students each year, and students are placed in rank order based on their scores on the exam (Dobbie & Fryer, 2014). Given the distribution of scores and the number of seats in the entering class, specialized high schools set a threshold score for admissions. Students also rank the schools that they wish to attend, so if they beat the cutoff score for their first-choice school, they are granted admission. If not, they drop to their second choice which may have a lower threshold score, and so on.

In other districts, the current admissions system eschews tests altogether, relying on a slew of other methods for evaluating academic performance and potential. For example, no standardized test scores are collected during the admissions process for any of the ASHSs in Washington, D.C. (DCPS, 2023). Instead, each school relies on its own mix of grade cutoffs, essays, recommendation letters, and interviews to determine its admissions offers.

1.2 Admissions Policy Changes

For decades, ASHSs across the US have faced criticism for not being racially/ethnically nor economically diverse (we provide more representation details in Section 2.5). Lowell High School in San Francisco, California, provides a useful case study for the controversies and forces that drive ASHSs to alter their approach to admissions. Lowell became an ASHS in 1966 after it became oversubscribed, screening students based on their GPA. In 1983, the consent decree from San Francisco NAACP v. San Francisco Unified School District mandated that no single race or ethnic group could make up more than 40 percent of enrollment.[7] By 1995, this policy led to Chinese American students being held to higher standards on the admissions test than other groups (Woo, 1995). In February 1996, the San Francisco Board of Education changed Lowell's policy to fill between 70-80 percent of admission slots based only on grades and test scores. The remaining 20-30 percent were then filled with consideration for socioeconomic status, extracurriculars, past challenges, and race (Woo, 1996). In 2001, following the settlement of Ho v. San Francisco Unified School District which disallowed race-based admissions decisions[8], Lowell first implemented a three-level admissions system that closely resembles the current policy.[9] For the 2024-25 school year, Lowell's admissions

[6] The New York City specialized high schools include: The Bronx High School of Science, The Brooklyn Latin School, Brooklyn Technical High School, Fiorello H. LaGuardia High School of Music & Art and Performing Arts, High School for Mathematics, Science and Engineering at City College of New York, High School of American Studies at Lehman College, Queens High School for the Sciences at York College, Staten Island Technical High School, and Stuyvesant High School.

[7] https://law.justia.com/cases/federal/district-courts/FSupp/576/34/2593333/

[8] https://clearinghouse.net/case/9799/

[9] https://go.boarddocs.com/ca/sfusd/Board.nsf/goto?open&id=AZS28M7BDEA9

system admits students in the same three bands.[10] First, 70 percent of seats are allocated to students based on middle school grades and score on the admissions test. Another 15 percent of seats are distributed to applicants who are evaluated based on a rubric that includes various indicators of economic and personal hardship, extracurricular pursuits, and service (students must still meet minimum grade and admissions test score requirements). The final 15 percent of seats are allocated to students who attend middle schools that are underrepresented at Lowell among the previous three Lowell cohorts (students must still meet a minimum score and must be nominated by their principal).

ASHSs' admissions vary by context. In this paper, we focus on four race-neutral affirmative action admissions policies that aim to diversify the demographics of ASHSs. First are Low-Income Quota admissions policies that prioritize students who are categorized as low-income (New York City). Second are Eligibility in the Local Context (ELC) policies, which allocate a minimum number of seats for the top-performing students from each middle school in the district (Alexandra, VA; Charleston, SC; Boston, MA; New York City; San Francisco, CA). Third are Priority Group policies that target students who meet certain special circumstances, such as low parental educational attainment, living in a low-income neighborhood, or attending a low-income middle school (Austin, TX). And fourth is a Submarginal Conditional Acceptance (SCA) policy, referred to as the Discovery Program in New York City, which allows for an alternative route to admission for disadvantaged students in the city's specialized high schools through an invitation-only summer training program. We provide a detailed description of each of these policies in Appendix B.[11]

1.3 Past Evidence on the Impacts of K–12 Affirmative Action

To our knowledge, research that examines the effects of race-neutral affirmative action policies on the demographic and academic composition of ASHSs has been limited to simulations. Table A1 summarizes the available studies. To be included, a study had to: (a) use data rather than only theoretical models; (b) focus on academically selective K–12 schools; and (c) analyze affirmative action policies that assign admissions priority based on student characteristics (as opposed to modifying application requirements such as GPA or test scores).

We identified six studies that meet these criteria through targeted Google Scholar searches using terms such as "affirmative action admissions" and "exam school admissions." The studies examine different urban settings—including Boston, Chicago, and New York City—and cover applications to 1st, 7th, and 9th grade programs. Some rely on data from a single application cohort (Corcoran & Baker-Smith, 2018; Ellison & Pathak, 2021), while others

[10] https://drive.google.com/file/d/1SGCkuFxJvcVIdaYJK8MWpvo1pLrfvyYT/view

[11] The full set of observed policy changes is described in Section 2.

use multi-year datasets spanning three to twelve cohorts (Cestau et al., 2017; Idoux, 2022; Rokkanen, 2015; Rucinski & Goodman, 2022). Across contexts, the simulations evaluate a range of admissions policies, including (a) no AA, (b) SES-based AA, (c) middle school class rank priority, (d) neighborhood-based class rank priority, and (e) race-based AA. They report impacts on both student composition (by SES and race) and average academic performance.

Throughout our paper, we situate our findings against these simulation-based studies, which currently provide the strongest available data-driven evidence on the potential impacts of race-based and race-neutral affirmative action in academically selective K–12 schools. To facilitate comparison across studies, we standardize the reported results. Demographic impacts are expressed in percentage points and percent changes, and all achievement impacts are converted into fractions of a standard deviation, regardless of the original metric.

2 Data

This study links three data sets to investigate the impacts of admissions policy changes at ASHSs. First, we use a database of ASHSs across the United States. Second, we create a database of ASHS admissions policy changes from 2010–2023. Finally, we use school enrollment data and standardized test scores to analyze the effects of policy changes.

2.1 Database of Academically Selective High Schools

In defining ASHSs and generating our list of them, we follow Finn and Hockett (2012). That is, we set six conditions that define an ASHS. Each school must be public, have a 12th grade, be a stand-alone school, teach a college-prep curriculum, select students on academic criteria, and have competitive admissions. Finn and Hockett (2012) identify 165 ASHSs in the United States. We augment this list with an additional 13 schools based on a search for historical lists of ASHSs and an assessment of current admissions policies for the top 150 high schools in the 2022 U.S. News and World Report rankings (2023).[12] We drop six schools from Finn and Hockett (2012) that are not found in the National Center for Education Statistics Common Core of Data, leaving us with a list of 172 ASHSs.

2.2 Admissions Policy Changes

We create a new panel data set of 122 admissions policy changes affecting 76 ASHSs between the late 2000s and early 2020s. We use official district and school websites to ascertain cur-

[12] Although our process of identifying ASHSs may not be exhaustive, this would not affect the internal validity of our main identification strategy.

rent school admissions policies. From there, we check for archived historical policies using the Internet Archive Wayback Machine. When available, we rely on these official admissions materials. Policy changes are recorded when there are observable adjustments to the application requirements or rubric for evaluating applications. However, when admissions policy documents are not regularly and thoroughly archived, we search local news publications for mentions of the ASHS and its admissions policy.

We find 56 instances of admissions policy changes related to SES status during our sample window from 2010–2023. A subset of these changes make up the treated units in our analysis. We denote schools as potential controls if they have no observed policy change of any type between 2010 and 2023.

2.3 School Demographics Data

We link our admissions policy data with the National Center for Education Statistics' Common Core of Data (CCD), a comprehensive database of public school enrollment by free- or reduced-price-lunch eligibility (FRPL), free-lunch eligibility (FL), and race/ethnicity (U.S. Department of Education, 2023).[13] The CCD provides race/ethnicity enrollment counts at the grade-by-school-by-year-level, which allows us to study changes at the 9th grade cohort-level. On the other hand, the FRPL and free-lunch enrollment counts data are only available at the school-by-year-level, limiting us to study socioeconomic compositional changes over time at the school-level.[14] We impute the average between the previous and following school-year information when a school has missing demographic enrollment information. From this data, we restrict the sample for analysis to include only selective schools classified as traditional schools rather than alternative schools, vocational schools, or college-based programs. This data allows us to calculate enrollment shares by socioeconomic status and race/ethnicity.

2.4 Standardized Test Score Data

We also compiled a panel dataset of school-level average ACT and/or SAT scores, a proxy for college readiness. To do this, we collected publicly available data published by state departments of education or individual school districts. We relied on the Internet Archive

[13] As suggested by the NCES, we calculate free-lunch based on the non-missing maximum between free-lunch eligible and direct certification counts (Sinclair & Chen, 2020).

[14] Through various initiatives, some schools have instituted universal school lunch programs over time, however Section 4.4 shows that our main results are robust to excluding observations and/or schools that report having 100% free-lunch eligibility.

Wayback Machine or FOIA requests when records were not directly available from official sources.

We map composite SAT scores into equivalent composite ACT scores with published concordance tables.[15] We use the ACT scoring scale because it remained consistent over time, while the SAT changed its scoring system in 2016. When both raw ACT and SAT scores are available, we attempt to analyze the primary exam for a given school and year. First, we use the exam with the most test takers for each school-by-year observation when this number is available. For schools that do not report the number of test takers, we prioritize based on score availability over time–we use the score from the exam with the least missing data. Next, when schools make clear switches between ACT and SAT as their primary exam, we allow their primary exam to change over time. We also interpolate average scores when there is a single year of missing data for the primary exam by taking the average of the one-year lead and lag. We use ACT/SAT scores to estimate college readiness for the corresponding year's school cohort of graduating seniors.[16]

We require a four-year lag to observe the testing outcome from an admissions policy change, because the reported average scores are based on the year that a cohort graduates from high school. This cuts our analytic sample of ASHSs down by about 29%. Specifically, we only have standardized test score data for 46% and 82% of the treated and control schools from our final analytic sample, respectively. Therefore, we limit our academic performance analysis to the sub-sample for which we have both demographic and ACT/SAT data.

2.5 Descriptive Statistics

2.5.1 Demographic Representation at ASHSs

Our compiled dataset allows us to illustrate the socioeconomic and racial representation at ASHSs across the US. Figure 1 shows that ASHSs are located in many parts of the United States, but are heavily concentrated in some of the largest urban districts in the country. Therefore, to appropriately describe the under- or over-representation of different

[15]For scores between 2005 and 2016 we use the concordance table found at: https://blog.prepscholar.com/act-to-sat-conversion. For scores starting in 2017, we use the concordance table found at: https://satsuite.collegeboard.org/media/pdf/guide-2018-act-sat-concordance.pdf

[16]Before the 2016-17 school year, both ACT and SAT score reporting was based on the graduating class. For example, this means that fall 2013 entrants are assigned to the average spring 2017 test scores. Beginning with the 2016-17 school year, there was a change in SAT score reporting such that, instead of reporting scores for the graduating class, averages were based on the test takers in a given year. We assume that most ASHS test takers are likely in their junior year rather than their senior year. Therefore, for cohorts starting before fall 2014 in schools where we prioritize the SAT instead of the ACT, we assign them spring test scores from three years later (their junior year). For example, for cohorts entering fall 2014 and later, we assign the SAT scores from their junior year (e.g. fall 2016 entrants are assigned to the spring 2019 average SAT scores).

demographic group, we compare the enrollment share of a demographic group at each ASHS relative to the enrollment share of that group in the school district the ASHS is located in.

ASHSs underrepresent economically disadvantaged students. Panel A in Figure 2 shows the gap between each ASHS's free-lunch enrollment share (start of the arrow) and the enrollment share of that group in the school district the ASHS is located in (end of the arrow). In our dataset, 72% of ASHSs would have to increase their FL enrollment shares to reach full representation. The average gap is -18 percentage points. Similarly, panel B shows that ASHSs underrepresent Black and Latino students, with an average gap of -20pp.

On the other hand, Asian and White students are overrepresented at ASHSs. Panel A in Figure 3 shows that 64% of ASHSs have a larger White enrollment share than the district they are located in, with an average representation gap of 9pp. Similarly, Panel B shows that 75% of ASHS have an overrepresentation of Asian students, with an average gap of 10pp.

2.5.2 Analytic Sample of Exam Schools

Our main goal in this paper is to compare treated ASHSs to ASHSs that did not implement changes to their admissions policies during our sample period. Table 1 compares our outcomes of interest across several relevant school groups during the 2015-16 school year—prior to any of the policy treatments that meet the criteria for inclusion in our analytic sample.[17]

Columns (1)-(3) of Panel A characterize our full sample of ASHSs. Column 1 describes ASHSs that appear in the NCES Common Core of Data, have non-missing enrollment-by-grade-by-race data in 2015, are classified as traditional high schools,[18] and have a 9th grade class. Column (2) averages across schools in our sample that were never treated by an admissions policy change during our sample period (2010 through 2022). Column (3) restricts to those schools that were treated by an admissions policy change that targeted prospective students by socioeconomic status.[19] Untreated and SES treated schools have similar shares of white students. However, compared to their counterparts, SES treated schools have a higher proportion of Asian students—30% compared to 16%, on average—while untreated schools have a higher share of underrepresented minority (URM) students—51% compared to 34%.

Columns (4)-(6) of Table 1 describe the main analytic sample, which excludes observations with missing race or FL enrollment data. We also restrict the SES treated schools to require that there are at least six years without a policy change before the relevant policy

[17] We discuss our decision rules for simultaneous or consecutive policy changes in appendix C.

[18] Other possible classifications include special education, vocational, alternative, and reportable programs. Schools are dropped from the sample for any year in which they are not classified as a traditional school.

[19] Twenty-four schools from Column (1) are excluded from Columns (2) or (3) because they had admissions policy changes during the sample period that were not directly related to applicants' SES.

change goes into effect. This leaves our final sample of 60 control and 28 SES treated ASHSs. We utilize an unbalanced panel in which schools are treated for anywhere between 1 and 4 periods depending on the existence of subsequent policy changes and data availability, due to policy recency—Table A2 lists the treated schools in our analysis sample.[20] Columns 5 and 6 of Table 1 show that entering classes at treated schools have relatively larger shares of Asian and White students, and smaller shares of FL or URM students than the control group.

Panel B of Table 1 presents summary statistics for the sub-sample of ASHSs for which we are able to compile standardized test score data. Columns (5) and (6) show that the treated schools, on average, have an ACT score of 31, while the control schools have an average score of 24. In Table A3, we show that this sub-sample exhibits similar demographic patterns to those of the main sample.

3 Empirical Strategy

We estimate the effects of implementing class-based affirmative action on the demographic composition and academic achievement of academically selective high schools. Our approach compares changes in the differences in outcomes between ASHSs that implement SES-based affirmative action admissions policy and ASHSs that continue with their status quo admissions during the sample period. To identify these effects, we leverage geographic and temporal variation in a staggered difference-in-difference (DID) approach.

More specifically, we estimate regressions of the following form:

$$Y_{sdt} = \alpha + \beta * AdmissionsChange_{st} * Post + \theta_t + \mu_s + \epsilon_{sdt} \quad (1)$$

where Y_{sdt} represents the outcome of interest at school s in district d in year t. Depending on the analysis, the outcome of interest is either the school's enrollment share for a specific demographic group (FL eligible, White, Asian, URM, Black, or Latino) or average ACT/SAT score. Our key variable of interest, $AdmissionsChange_{st}$, is a measure of treatment dosage in the corresponding year. Our race enrollment shares and ACT/SAT data are available at the cohort-by-school level, capturing immediate policy impacts on each incoming cohort ($AdmissionsChange_{st} = 1$ for treated schools when analyzing race- and test-based outcomes). On the other hand, our FL-eligibility data is only available at the school-by-year level. This means that in the year of the policy shock, only one out of four

[20] While 75% of our treated ASHSs are from NYC, in robustness checks, we show that our policy impact estimates are consistent with or without treated New York City ASHSs.

cohorts in a high school is treated ($AdmissionsChange_{st} = 0.25$); by the fourth year, all cohorts in the school-by-year observation are treated ($AdmissionsChange_{st} = 1$). [21] We set $AdmissionsChange_{st}$ to 1 for all treated units during the pre-policy periods. Our control group of untreated schools is "uncontaminated" in two main ways. First, in the sense that it only includes ASHSs that did not experience any policy change in our sample period ($AdmissionsChange_{st} = 0$ for every year in the sample).[22] Second, it excludes non-ASHSs that may have experienced spillover-effects, where a reform could have pushed students from ASHSs into non-ASHSs, and vice-versa.[23]

This model includes year fixed effects, θ_t, to control for any time-varying changes in school demographics that are experienced across schools. We also include school fixed effects, μ_s, which control for time-invariant quality differences between schools. The school fixed effects account for any underlying school characteristics, such as prestige or desirability, that are known by school applicants before a policy change. The inclusion of both year and school fixed effects means the estimates are identified from within-school changes in demographics after an admissions policy change.

The coefficient of interest, β, represents the DID estimate of the impact of a full-dose policy change on the outcome of interest after implementation, relative to the outcome prior to the policy change. Note that the model excludes $AdmissionsChange_{st}$ without the post indicator because, like other time-invariant school characteristics, this is absorbed by the school fixed effects. Standard errors are clustered at the district-level to account for similarity of policies and treatment timing across ASHSs within a district.

Policy implementation may not be exogenous. For example, ASHSs may decide to change their admission policies because of stagnation or deterioration in their diversity, or because of new social pressures within the district. To ensure we avoid potential bias in our estimates due to school-level variation in pre-implementation trends, we follow the literature and directly control for pre-trends by applying a two-step de-trending method (East et al., 2023; Goodman-Bacon, 2021).[24] Specifically, we estimate linear outcome pre-trends for each school using only data from before the policy change, which we extrapolate for each school through

[21] In robustness section D, we further discuss the implications and benefits of using a dose-response approach in our setting.

[22] We allow ASHSs that had multiple admissions policy changes, as long as they meet our requirement that the treatment timing does not overlap with a previous or future admissions policy change at the same school. That is, we require that the policy change must be at least six years after any previous policy change and at least one year prior to any future policy change. In practice, no school appears more than once in our analysis sample, though many schools have multiple policy changes.

[23] In robustness tests, we also exclude control ASHSs within the same district as treated ASHSs that may have experienced spillover effects.

[24] In Section 4.4, we show that not controlling for pre-trends leads to negatively biased estimates with qualitatively similar patterns.

the post-shock period. Next, we subtract the fitted values from the observed values for each outcome, removing any differential pre-trends. This does not change the estimated trend breaks; it only modifies the orientation of the event-study coefficients (Goodman-Bacon, 2021).

Two main assumptions underlie our DID approach. First, any differential pre-trends must be approximately linear to be accounted for by our specification. This assumption is analogous to parallel pre-trends in the traditional DID specification. We test this assumption by conducting event study analyses that explore the yearly policy change effect on the outcomes of interest over time, after removing school-specific linear pre-trends. These event studies provide information on whether there are any remaining nonlinear differences across ASHSs in the pre-period.

We estimate a standard event study using regressions of the following form:

$$Y_{sdt} = \lambda + \sum_{\substack{\tau=-6,\\ \tau\neq -1}}^{3} \Big(\gamma_\tau * AdmissionsChange_{st} * 1(RelativeYear = \tau)\Big) + \theta_t + \mu_s + \epsilon_{sdt} \quad (2)$$

where the γ_τ coefficients represent the change in the relationship between an eventually/fully treated unit of observation and the outcome of interest in relative period τ compared to the relationship in relative period -1, which is the reference period.

Recent econometric studies highlight potential issues with the two-way fixed effects (TWFE) model and its corresponding event studies (Roth et al., 2023). Sun and Abraham (2021) highlight that the traditional TWFE event study yields consistent results only if certain stringent assumptions about homogeneous treatment effects across treated groups and/or across time are met. Therefore, to address concerns about the robustness of our main results and to allow for heterogeneity in treatment effects across time and treated units, we present event studies that employ four distinct robust estimators introduced by Callaway and Sant'Anna (2021), Sun and Abraham (2021), Wooldridge (2021), and de Chaisemartin and D'Haultfœuille (2020). With slightly different approaches and assumptions, each of these estimators delivers consistent estimates even in the presence of heterogeneous treatment effects across time and/or treated units, by shutting down DID comparisons between newly treated and already treated units.[25]

The DID approach also assumes that there are no contemporaneous unobservable shocks that are correlated with the treatment and the outcome of interest. Although this assumption is not directly testable, our catalog of ASHS admissions policy changes allows us to exclude

[25]See Roth et al. (2023) for a more thorough description of the differences between these four methods.

schools with multiple simultaneous admissions adjustments from our analysis. For example, our sample of treated schools excludes all Chicago ASHSs because the policy change that took place in 2009 replaced a race-based AA policy with an SES place-based policy. While we would be interested in the impacts of the SES place-based policy in Chicago, the termination of the race-based admissions policy would contemporaneously impact admissions in a way that would obscure the impact of the SES-based policy.[26]

4 Impacts on Demographic Enrollment Shares

We begin our empirical analysis by estimating policy impacts on ASHSs' SES and racial/ethnic composition. We then compare our results to the existing literature and explore policy-type heterogeneity.

4.1 Impacts on Enrollment Shares, by Free-Lunch Status

Figure 4 presents an event study based on equation 2, which estimates the dynamic relationship between the activation of an SES-based AA admissions policy and FL enrollment shares at ASHSs from six years before a policy change and up to three years after implementation. Each dot represents the estimated coefficient, and the bands represent the 95% confidence interval. The dashed vertical line separates the pre- and post-policy shock periods. Because the outcome of interest is a measure between 0 and 1, each γ_τ represents the percentage point difference in the relationship between the eventual or implemented admissions policy change and the outcome of interest in period τ, relative to the most recent pre-shock year ($\tau = -1$).

The event study shows that after accounting for differences in school-specific linear pre-trends the estimates of the relationship between relative time and FL enrollment shares are close to zero and relatively flat during the pre-shock period. Thus, the figure provides supporting evidence for the linear pre-trends assumption, suggesting negligible remaining non-linear differences across treatment and control ASHSs in the pre-period. The implementation of an SES-based AA admissions policy change in period 0 is associated with a visually clear and significant increase in FL enrollment shares. The abrupt increase makes it clear that our estimates represent policy effects rather than the continuation of a pre-existing trend.[27]

[26] We discuss cases and our decision rules for simultaneous/consecutive policy changes in appendix C.

[27] As a robustness check, in section 4.4.2, we show that event studies based on robust estimators that allow for heterogeneity in treatment effects across time and treated units have very similar dynamics to those estimated by our main approach.

Panel A in Table 2 presents analogous DID estimates based on equation 1 for our overall sample. Column (1) show that controlling for school-specific pre-trends, there is an average increase of 13.5 percentage points (pp) in FL enrollment shares in response to a full dose of treatment, which is statistically significant at the 5% level. These results are for an unbalanced sample that includes all control schools and treated schools with one to four post-shock observations. Taking these estimates at face value, they suggest a 40% to 46% increase in FL enrollment shares. These findings show that the admissions policy changes are effective at increasing the representation of the lowest-income students at ASHSs.[28]

The forest plot in Figure A1 presents our estimated overall policy impact along with race-neutral (filled points) and race-based (hollow points) simulation-based predictions. We estimate a slightly larger policy impact on FL enrollment shares in terms of percentage changes relative to the median simulated impact for race-neutral (dashed vertical line) and for race-based (dotted vertical line) admissions policies. However, these differences are not statistically significantly different from each other.[29]

4.2 Impacts on Enrollment Shares, by Race/Ethnicity

The evidence above strongly suggests that the implementation of SES-based affirmative action policies had the first-order effect of improving the representation of low-income students, as proxied by FL eligibility. While these policies were intended to be race-neutral, there may be differential impacts by race/ethnicity due to the widely documented strong correlation between family income and race/ethnicity (Chetty et al., 2020).

Panel (a) of Figure 5 presents event study estimates for 9th grade cohorts' URM enrollment shares before and after SES-based admission policy implementation, defined as the sum of Black and Hispanic enrollment shares. Accounting for differences in linear pre-trends, there are no remaining differential pre-trends between treatment and control ASHSs. In treated ASHSs, the period following the policy change exhibits a clear break in trend, with an immediate and statistically significant increase in URM enrollment shares within the first entering 9th grade cohort that is admitted under the new admissions policy.

Panels (b) through (d) in Figure 5 present event studies of the estimates for enrollment shares in 9th grade cohorts over time, by non-URM groups. Panel (b) shows that the increase in URM enrollment shares, discussed above, is almost entirely driven by an abrupt

[28] In Appendix Table A4, we limit the analysis to include all pre-shock periods but only the initial post-shock period, which ensures a balanced sample. Using this more limited analysis sample, column (1) shows a slightly larger, but less precise, impact of about 15.6pp in FL enrollment shares.

[29] Appendix Figure A2 shows that, in terms of percent change, our estimate is very similar to, although slightly lower than the median of the race-neutral and race-based simulations.

and significant decline in White enrollment shares. Importantly, panels (c) and (d) show no significant change in Asian and other race enrollment shares.

Table 2 summarizes the corresponding DID results based on equation 1. Panel A shows the results for the unbalanced sample across the four post-shock periods. Column (1) shows that the SES-based admissions policies lead to a 6.5 percentage point increase in URM enrollment shares, which is statistically significant at the 1% level—a 28% increase relative to the mean. On the other hand, there is a 4 percentage point decline in White enrollment shares (12%), which is not statistically significant. Column (3) shows that, on average, there is also insufficient evidence to reject the null hypothesis that SES-based admissions policies have no impact on Asian enrollment shares. The URM and White coefficients do not exactly counteract each other because there is a slight decline in other race enrollment shares of 2.4pp, which is not statistically significant.[30]

The forest plot in Figure A3 shows that while our estimated overall policy impact on URM enrollment shares are positive and statistically significant, it is at the lower-end of the distribution relative to simulation-based predictions. Our estimate is significantly smaller than the median predictions from race-neutral and race-based AA simulations.[31] This difference may be due to simulation studies relying on various assumptions, including not accounting for potential differences in families' responsiveness or social network constraints when navigating school choice systems (Barrow, 2002; Lawrence & Mollborn, 2017; Posey-Maddox et al., 2025).

4.3 Heterogeneity by policy type

Table A5 displays the estimated policy impacts on FL enrollment shares (panel A), URM shares (panel B), White shares (panel C), Asian shares (panel D), and other race enrollment shares (panel E) for each of the four policy types (shown in each column).[32] Panel A shows that the Low-Income Quotas and Geographic Ranking policies had large and statistically significant impacts on FL enrollment shares, while middle school class ranking and the SCA policies had less half as large of impacts and were not statistically significant. This difference may be due to Class Rank and SCA policies being academically stricter by targeting only students at the top of their eighth grade class or scoring slightly below the cutoff, while the other two policies essentially create new admissions cutoffs for students meeting SES

[30] Table A4 shows the patterns are qualitatively similar for the balanced sample.

[31] Appendix Figure A4 shows a similar pattern in terms of percent change.

[32] In appendix B, we provide more details on each of the SES-based admissions policies that were implemented by ASHSs in our analytic sample.

criteria.[33]

Panel B of Table A5 shows that the estimated impacts on URM enrollment shares from each of the four policy types are not statistically different from each other.[34] Table A5 further shows suggestive evidence that The Low-Income Quotas and Geographic Ranking admissions policies' positive impacts on URM enrollment shares are driven by declines in White enrollment shares. However, we find suggestive evidence that the Class Rank and SCA policies' impacts on URM may be driven by declines in Asian and other race enrollment shares.

In terms of percentage changes, we estimate policy impacts on FL enrollment shares ranging between 50% to 62% for the LIQ, Class Rank, and Geographic Ranking policies, but only a 20% effect for SCA.[35] Comparing between the impacts on FL and URM effects, the LIQ and Geographic ranking policies have half the impact on URM than what they had on FL enrollment shares. On the other hand, Class Ranking and SCA policies have more similar impacts on URM and FL enrollment shares.

Our estimates notably differ from smaller policy impacts observed at the college level. For example, Kapor (2020) estimates a 5.4% effect from the Texas' ELC admissions policy on enrollment of students from high-poverty schools, while Bleemer (2023) finds a negligible effect from California's ELC policy on low-SES student enrollment. Moreover, our estimates of the effects on URM enrollment are larger than those documented in the college context (Bleemer, 2023; Kapor, 2020). These contrasts underscore the importance of examining SES-based AA in selective high schools, where such policies seem to be more effective than similar initiatives in college admissions.

4.4 Robustness

4.4.1 Dose-Response vs. Traditional DID and Event Study Estimates

We now test the sensitivity of our main findings by conducting a series of robustness checks. First, our results are robust to not incorporating a dose-response to our DID approach. As described in section 2, our socioeconomic enrollment data is only available at the school-by-year-level. Our main empirical approach incorporates dose-response to address the fact that our socioeconomic unit of observation is only partially impacted in the first three years that

[33]The forest plot in Appendix Figure A5 shows that the simulation-based median for each policy type is within the 95% CI of our corresponding estimated impact on FL enrollment shares, with the exception of our LIQ estimate, which is statistically significantly larger than simulation-based median at the 5% level.

[34]The forest plot in Appendix Figure A6 shows that our estimated policy impacts on URM enrollment shares are significantly lower than the median of their corresponding simulated predictions.

[35]Appendix Figures A7 and A8 shows similar patterns relative to the literature in terms of percentage change.

the policy is implemented. With proper scaling, this approach is mathematically equivalent to the traditional approach without dosage. Furthermore, our preferred approach provides estimates that are more easily interpretable and comparable than the traditional DID approach, given the limitations of the socioeconomic enrollment data.[36] We provide more details in Appendix D.

4.4.2 Robust Estimators for Staggered DID

Second, our estimates are robust to bias from heterogeneous treatment effects across groups or time. Sun and Abraham (2021) highlight that the traditional TWFE event study yields consistent results only if certain stringent assumptions about homogeneous treatment effects are met. To allow for heterogeneity in treatment effects across time and treated units, we present event studies that employ robust estimators introduced by Callaway and Sant'Anna (2021), Sun and Abraham (2021), Wooldridge (2021), and de Chaisemartin and D'Haultfœuille (2020).

Figure A9 shows that for all our outcomes of interest the pre-trends and post-shock effects are quite consistent across the four robust estimators as well as the original two-way fixed effects specification, suggesting that the underlying assumptions are not strongly influencing our results. The robust estimators cannot accommodate dose-response DID, therefore, the event study for FL enrollment shares in panel (a) is estimated in the traditional way, without dosage. As described in Appendix D, FL event studies that do not incorporate dosage will differ from our dose-response estimates in periods +0, +1, and +2, however, these results are equivalent with proper rescaling.

4.4.3 Robustness to choice of sample and specification

A third potential concern is that our estimated effects are driven by only New York City schools.[37] To test this, we replicate our analysis on two separate samples. First, we exclude all treated New York ASHSs. Then, we include only treated ASHSs from New York, with the same set of control schools used in both samples. Appendix table A6 presents quantitatively and qualitatively similar estimates of the impact of implementing SES-based AA on the low-income and racial composition at non-NYC ASHSs (in panel A) and at NYC-only ASHSs (in panel B). However, due to the smaller sample size, the estimates are noisier and have larger standard errors.

[36] The dose-response DID approach estimates coefficients that are interpreted in the same way for both socioeconomic outcomes (school-by-year-level data) and for race/ethnicity outcomes (grade-by-school-by-year-level data).

[37] As shown in table A2, the majority of treated ASHSs in our analytic sample are in New York City.

The only coefficient that changes direction is the impact on Asian enrollment shares, which is negative when focusing on non-NYC treated ASHSs and positive when limiting to NYC treated ASHSs. However, the standard errors on these estimates are larger than the coefficients in both cases—still failing to reject the null hypothesis. This test suggests that the impacts we find are broadly representative of ASHSs, and not driven solely by NYC ASHSs.

Next, Appendix Table A7 shows the effects of implementing an SES-based AA admissions policy for various sample restrictions. Column (1) shows the main results for comparison. To address the concern that perhaps our main results are driven by pandemic-era admissions policy changes, columns (2) and (3) show that we find similar estimates when we limit to pre-pandemic treated units and drop all observations before 2020 or before 2021. Furthermore, through various initiatives, some schools, districts, and states have instituted universal free lunch programs, potentially impacting some of the schools in our sample. In columns (4) and (5) we estimate similar coefficients when we exclude all observations that report 100% FL eligibility or all ASHSs that ever report having a 100% FL eligibility.

Another potential concern is that in districts with multiple treated ASHSs, admissions policy changes in the most in-demand ASHS may have trickle-down effects for ASHSs that are less in-demand. If higher-ranked ASHSs admit more under-represented students, lower-ranked ASHSs may not be able to attract as many AA-targeted students, whether or not they also instituted an SES-based AA admissions policy. To address this concern, in column (6) of Table A7, we keep only the ASHS with the highest average ACT scores if there are multiple ASHSs in the district.[38] Additionally, in column (7) we drop all control ASHSs in districts with treated ASHSs from the sample to address the possibility that untreated ASHSs have their cohort composition impacted by the admissions policies at treated schools. In both subsamples, we estimate similar, although slightly larger, coefficients compared to our main sample.

Next, in Table A8 we estimate similar coefficients when we control for year-to-year district-level racial composition changes (column 2) or year-to-year district-level income composition changes (column 3). This provides evidence that we are estimating school-specific effects rather than district-wide demographic changes. Additionally, in Table A9 we show that using a stacked DID specification, as proposed in Wing et al., 2024, yields similar demographic results (shown in column 2). Lastly, in column (3), we show that failing to account for school-level enrollment pre-trends yields misleading smaller estimates of the policy impacts.

[38]Effectively, this drops non-Stuyvesant ASHSs in NYC and non-Boston Latin ASHSs in Boston.

4.4.4 Secular Trends and Reversion to the Mean

Another potential concern is that our estimated policy impacts reflect secular trends or reversion to the mean that would have happened even without policy implementation. To investigate this, we conduct placebo tests where we randomly assign treated ASHSs to implementation years and test whether the new estimated effects are as large as our baseline policy impact estimates. In Appendix table A10, we distribute the treated ASHSs equally across our implementation years between 2016 and 2022.

We randomize placebo treatment dates 100 times for each outcome and calculate the percent of cases that generate policy impact estimates larger than our baseline estimates. Therefore, these results represent p-values for the null hypothesis that any combination of policy implementation dates would generate as extreme of policy impact estimates as our results. In table A10, we reject the null of such a hypothesis for estimates of Free Lunch share, URM share, and White enrollment share. We do not reject the null of such a hypothesis for Asian or other race enrollment shares. This is consistent with our main estimates for Asian and other race enrollment shares because they are not statistically significant.

These results suggest our main estimates are not identified based on secular trends. Although reversion to the mean may play some part in enrollment share changes, these placebo tests show that our main estimates are strongly linked to whether an SES-based admissions policy is implemented and to the specific timing with which the policy is implemented. These findings support the validity of our estimation strategy. Furthermore, if reversion to the mean is playing a role in our estimates, this further suggests that we are estimating a lower bound of the actual policy impact.

5 Impacts on ACT/SAT Scores

Next, we examine SES-based AA admissions policies affect average test scores, as proxied by the ACT/SAT, at the cohort-level within schools. To do this, we focus on the sub-sample of ASHSs for which we have standardized test score data, as described in Section 2. Since the scores are based on cohorts taking the test during their senior year, we require a four-year lead to observe outcomes with the correct timing. This reduces our sample by about 29%, leaving standardized test score data for 46% of treated schools and 82% of control schools in the final analytic sample. Nonetheless, we provide supporting evidence for the generalizability of our results on test scores.

5.1 Results

Panel B in Table 2 shows that the test scores sample exhibits very similar compositional impacts to those from the main analytic sample, with an increase of 14.1 percentage points (41.6%) in FL shares and a 5.6 percentage point (24.1%) increase in URM enrollment shares, both of which are statistically significant at the 1% level. Asian enrollment shares have a positive coefficient of 0.1pp, but this is not statistically significant. Again, the gains in URM enrollment shares come from declines in White and other race enrollment shares.[39]

The event study in Figure 6 shows that even with significant gains in FL and URM enrollment shares, the SES-based admissions policies have a small impact on the standardized test scores of the corresponding cohorts at treated ASHSs. Panel A of Table 3 shows that the average ACT score for the treated cohorts decreased by 0.53 points, or 1.7%, which is not statistically significant. Using the 95% confidence interval, our estimate rules out an increase in average ACT scores of greater than +0.11 points and rules out declines greater than -1.17 points.

In more general terms, we estimate an overall policy impact of -0.10σ, which is smaller than the majority of simulation-based predictions.[40] The forest plot in Figure A10 shows that our estimated effect is less negative than the median impact for race-neutral and race-based admissions policies. However, only the race-neutral median falls below our 95% confidence interval.

Table A11 presents the estimated effects on test scores split by policy type.[41] Panel A shows that low-income quotas had the most negative effects on average test scores, with a decline of about 1.4 ACT points, or 0.28σ. However, the other two policies have negligible effects on test scores that are not statistically significant. The forest plot in Figure A12 shows that the simulation-based median for each policy type is within the 95% CI of our corresponding estimated impact.[42] We estimate that the SCA policy has the smallest impact on test scores, but we do not have simulation-based predictions for this type of policy to compare to. Given differential impacts on FL and URM enrollment shares examined in Section 4, we present a more comparable academic impact metric, in Section 6, that accounts for differential demographic changes in Section 6.

[39] We present the corresponding event studies in Figure A11.

[40] To make estimates comparable to each other, we convert all impacts to be relative to their standard deviations.

[41] We cannot separately estimate effects for geographic ranking admissions policies because, in our sample, these took place too recently to capture test score outcomes data.

[42] More specifically, the 95% confidence interval for our low-income quotas estimate captures the estimates of nine out of 13 simulation-based predictions. Similarly, the confidence interval for our class-ranking policy estimate captures two out of three simulation predictions.

5.2 Robustness

While our sample of treated schools with ACT/SAT data is smaller than our overall sample of ASHSs, we conduct robustness tests to examine the consistency of the results. Figure A13 shows that the pre-trends and post-shock effects are quite consistent across four robust estimators as well as the original two-way fixed effects specification, suggesting that the underlying assumptions are not strongly influencing our results.

Since five of the seven treated ASHSs with available test score data are in NYC, one may be concerned that the sample choice may be more impactful for the test score results. In Table A12, we re-estimate the test score regressions for separate samples of treated ASHSs in and outside NYC, again with the same set of control schools used in both samples. In column (2), we find that the SES-based AA outside NYC led to a decline of 0.91 (3.2%) on average ACT scores, which is statistically significant at the 10% level. At treated NYC schools, we estimate a 0.42 (1.4%) decline in scores, which is not statistically significant (shown in column 3). Therefore, our estimates remain consistent with our main findings that SES-based admissions policies have a small or negligible impact on ACT scores. Moreover, in column (4) we show that a stacked DID specification leads to a decline of 0.999 points, or 3.5%, which is still within the 95% confidence interval of our main estimate.

6 Trade-offs between Diversity and Test Scores

The policy impacts on test scores presented in Section 5 do not account for the differential impacts on FL and URM enrollment shares. Therefore, in this section, we focus on the trade-off between comparable changes in diversity and average academic achievement.

6.1 Impacts from demographic change on average test scores

We employ an instrumental variables (IV) approach to estimate the effects on average test scores from a change in FL or URM enrollment shares. Panels B and C in Table 3 present 2SLS regression estimates of the trade-off between average test scores and demographic enrollment shares, using the SES-based admissions policy shocks as the IV. Overall, the 2SLS estimate in panel B indicates that increasing FL enrollment shares by 10 percentage points decreases ASHSs' average ACT score by 0.96 point or 0.18σ. Similarly, Panel C shows that increasing URM shares by 10 percentage points decreases schools' average ACT score by 0.94 points, or 0.18σ.

The forest plot in Figure A14 shows that we estimate a smaller negative effect on test scores from an increase in FL shares than the majority of simulation-based predictions.

However, Figure A15 shows that our estimate of the impact of an increase in URM shares on test scores is right at the median of the distribution of race-neutral simulation-based predictions.[43]

An alternative approach to estimate the effect on average test scores from a change in FL or URM enrollment shares is to directly estimate a linear regression from the simulation-based predictions as data points of the relationship. In Figure 7 we plot the raw, standardized effect of each policy on average test scores against the change in low-SES (left panel) or URM (right panel) enrollment shares caused by the same policy. The filled dots represent race-neutral simulations, while the hollow points are race-based policy simulations. For reference, we plot our overall 2SLS estimates with a red diamond. We overlay a linear regression on these data and cluster standard errors by study (we exclude our overall estimated impact from this calculation). This approach estimates that a 10 percentage point increase in low-SES or in URM shares leads to declines of 0.22σ or 0.17σ in average test scores, respectively. These estimates are in line with our estimated impacts of 0.18σ from Table 3.

6.1.1 How would matching district demographics impact test scores?

As noted in Section 2.5, ASHSs often underrepresent free lunch eligible students, Black students, and Latino students while overrepresenting White and Asian students relative to the racial and socioeconomic composition of their districts. Our 2SLS estimates of the impact from changes in FL or URM enrollment shares on test scores allow us to consider how ASHSs' average test scores would change if they implemented admissions policies that led to their entering cohorts perfectly matching their districts' SES or racial composition.

First, we calculate the district-level share of underrepresented minority students and FL-eligible students for the 2015-2016 school year for every district with an untreated ASHS in our sample and for the year prior to treatment for every district with a treated ASHS in our sample. From these district-level shares, we calculate the representation gap for each ASHS by subtracting its enrollment shares from its districts' enrollment shares in the same year. Finally, we multiply the gap by the coefficients estimated in Panels B (for FL-eligible share) and C (for URM share) of Table 3 to generate the test score impact of matching district composition.

Figure A16 and Figure A17 plot the simulated effect of matching district FL shares on school-wide average ACT scores and matching district URM shares on school-wide average ACT scores, respectively. Panel (a) of each figure shows the predicted impact on raw scores

[43] We show that the median effect on test scores from race-neutral simulations is more negative than the median effect from race-based simulations.

and Panel (b) shows the predicted impacts on school average ACT percentile.[44] While each raw score figure shows significant declines in school average ACT scores under this policy simulation, the percentiles figure shows that the test score declines lead to much more modest changes in school average ACT percentile ranking. That is, even if they matched district demographics, ASHSs could maintain their status among the highest-scoring schools in the country.

6.2 Heterogeneity by policy type

In panels B and C of Table A11, we examine heterogeneity in the trade-offs between impacts on test scores and impacts on FL or URM enrollment shares, by policy type. SCA has the smallest impact on academic achievement, with estimated declines of 0.05σ. This is followed by the Class-Ranking and Geography-Ranking admissions policies. Low-income quotas have the largest impacts on academic achievement, with estimated declines of 0.29σ to 0.31σ. One potential reason for SCA and Class-Ranking admissions policies having the least negative impacts on academic achievement is that both policies are more strictly beneficial to applicants based on quantitative academic measures.

The forest plots in Figures A18 and A19 show that our estimates by policy-type are generally consistent with simulation-based predictions of the trade-off between diversity and academic achievement, capturing all but one to two predictions. However, our 2SLS approach generates large standard errors.

6.3 Discussion

While our data allow us to examine changes in average academic performance during high school, we are unable to directly examine how these policies impact incoming students' achievement. Studies using regression discontinuity (RD) in Boston, Chicago, and New York suggest that ASHSs have negligible impacts on marginally admitted students' standardized test scores (Abdulkadiroglu et al., 2014; Angrist et al., 2023; Barrow et al., 2020; Dobbie & Fryer, 2014). However, emergent research suggests that income- or race-based AA that admits lower-achieving students may positively impact marginal students' high school test scores by providing them access to exam schools (Idoux, 2022; Rokkanen, 2015).[45]

[44] School average ACT percentiles are calculated based on each ASHS's percentile ranking relative to school-wide average ACT scores for high schools in AL, CA, DC, DE, FL, GA, IL, KS, KY, LA, MA, MO, NC, NV, NJ, NY, PA, SC, TN, TX, and WI.

[45] These studies focus on marginal students farther from the admissions cutoff or marginal students who face weaker high school alternatives in Boston.

Because our estimated test score impacts align with predictions based on incoming average achievement, our results suggest that ASHSs do not improve academic growth enough to offset declines in incoming student performance from race-neutral AA policies. This conclusion is suggestive, as our data do not directly measure changes in incoming achievement. Further research should examine how SES-based AA admissions policies influence the academic outcomes of students affected by these reforms.

7 Conclusion

Selective educational institutions provide students with access to competitive learning environments, substantial resources, and high-achieving peers (Chetty et al., 2023; Dobbie & Fryer, 2014; Finn & Hockett, 2012). For decades, these institutions have struggled to diversify their campuses and expand access to high-quality education for students from different social backgrounds. While universities have been at the center of this debate, the topic is equally controversial for admissions to highly-coveted K–12 academically selective schools.

In this paper, we first provide a descriptive analysis of the SES and racial/ethnic diversity at public, academically selective high schools across the United States. We then provide the first empirical evidence on the effects of socioeconomic class-based admissions policies on the demographic composition at these schools. To do this, we create a novel national dataset of admissions policies at academically selective high schools (ASHSs) and changes in these policies over time, which we link to national 9th-grade cohort and high school-level demographic data.

We show that selective schools are concentrated in large cities that serve a significant fraction of the US student population. We find that, on average, low-income, Black, and Latino students are underrepresented at ASHSs compared to non-ASHSs in districts with at least one ASHS. This under-representation has been shrinking for low-income students, but growing for URM students over time. We then use geographic and temporal variation to show that SES-based affirmative action admissions policies lead to increases in low-income and URM enrollment shares. These effects are almost entirely driven by declines in White enrollment shares, with very little, if any, changes in Asian enrollment shares.

Importantly, we show that SES-based admissions policies that directly target low-SES students through Low-Income Quotas, ELC, and Priority Groups have the largest and most statistically significant impacts on free-lunch enrollment shares, while SCA has smaller effects. In terms of race/ethnicity, all four admissions policies that we investigate have positive and statistically significant effects on the URM enrollment shares at ASHSs.

Lastly, we find suggestive evidence that, on average, SES-based admissions policies lead to

null, or slightly positive, impacts on ACT/SAT scores, even though they admit students with lower past performance levels. Our findings support predictions that have been proposed by relevant theoretical and simulation-based research.

Data limitations and the recency of these policy changes, impede us from studying the impacts on advanced course enrollment, bullying, and behavioral issues. We leave the exploration of these important and policy-relevant impacts to future work.

This paper fills a gap in the economics and education literature and provides insights for policymakers and administrators at selective educational institutions into the potential effects of SES-based affirmative action admissions policies, which do not explicitly rely on race—making them permissible under the latest Supreme Court rulings against race-based affirmative action.

References

Abdulkadiroglu, A., Angrist, J. D., & Pathak, P. A. (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica*, *82*(1), 137–196. https://doi.org/10.3982/ECTA10266

Angrist, J. D., Pathak, P. A., & Zarate, R. A. (2023). Choice and consequence: Assessing mismatch at Chicago exam schools. *Journal of Public Economics*, *223*, 104892. https://doi.org/10.1016/j.jpubeco.2023.104892

Arcidiacono, P., & Lovenheim, M. (2016). Affirmative Action and the Quality–Fit Trade-off. *Journal of Economic Literature*, *54*(1), 3–51. https://doi.org/10.1257/jel.54.1.3

Barrow, L. (2002). School choice through relocation: Evidence from the Washington, D.C. area. *Journal of Public Economics*, *86*(2), 155–189. https://doi.org/10.1016/S0047-2727(01)00141-4

Barrow, L., Sartain, L., & de la Torre, M. (2020). Increasing Access to Selective High Schools through Place-Based Affirmative Action: Unintended Consequences. *American Economic Journal: Applied Economics*, *12*(4), 29.

Black, S. E., Cortes, K. E., & Lincove, J. A. (2015). Academic Undermatching of High-Achieving Minority Students: Evidence from Race-Neutral and Holistic Admissions Policies. *American Economic Review*, *105*(5), 604–610. https://doi.org/10.1257/aer.p20151114

Bleemer, Z. (2023). Affirmative Action and its Race-Neutral Alternatives. *Journal of Public Economics*. https://doi.org/https://doi.org/10.1016/j.jpubeco.2023.104839

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Cestau, D., Epple, D., & Sieg, H. (2017). Admitting students to selective education programs: Merit, profiling, and affirmative action. *Journal of Political Economy*, *125*(3), 761–797. Retrieved October 15, 2017, from http://www.journals.uchicago.edu/doi/abs/10.1086/691702

Chetty, R., Deming, D. J., & Friedman, J. N. (2023). Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges. *NBER Working Paper*. https://doi.org/10.3386/w31492

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2020). Race and Economic Opportunity in the United States: An Intergenerational Perspective. *The Quarterly Journal of Economics*, *135*(2), 711–783. https://doi.org/10.1093/qje/qjz042

Corcoran, S. P., & Baker-Smith, E. C. (2018). Pathways to an Elite Education: Application, Admission, and Matriculation to New York City's Specialized High Schools. *Education Finance and Policy*, *13*(2), 256–279. https://doi.org/10.1162/edfp_a_00220

Cortes, K. E. (2010). Do bans on affirmative action hurt minority students? Evidence from the Texas Top 10% Plan. *Economics of Education Review*, *29*(6), 1110–1124. https://doi.org/10.1016/j.econedurev.2010.06.004

DCPS. (2023). *Selective High School Application Update for School Year 2022-2023* (tech. rep.). District of Columbia Public Schools. https://enrolldcps.dc.gov/sites/dcpsenrollment/files/page_content/attachments/English_SY22-23%20DCPS%20Selective%20HS%20Admissions%20Information%20Overview.pdf

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Dobbie, W., & Fryer, R. G. (2014). The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics*, *6*(3), 58–75. https://doi.org/10.1257/app.6.3.58

East, C. N., Miller, S., Page, M., & Wherry, L. R. (2023). Multigenerational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health. *American Economic Review*, *113*(1), 98–135. https://doi.org/10.1257/aer.20210937

Ellison, G., & Pathak, P. A. (2021). The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools. *American Economic Review*, *111*(3), 943–975. https://doi.org/10.1257/aer.20161290

Feerick Center for Social Justice. (2019). *Screened Out: The Lack of Access to NYC Screened Program Admissions Criteria* (tech. rep.). Fordham Law School. Retrieved August 27, 2024, from https://www.fordham.edu/media/home/schools/school-of-law/pdfs/screened-out-accessible.pdf

Finn, C. E., & Hockett, J. A. (2012, September). *Exam Schools*. Princeton University Press. Retrieved July 8, 2023, from https://press.princeton.edu/books/hardcover/9780691156675/exam-schools

Goodman, J., Gurantz, O., & Smith, J. (2020). Take Two! SAT Retaking and College Enrollment Gaps. *American Economic Journal: Economic Policy*, *12*(2), 115–158. https://doi.org/10.1257/pol.20170503

Goodman-Bacon, A. (2021). The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes. *American Economic Review*, *111*(8), 2550–2593. https://doi.org/10.1257/aer.20171671

Hoxby, C. M. (2003). School choice and school competition: Evidence from the United States. *Swedish Economic Policy Review*.

Hoxby, C. M., & Avery, C. (2013). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students. *Brookings Papers on Economic Activity*, *2013*, 1–65. Retrieved February 12, 2025, from https://www.brookings.edu/wp-content/uploads/2016/07/2013a_hoxby.pdf

Idoux, C. (2022). Can Selective School Admission Reforms Increase Student Achievement? *Working Paper*.

Kapor, A. (2020). Distributional Effects of Race-Blind Affirmative Action [Number: 2020-51 Publisher: Princeton University. Economics Department.]. *Working Papers*. Retrieved January 16, 2024, from https://ideas.repec.org//p/pri/econom/2020-51.html

Klasik, D., & Cortes, K. E. (2022). Uniform admissions, unequal access: Did the top 10% plan increase access to selective flagship institutions? *Economics of Education Review*, *87*, 102199. https://doi.org/10.1016/j.econedurev.2021.102199

Lawrence, E., & Mollborn, S. (2017). Racial/Ethnic Patterns of Kindergarten School Enrollment in the United States [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/socf.12352]. *Sociological Forum*, *32*(3), 635–658. https://doi.org/10.1111/socf.12352

Mader, N., Kramer, A., & Butel, A. (2018, November). *Promising Outcomes, Limited Potential: Diversity in Admissions in New York City Public Schools* (tech. rep.). Center for New York City Affairs. New York, NY.

Perez, J. (2018). Competition remains fierce for top programs as CPS sends out high school offers. *Chicago Tribune*. Retrieved August 27, 2024, from https://www.chicagotribune.com/2018/03/30/competition-remains-fierce-for-top-programs-as-cps-sends-out-high-school-offers/

Posey-Maddox, L., Powell, S. N., Roda, A., Lenhoff, S. W., & Miller, E. O. (2025). Advantaged Families' Opportunity Hoarding in U.S. K–12 Education: A Systematic Review of the Literature [Publisher: American Educational Research Association]. *Review of Educational Research*, 00346543241304766. https://doi.org/10.3102/00346543241304766

Rokkanen, M. (2015). *Extrapolation and bandwidth choice in the regression discontinuity design* [Doctoral dissertation, Massachusetts Institute of Technology]. Retrieved October 15, 2017, from https://dspace.mit.edu/handle/1721.1/90131

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2023.03.008

Rucinski, M., & Goodman, J. (2022). Racial Diversity and Measuring Merit: Evidence from Boston's Exam School Admissions. *Education Finance and Policy*, *17*(3), 408–431. https://doi.org/10.1162/edfp_a_00343

Sinclair, B., & Chen, C.-S. (2020, August). Understanding School Lunch Eligibility in the Common Core of Data. https://nces.ed.gov/blogs/nces/post/understanding-school-lunch-eligibility-in-the-common-core-of-data

Stuyvesant High School. (2023). Admissions. Retrieved January 22, 2024, from https://stuy.enschool.org/m/pages/index.jsp?uREC_ID=126615&type=d

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

U.S. Department of Education. (2023). Common Core of Data [Education Data Portal (Version 0.20.0), Urban Insitute Urban Institute, made available under the ODC Attribution License]. https://educationdata.urban.org/documentation/

U.S. News and World Report. (2023). Best U.S. High Schools (2022-2023). https://www.usnews.com/education/best-high-schools/national-rankings

Wing, C., Freedman, S. M., & Hollingsworth, A. (2024, January). Stacked Difference-in-Differences. https://doi.org/10.3386/w32054

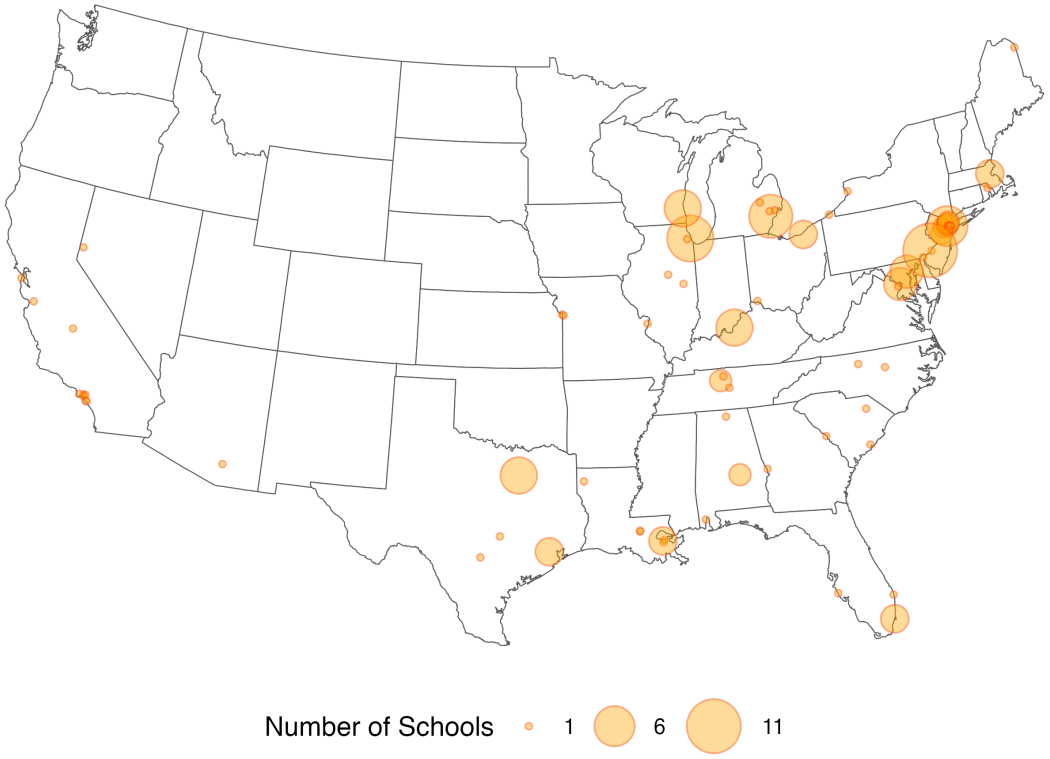
Woo, E. (1995). Caught on the Wrong Side of the Line?: Chinese Americans must outscore all other groups to enter elite Lowell High in San Francisco, sparking an ugly battle over diversity and the image of a 'model minority.' *Los Angeles Times*. Retrieved February 13, 2025, from https://www.proquest.com/docview/2021890024?sourcetype=Historical%20Newspapers

Woo, E. (1996). Racial Policy for Admission to S.F. School Altered [Num Pages: 2 Publisher: Los Angeles Times Communications LLC]. *Los Angeles Times*, NaN–NaN. Retrieved February 13, 2025, from https://www.proquest.com/docview/2047718203/abstract/357B736FE59B44B8PQ/1

Wooldridge, J. M. (2021, August). Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. https://doi.org/10.2139/ssrn.3906345

Figures

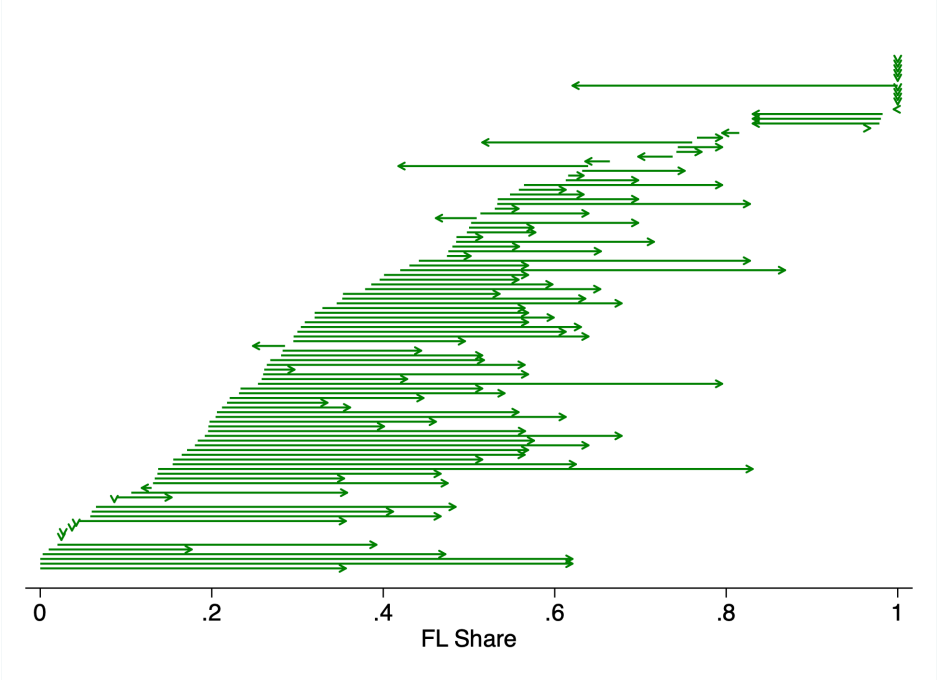
Figure 1: Map of school districts with at least one academically selective high school (ASHS)



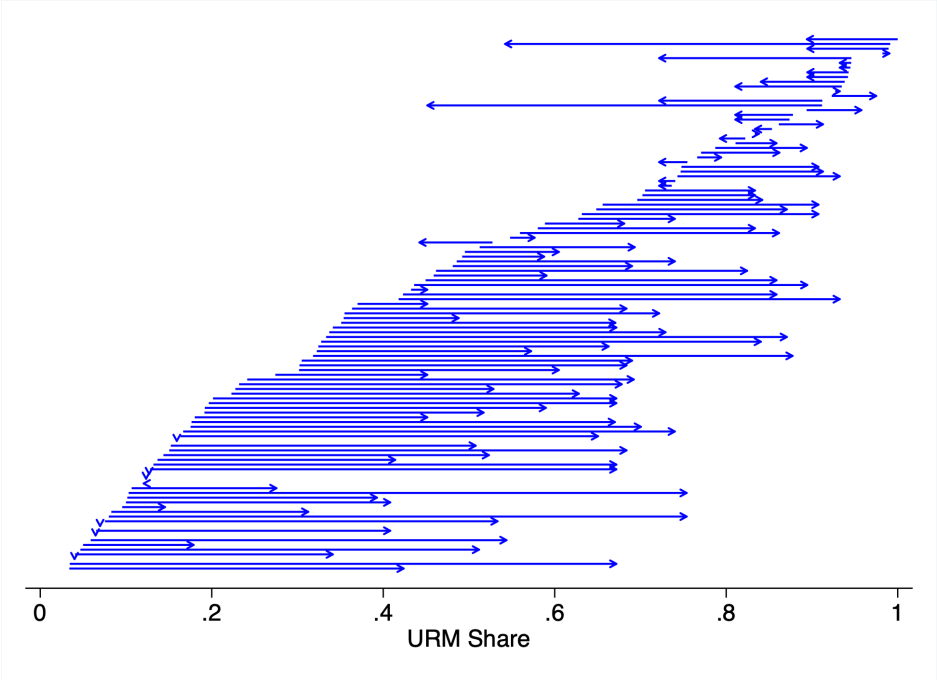
Notes: Each circle represents a separate school district. The circle size represents the number of 9th grade serving ASHSs in the corresponding school district between 2005 and 2021.

Figure 2: Gap Between ASHS and District Enrollment Share (typically underrepresented)

(a) FL Enrollment Share Gaps



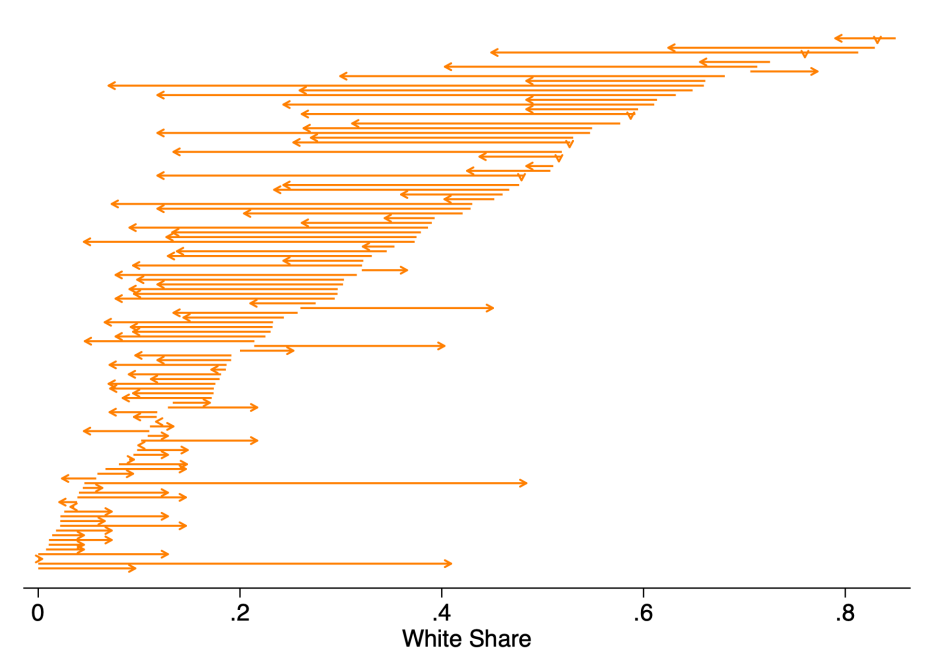
(b) URM Enrollment Share Gaps



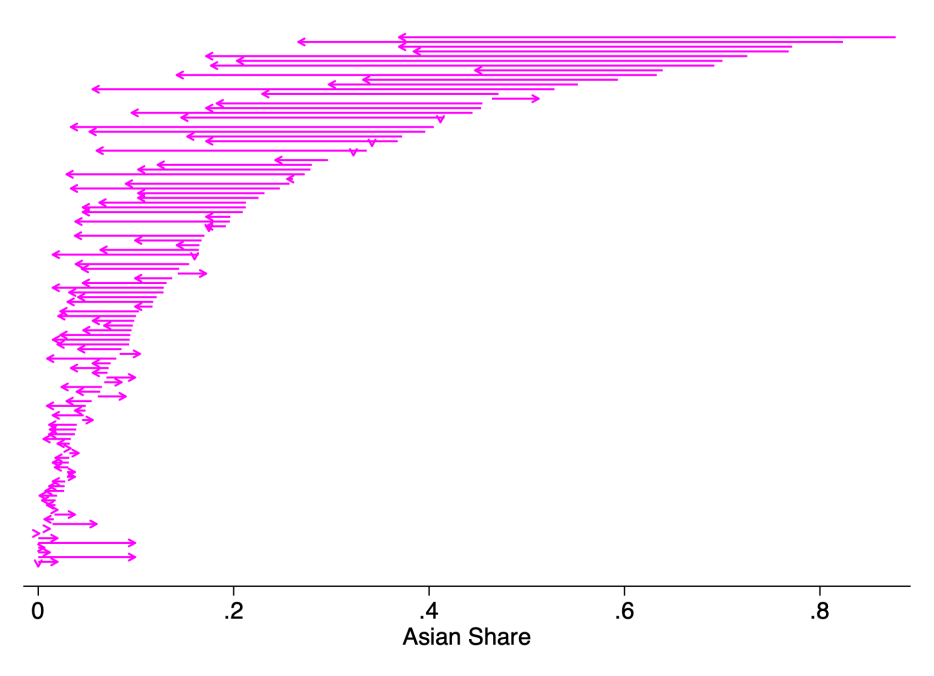
Notes: Based on sample of ASHSs in 2015-16. Each starting point represents each specific ASHS's actual enrollment shares for the corresponding group, while the arrow is the enrollment share for that same group in the corresponding school district.

Figure 3: Gap Between ASHS and District Enrollment Share (typically overrepresented)

(a) White Enrollment Share Gaps

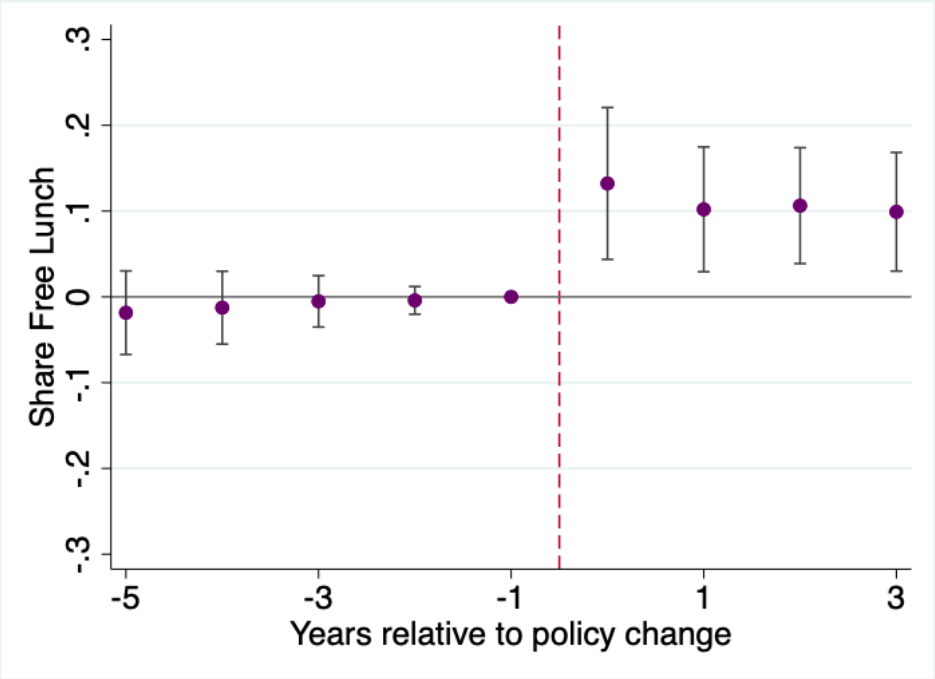


(b) Asian Enrollment Share Gaps



Notes: Based on sample of ASHSs in 2015-16. Each starting point represents each specific ASHS's actual enrollment shares for the corresponding group, while the arrow is the enrollment share for that same group in the corresponding school district.

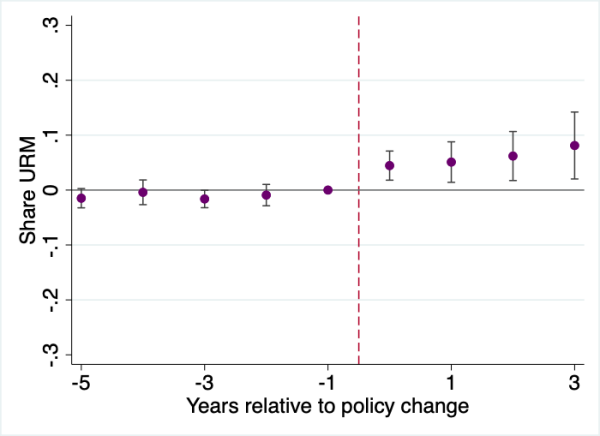
Figure 4: Effects on Free Lunch Enrollment Shares



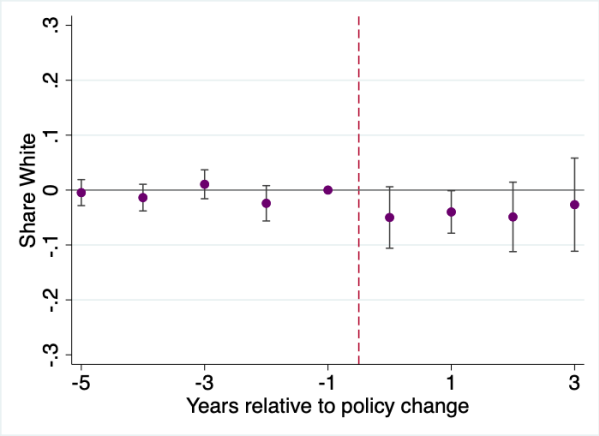
Notes: Event study is based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure 5: Effects on Racial/Ethnic Composition

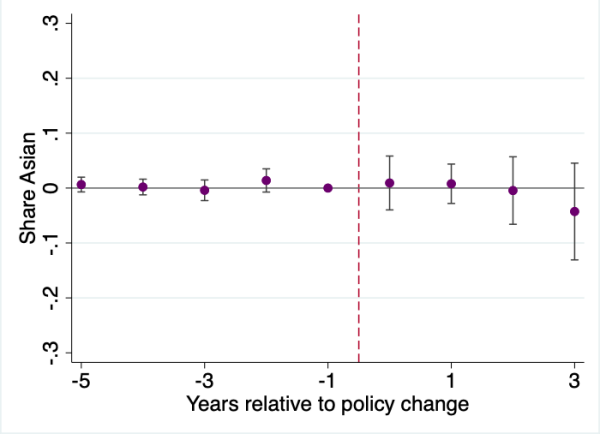
(a) URM Enrollment Shares



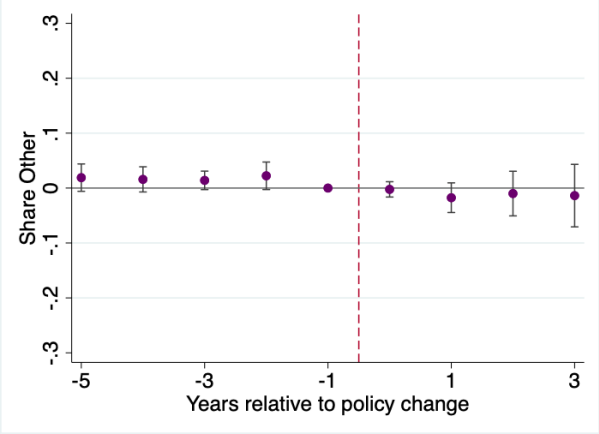
(b) White Enrollment Shares



(c) Asian Enrollment Shares

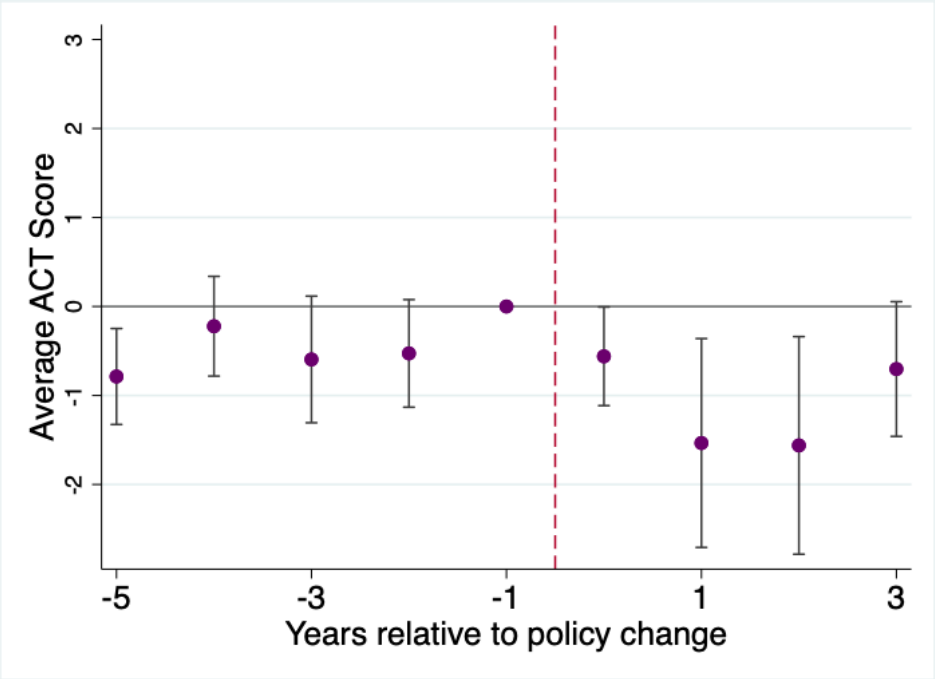


(d) Other Race Enrollment Shares



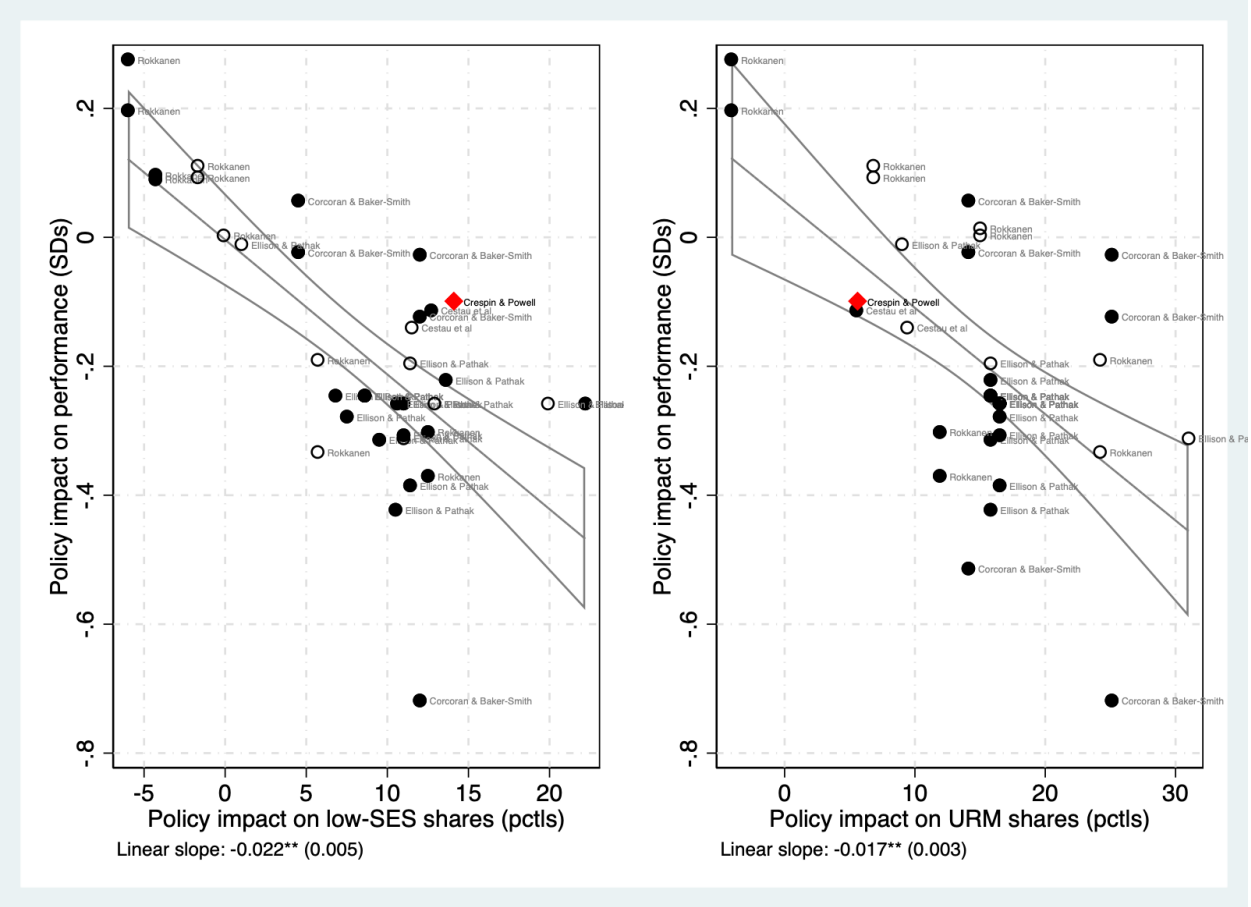
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure 6: Effects on Average Composite ACT Score



Notes: Event study is based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. We bin all observations with relative time ≥ 4 for estimation, but we do not report the estimates due to lack of balance.

Figure 7: Comparing FL/Performance tradeoffs to URM/Performance tradeoffs



Notes: Each coordinate represents the policy impact on academic achievement (y-axis) or by the policy impact on low-SES or URM enrollment shares (x-axis) for each study-policy. Point labels represent authors. The red diamonds represent our own estimated policy impacts. The linear slopes are estimated using OLS regressions with standard errors clustered by study.

Tables

Table 1: Summary statistics of ASHSs, main sample (2015–16 SY)

	Full Sample			Final Analysis Sample		
	ASHSs	Untreated	SES Treated	ASHSs	Control	SES Treated
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Main sample with available demographic data (average enrollment)						
Free Lunch Shares	0.45	0.38	0.45	0.36	0.38	0.32
	(0.32)	(0.29)	(0.28)	(0.26)	(0.29)	(0.16)
URM Shares	0.48	0.51	0.34	0.42	0.51	0.23
	(0.32)	(0.30)	(0.28)	(0.30)	(0.31)	(0.19)
White Shares	0.30	0.30	0.31	0.32	0.30	0.36
	(0.23)	(0.24)	(0.20)	(0.23)	(0.24)	(0.20)
Asian Shares	0.19	0.16	0.30	0.22	0.16	0.36
	(0.21)	(0.21)	(0.21)	(0.23)	(0.21)	(0.21)
Other Race Shares	0.04	0.03	0.05	0.04	0.03	0.05
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
9th Grade Enroll.	243.45	226.32	315.05	247.43	222.84	301.00
	(210.65)	(166.83)	(297.23)	(220.71)	(165.92)	(305.38)
School Enroll.	1025.72	949.13	1329.85	1052.99	929.36	1322.32
	(811.34)	(633.48)	(1104.45)	(847.67)	(619.15)	(1174.86)
N	130	62	40	89	61	28
Panel B: Subsample with available ACT/SAT data						
ACT Score	24.87	24.40	27.18	25.76	24.35	31.05
	(5.29)	(4.99)	(4.43)	(5.36)	(5.06)	(2.24)
N	117	53	39	63	50	13

Note: Summary statistics are based on sample described in section 2.5. In the full ASHSs demographic sample, Central Magnet School, Hume - Fogg High, Martin Luther King Jr School, Merrol Hyde Magnet School, and The Davidson Academy of Nevada are missing Free Lunch data in 2015. This means that the Share Free Lunch in column 1 is based on 125 observations and the Share Free Lunch in column 2 is based on 61 observations. Because Merrol Hyde Magnet School is one of the control units in the final sample, the Share Free Lunch in Column 4 is based on 88 observations, and the Share Free Lunch in Column 5 is based on 60 observations. Standard deviations are shown in parentheses.

Table 2: Effects of SES-based Admissions Policy on Demographic Enrollment Shares

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
Panel A: Overall Sample					
SES Policy × Post	0.135**	0.065***	-0.042	-0.001	-0.024
	(0.056)	(0.017)	(0.026)	(0.022)	(0.018)
Percent Effect	39.8%	27.9%	-12.3%	-0.3%	-36.4%
Observations	898	898	898	898	898
Panel B: Test Scores Sample					
SES Policy × Post	0.141***	0.056***	-0.044*	0.001	-0.017
	(0.052)	(0.014)	(0.023)	(0.017)	(0.015)
Percent Effect	41.7%	35.2%	-12.5%	0.2%	-27.0%
Observations	501	501	501	501	501

Notes: Estimation is based on equation 1. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

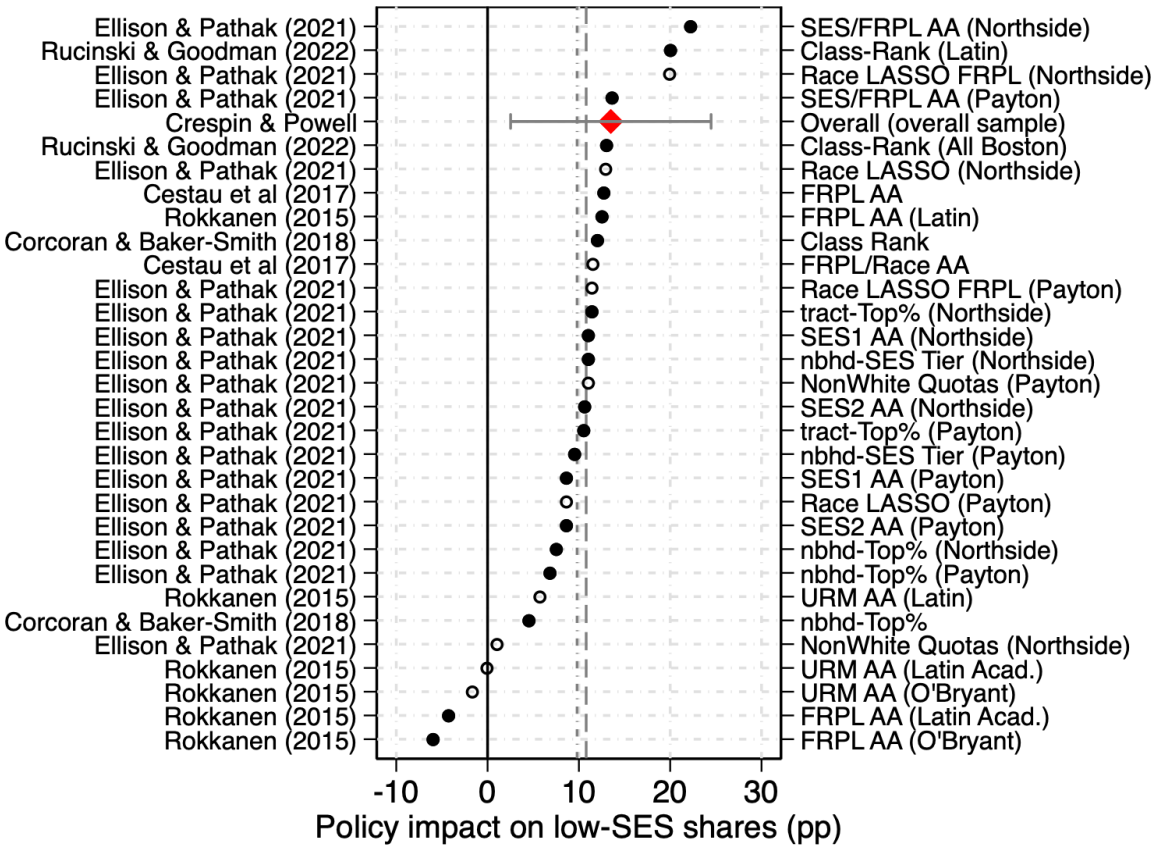
Table 3: Effects on Academic Performance

	Avg. ACT Score
	(1)
Panel A: Reduced form estimates	
SES Policy × Post	-0.529
	(0.326)
Percent Effect	-1.7%
Standardized Effect	-0.10sd
Observations	501
Panel B: IV estimates (SES composition)	
FL Enrollment Shares	-0.959
	(0.593)
Percent Effect	-3.1%
Standardized Effect	-0.18sd
Observations	501
Panel C: IV estimates (racial composition)	
URM Enrollment Shares	-0.940*
	(0.538)
Percent Effect	-3.0%
Standardized Effect	-0.18sd
Observations	501

Notes: Estimation in panel A is based on equation 1. Panels B and C report 2SLS coefficients where SES Policy implementation is the instrument and either FL Enrollment Shares or URM Enrollment Shares is the endogenous variable, respectively. IV estimates are multiplied by 10 for ease of presentation. Average ACT Scores are based on the average between ACT and the converted SAT scores, when available, as described in section 2. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standardized change shows how the average ACT score changed relative to the standard deviation of the treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

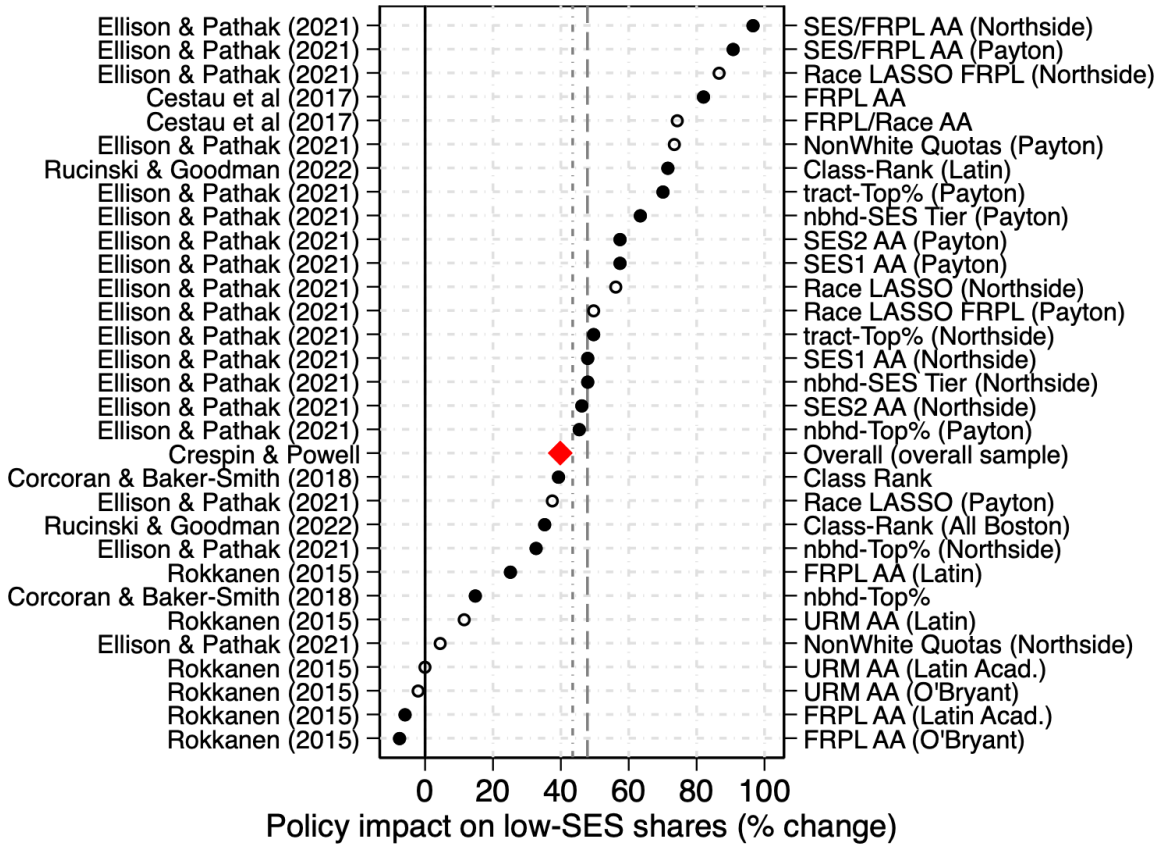
Appendix A Figures and Tables

Figure A1: Forest Plot, Policy Impacts on low-SES Enrollment Shares



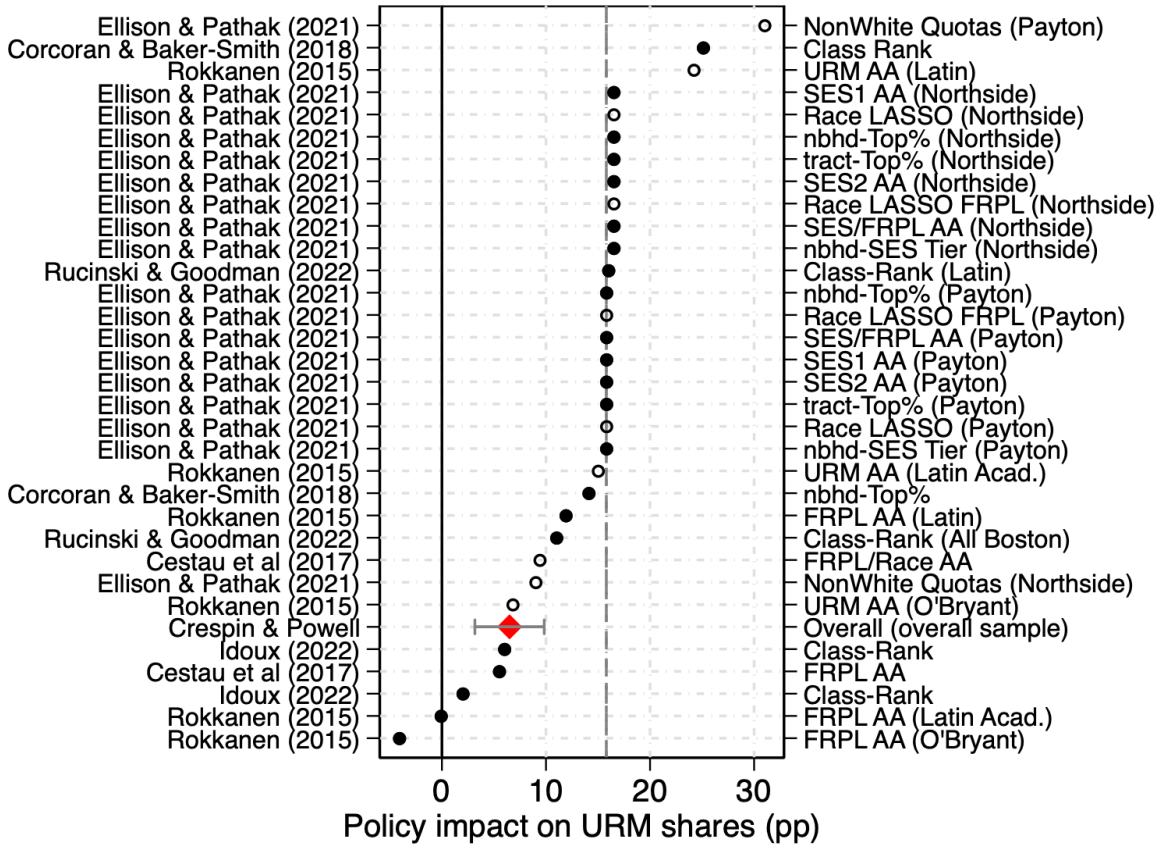
Notes: Each estimate represents the estimated policy impact on low-SES enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, filled in circles represent race-neutral simulations, and hollow circles represent race-based simulations. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A2: Forest Plot, Policy Impacts on low-SES Enrollment Shares (in percent change)



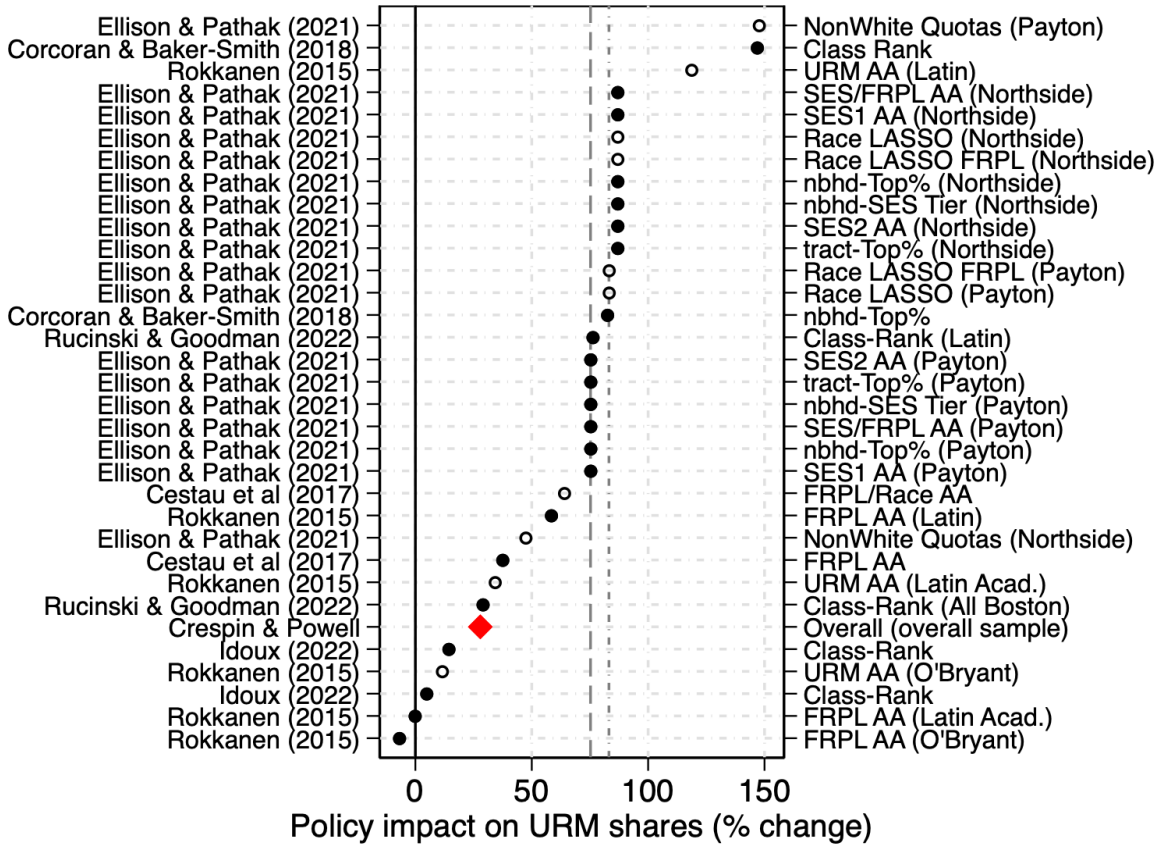
Notes: Each estimate represents the estimated policy impact on low-SES enrollment shares in terms of percent change. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, filled in circles represent race-neutral simulations, and hollow circles represent race-based simulations. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A3: Forest Plot, Policy Impacts on URM Enrollment Shares



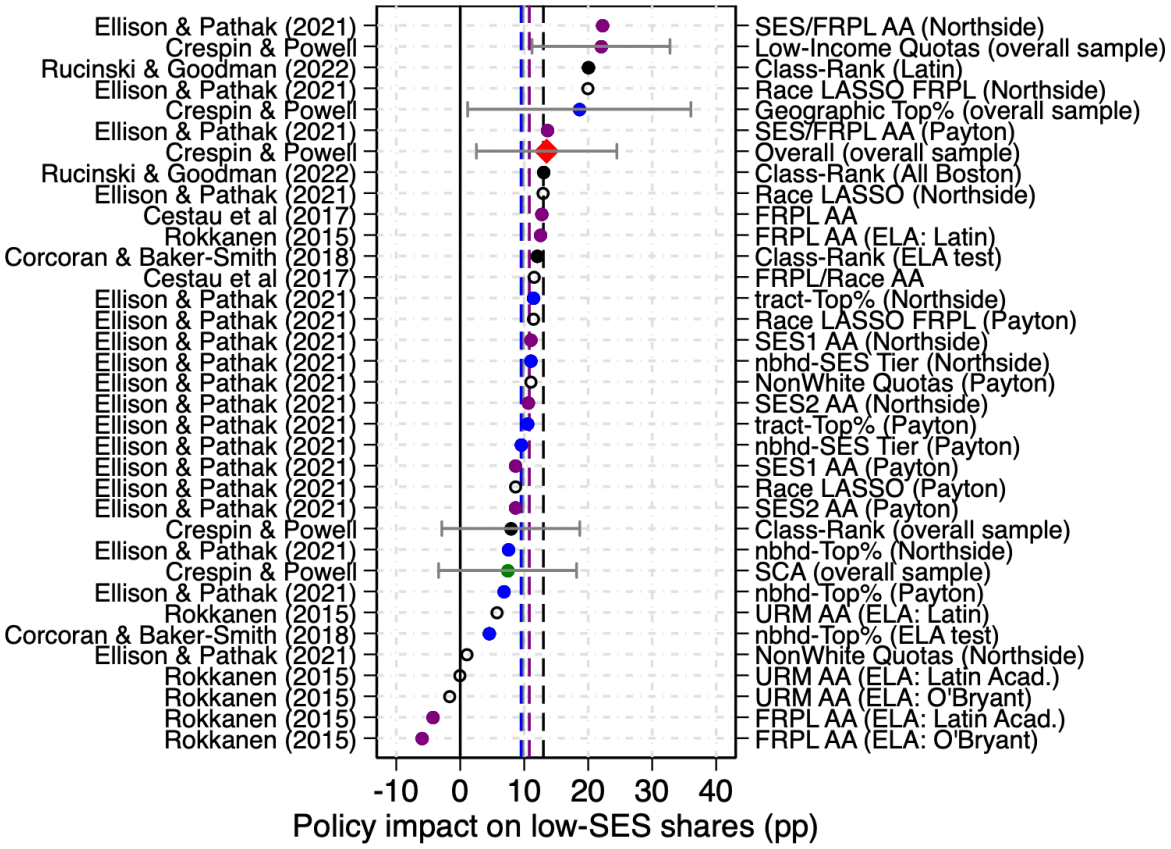
Notes: Each estimate represents the estimated policy impact on URM enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, filled in circles represent race-neutral simulations, and hollow circles represent race-based simulations. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A4: Forest Plot, Policy Impacts on URM Enrollment Shares (in percent change)



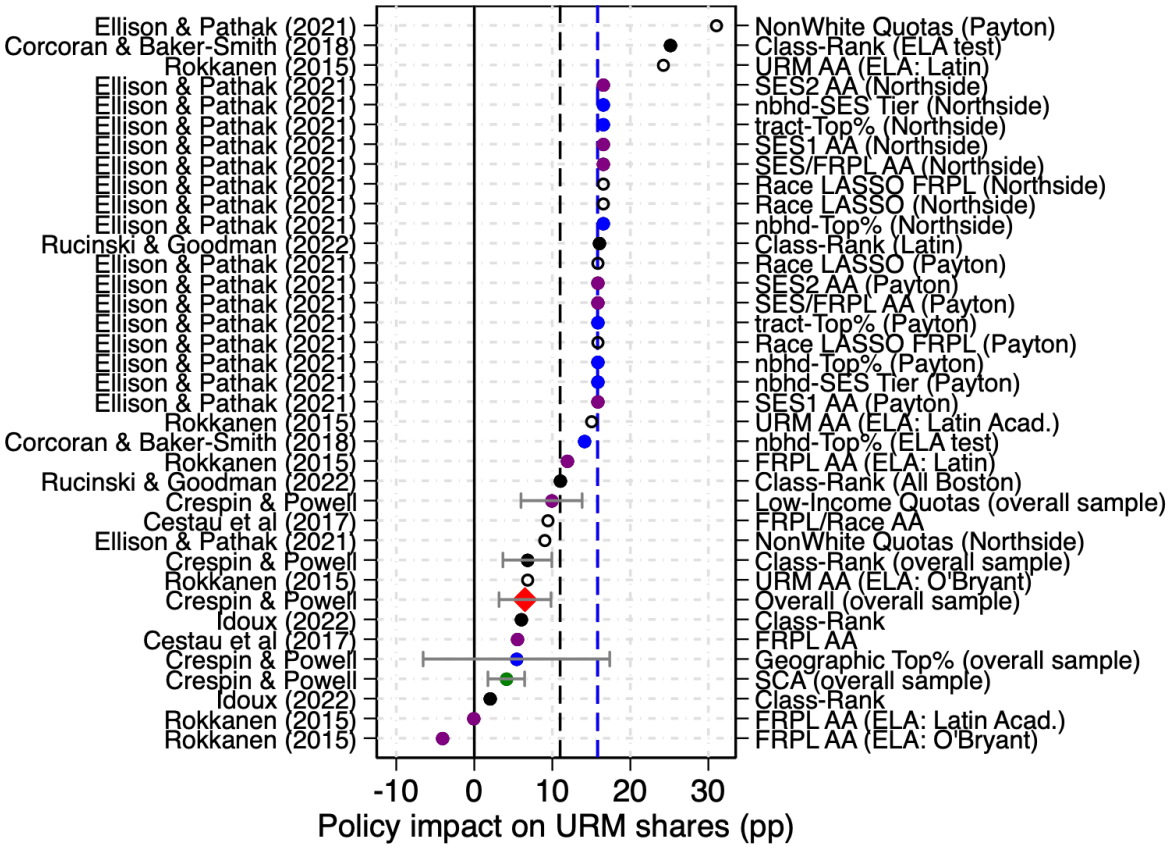
Notes: Each estimate represents the estimated policy impact on URM enrollment shares in terms of percent change. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, filled in circles represent race-neutral simulations, and hollow circles represent race-based simulations. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A5: Forest Plot, Policy Impacts on low-SES Enrollment Shares



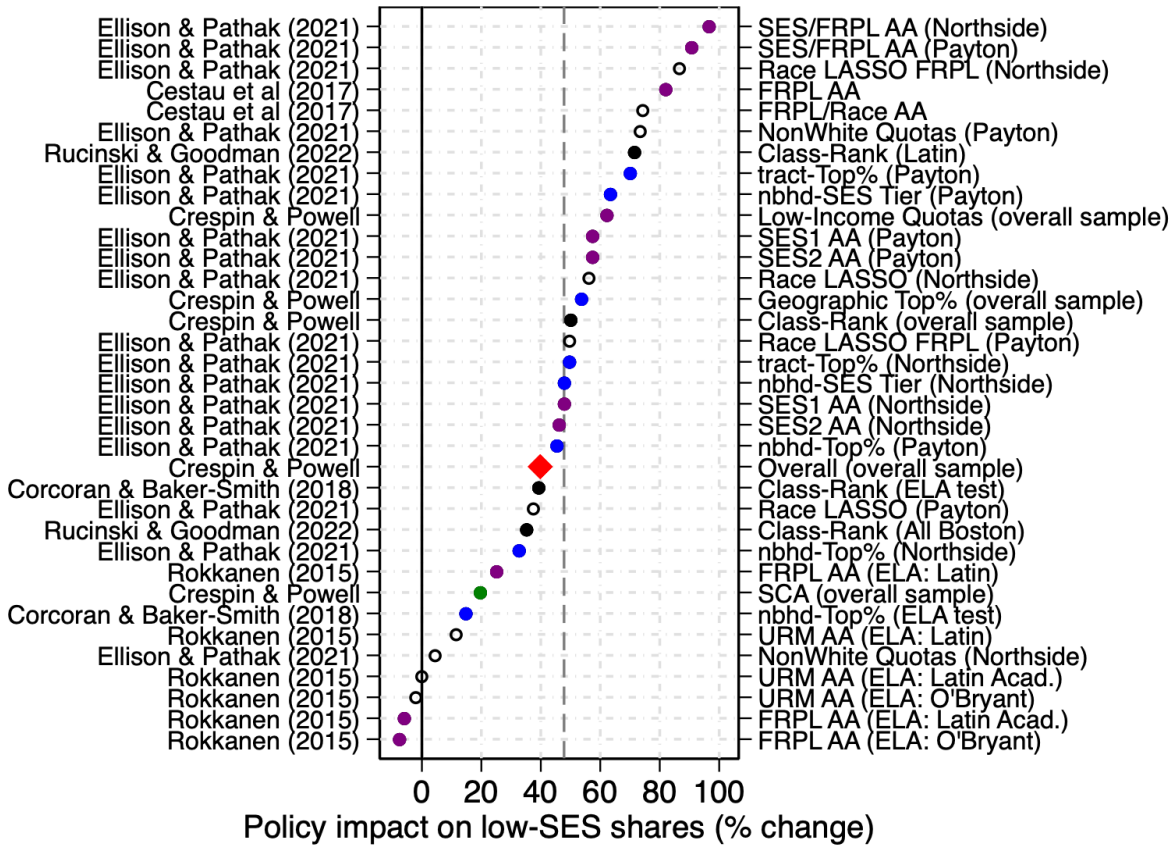
Notes: Each estimate represents the estimated policy impact on low-SES enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A6: Forest Plot, Policy Impacts on URM Enrollment Shares



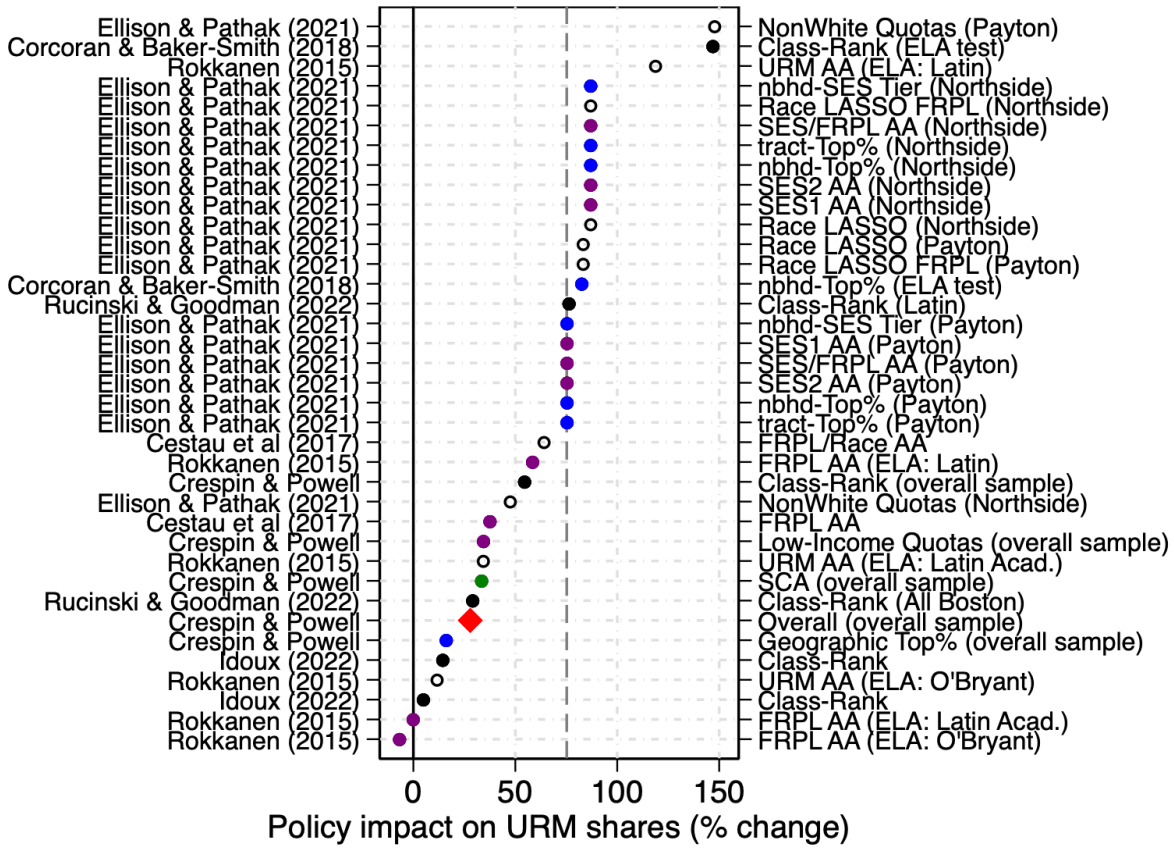
Notes: Each estimate represents the estimated policy impact on URM enrollment shares in terms of percentiles. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The error bars are based on the 95% confidence interval for each estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A7: Forest Plot, Policy Impacts on low-SES Enrollment Shares (in percentage change)



Notes: Each estimate represents the estimated policy impact on low-SES enrollment shares in terms of percentage change. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

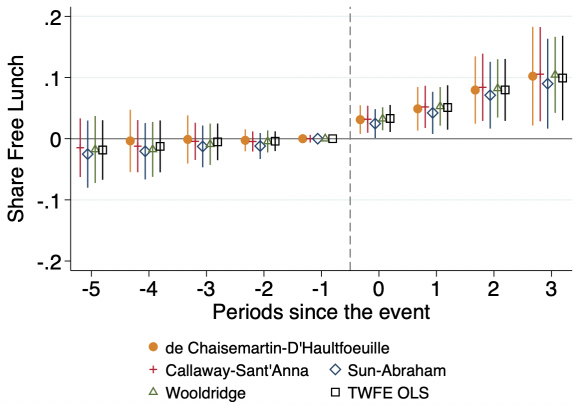
Figure A8: Forest Plot, Policy Impacts on URM Enrollment Shares (in percentage change)



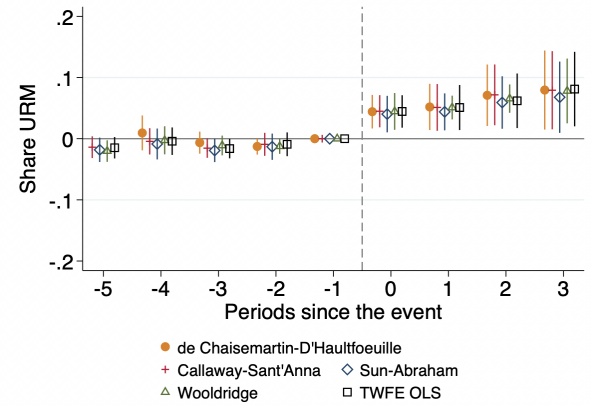
Notes: Each estimate represents the estimated policy impact on URM enrollment shares in terms of percentage change. The right-side y-axis details the policy type each estimate is based on. The red point represents our overall estimate, purple points present estimates for low-SES-based affirmative action policies, blue points present geographic-based ranking policies, black points present middle school class rank policies, while the green dot presents our SCA policy estimate. The vertical dashed gray line represents the median for the simulation-based estimates.

Figure A9: Event Study Estimates of Impacts on School Demographics – Using various estimators (without dose-response)

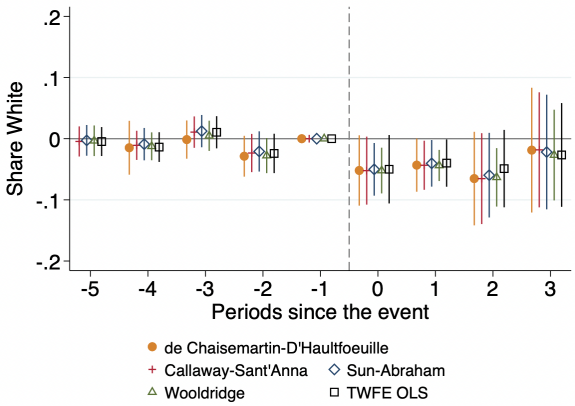
(a) Free Lunch Enrollment Shares



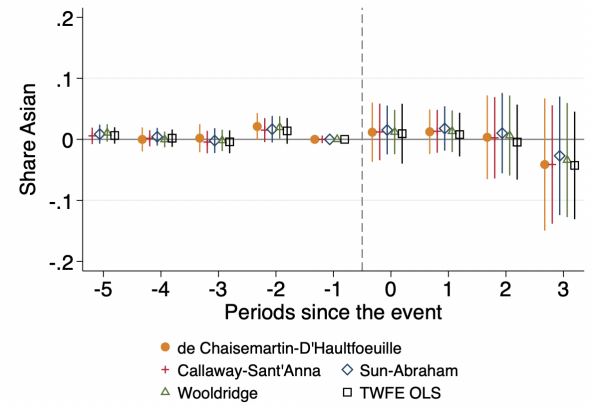
(b) URM Enrollment Shares



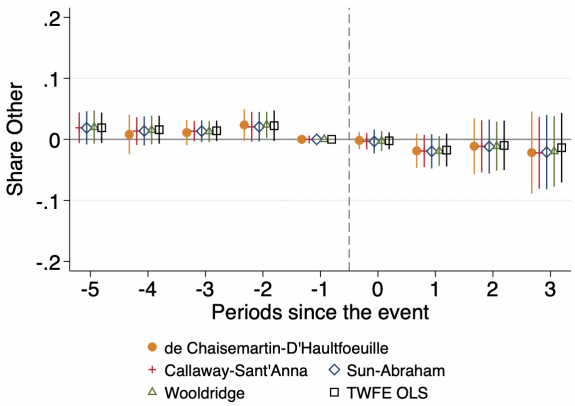
(c) White Enrollment Shares



(d) Asian Enrollment Shares

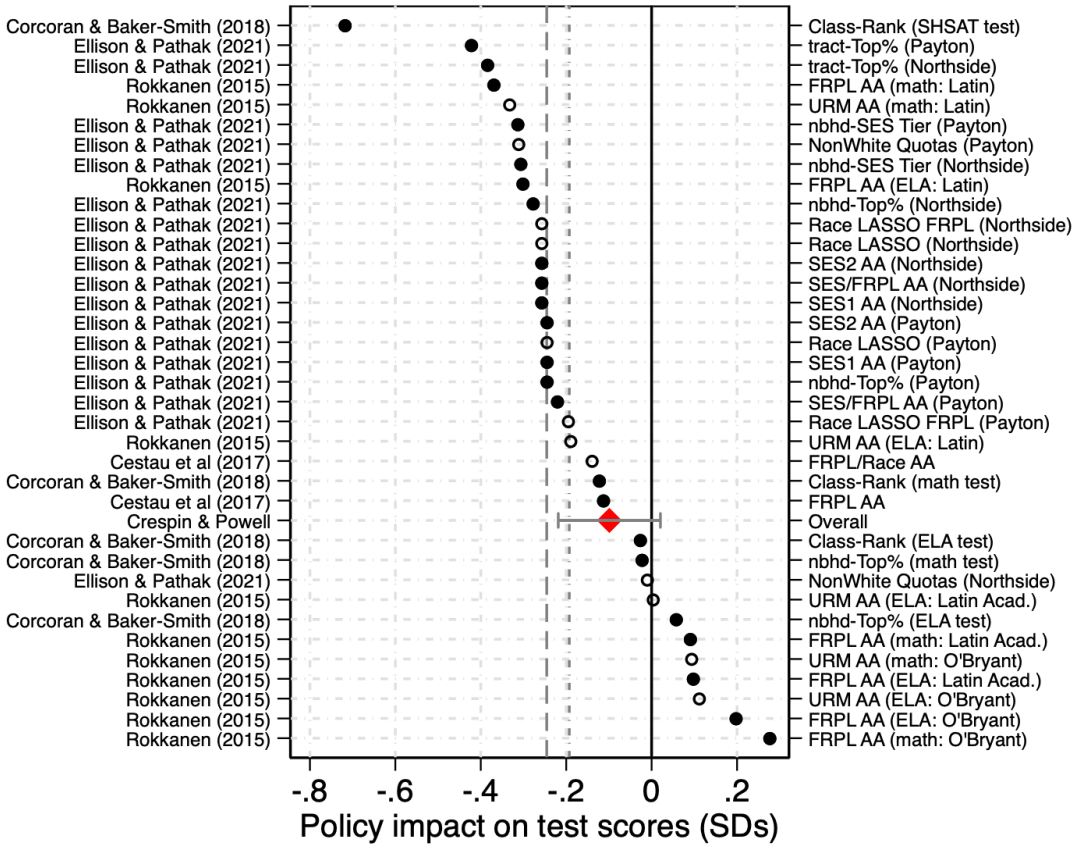


(e) Other Race Enrollment Shares



Notes: Each color/shape event study is based on a separate regression estimation (without dose response): black squares present TWFE estimates based on equation 2, orange circles present robust estimates by de Chaisemartin and D'Haultfœuille (2020), red crosses present robust estimates by Callaway and Sant'Anna (2021), blue diamonds present robust estimates by Sun and Abraham (2021), and green triangles present robust estimates by Wooldridge (2021). We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

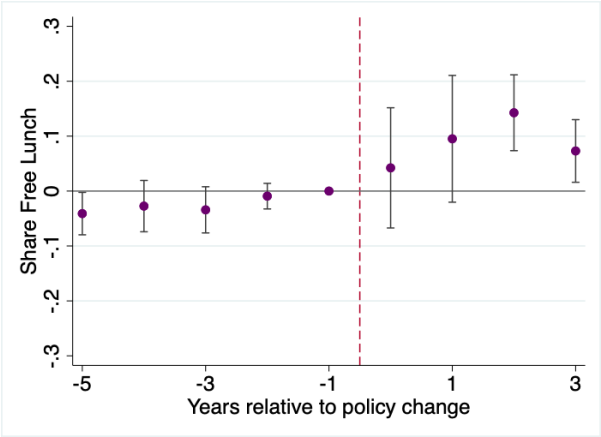
Figure A10: Forest Plot, Policy Impacts on Average Test Scores



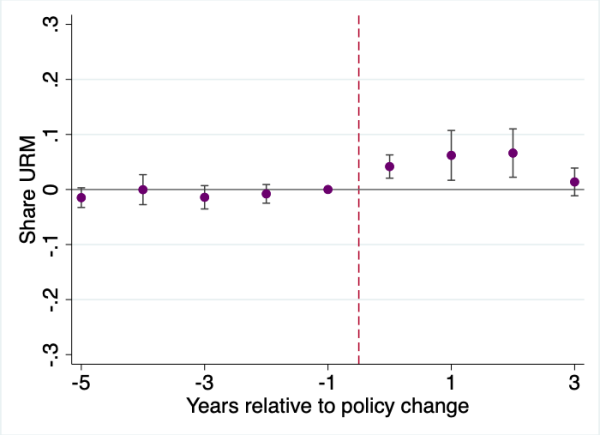
Notes: Each estimate represents the estimated policy impact on average test scores. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate. Filled points represent simulation-based race-neutral policy estimates, while hollow points are based on race-based policy simulations.

Figure A11: Effects on Racial/Ethnic Composition— Test Scores Sample

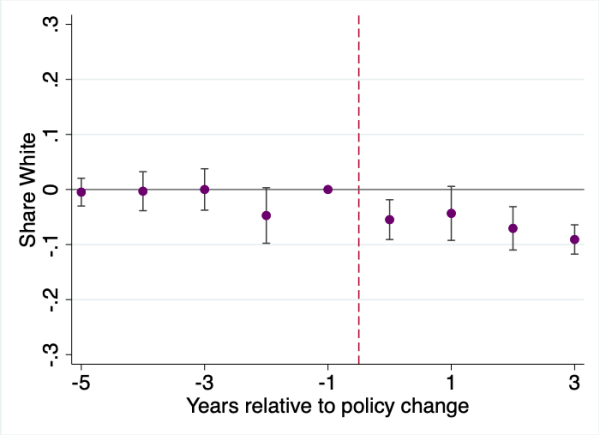
(a) Free Lunch Enrollment Shares



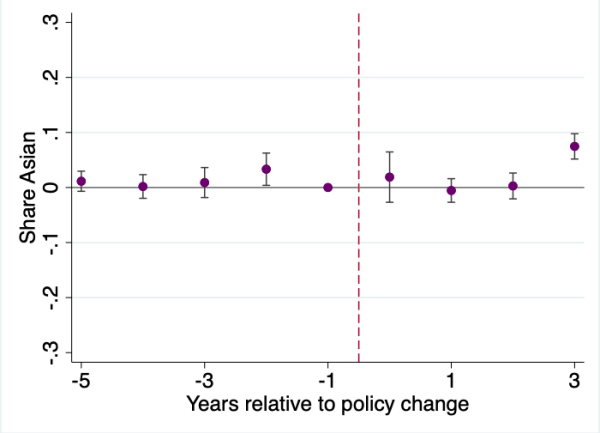
(b) URM Enrollment Shares



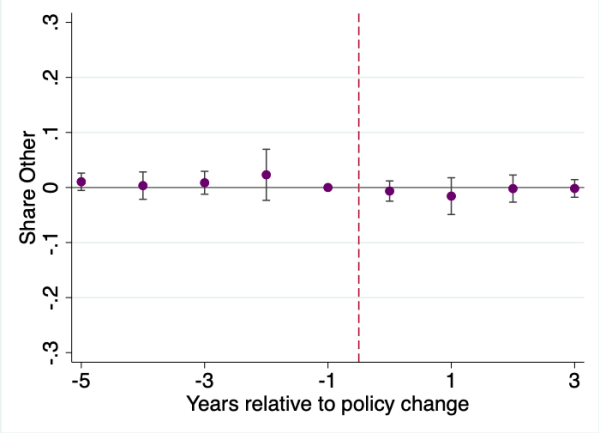
(c) White Enrollment Shares



(d) Asian Enrollment Shares

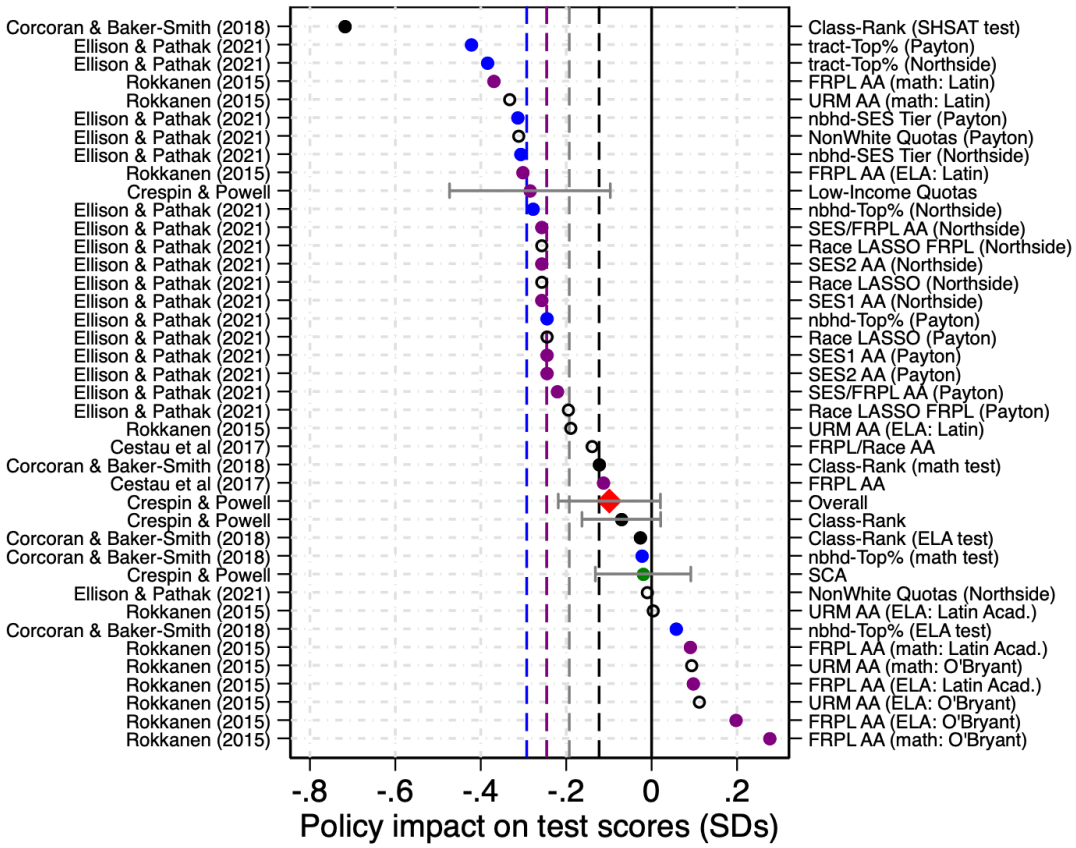


(e) Other Race Enrollment Shares



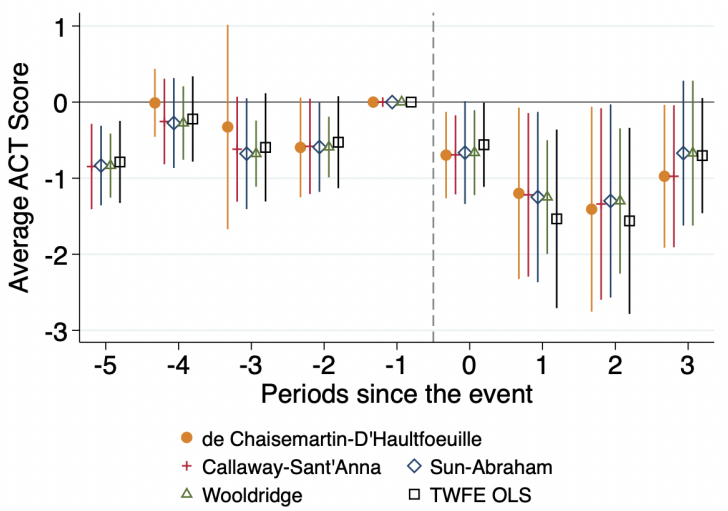
Notes: Event studies are based on equation 2. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend. Each event study is based on a separate regression estimation. We bin all observations with relative time $\geq$4 for estimation, but we do not report the estimates due to lack of balance.

Figure A12: Forest Plot, Policy Impacts on Average Test Scores (by specific policy type)



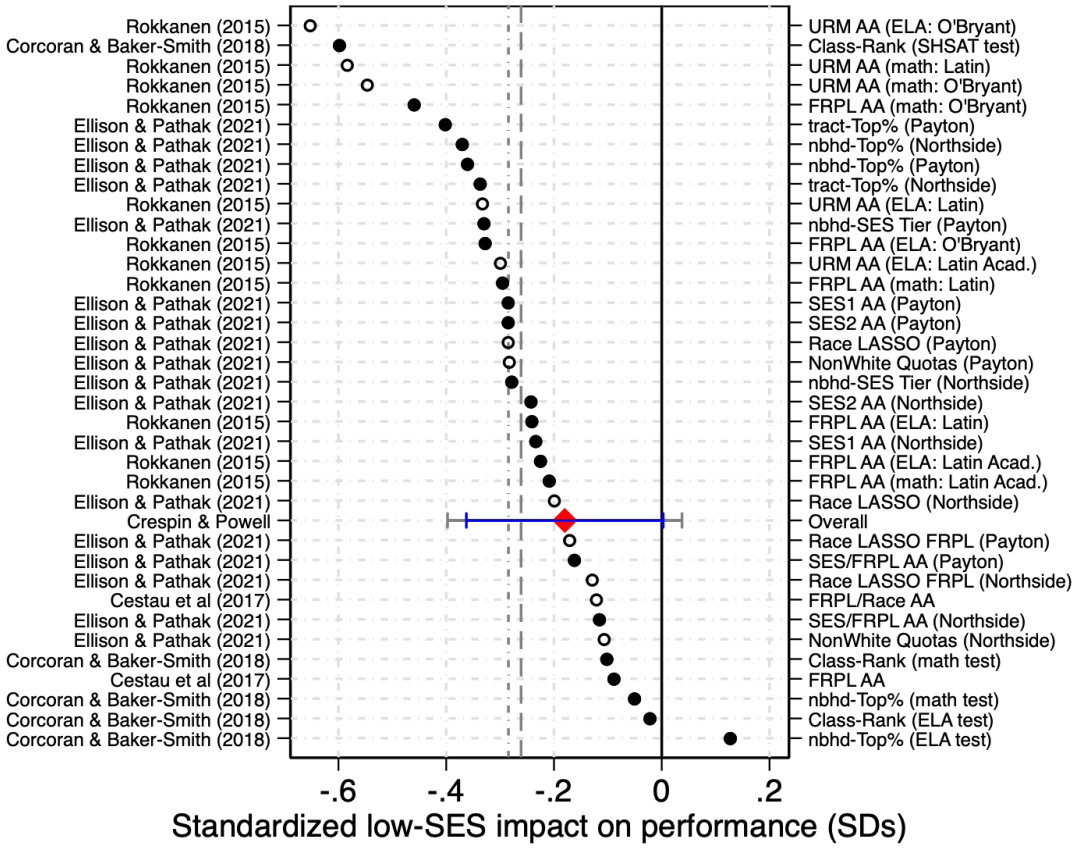
Notes: Each estimate represents the estimated policy impact on average test scores. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate.

Figure A13: Event Study Estimates of Impacts on ACT Scores – Using various estimators



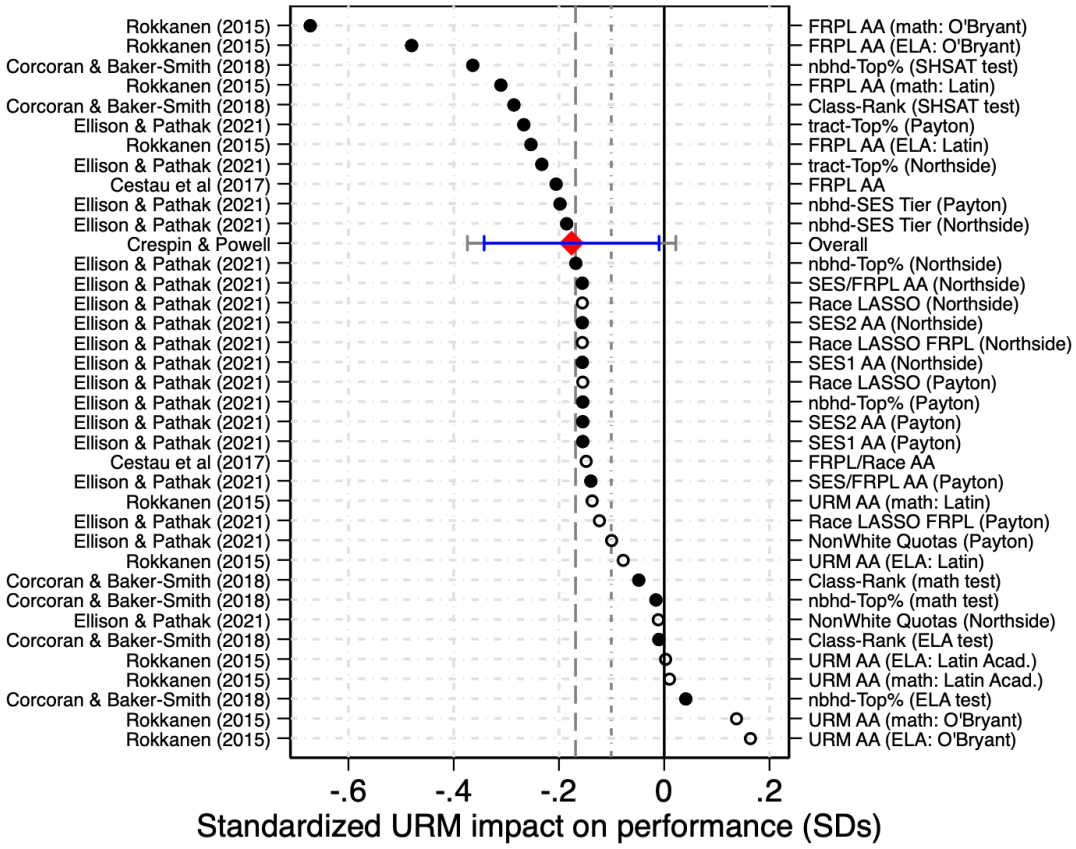
Notes: Each color/shape event study is based on a separate regression estimation (without dose response): black squares present TWFE estimates based on equation 2, orange circles present robust estimates by de Chaisemartin and D'Haultfœuille (2020), red crosses present robust estimates by Callaway and Sant'Anna (2021), blue diamonds present robust estimates by Sun and Abraham (2021), and green triangles present robust estimates by Wooldridge (2021).

Figure A14: Forest Plot, FL Enrollment Impacts on Academic Achievement



Notes: Each estimate represents the impact of a 10 percentage point increase in FL enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate.

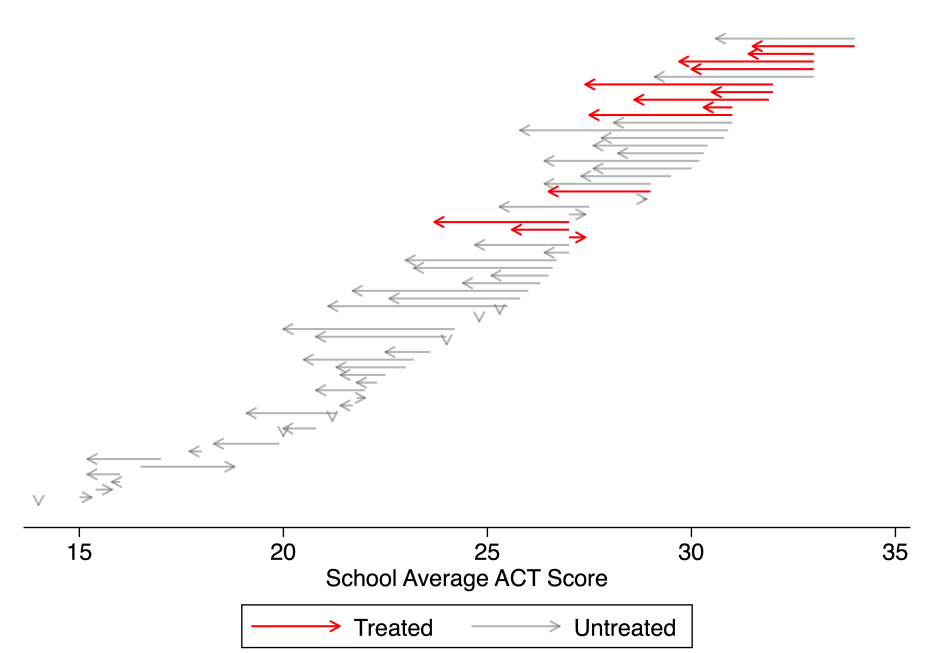
Figure A15: Forest Plot, URM Enrollment Impacts on Academic Achievement



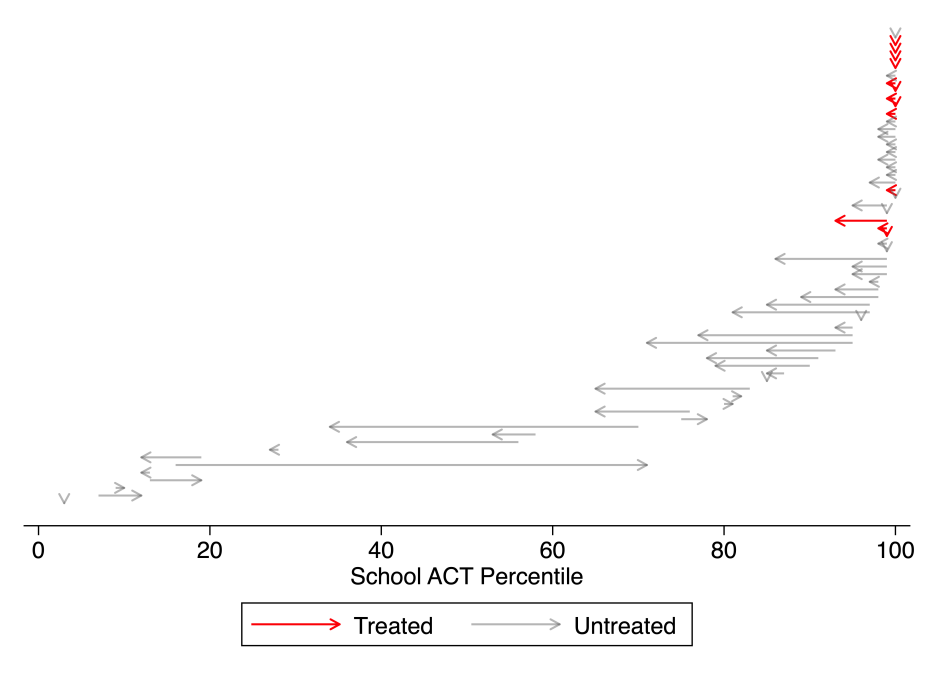
Notes: Each estimate represents the impact of a 10 percentage point increase in URM enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate.

Figure A16: Simulated Effect of Matching District FL Share on Test Scores

(a) Test Score Effect



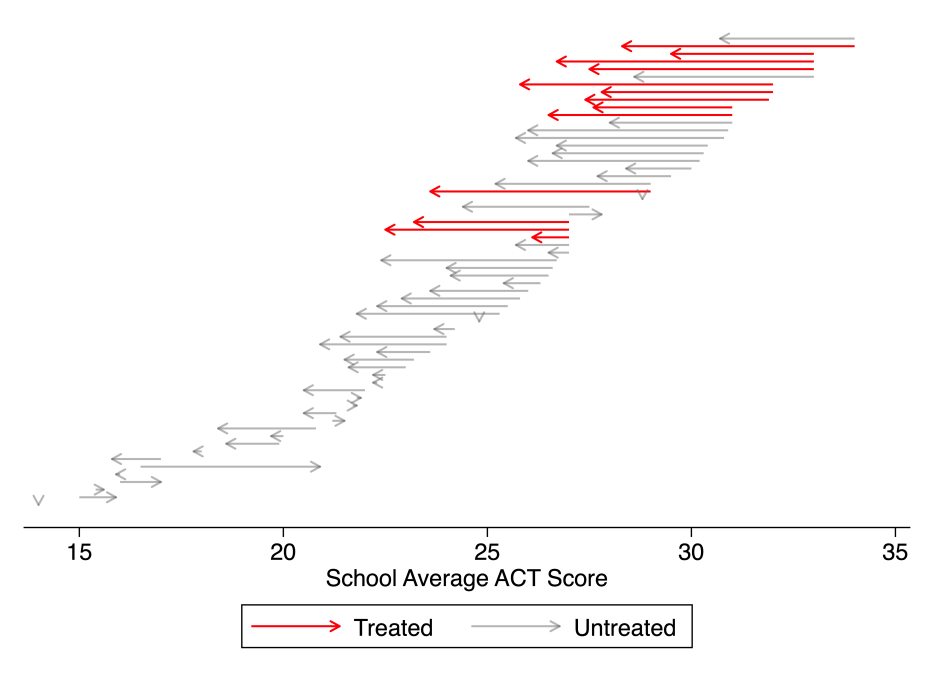
(b) School Average Test Percentile Effect



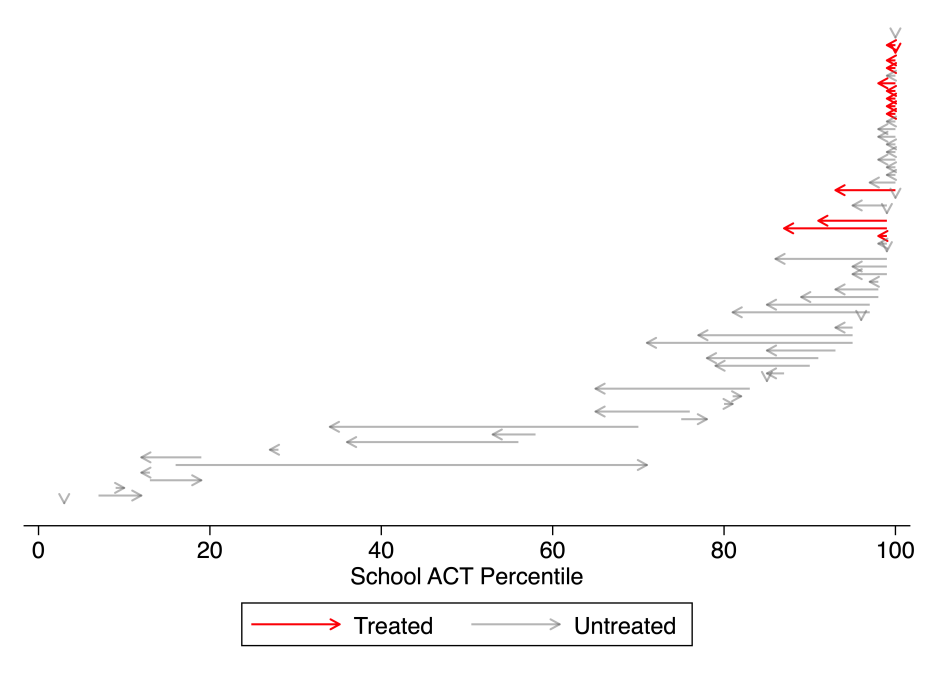
Notes: Based on sample of untreated ASHSs in 2015-16 and treated ASHSs in the year prior to treatment. Each starting point represents each specific ASHS's actual average ACT score for the corresponding group, while the arrow is the simulated average score after matching district free lunch share.

Figure A17: Simulated Effect of Matching District URM Share on Test Scores

(a) Test Score Effect

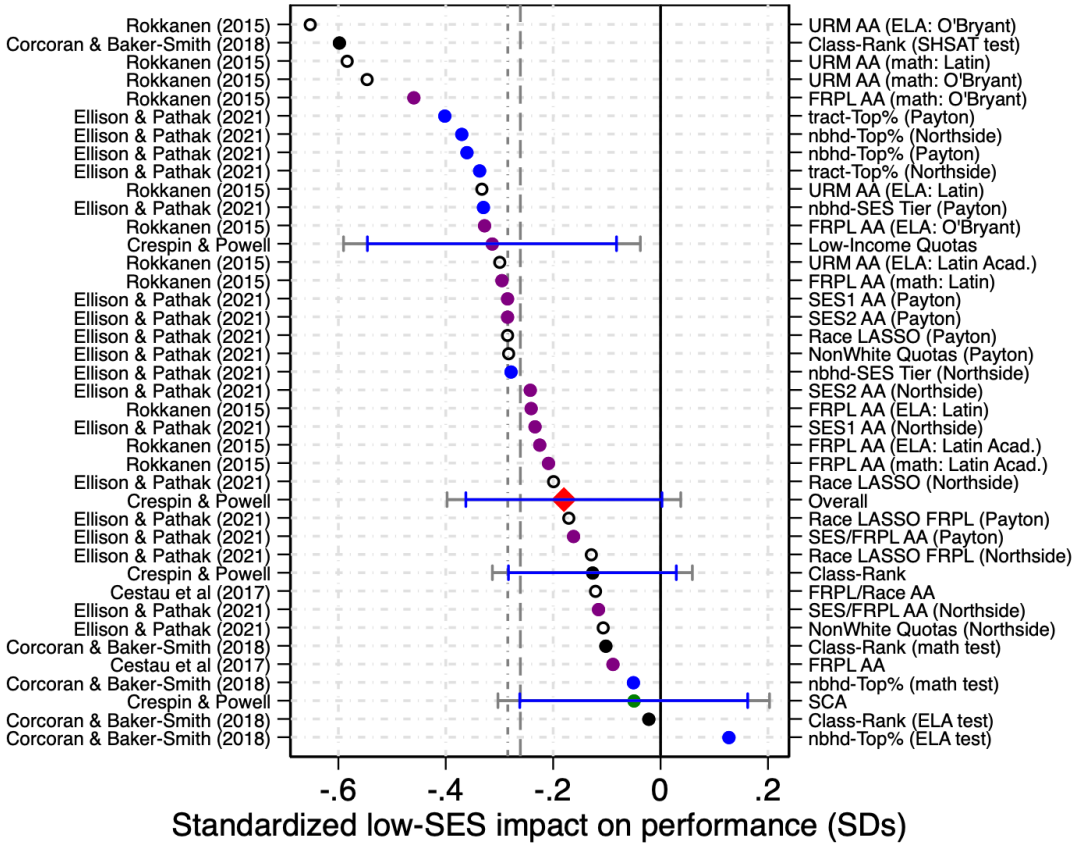


(b) School Average Test Percentile Effect



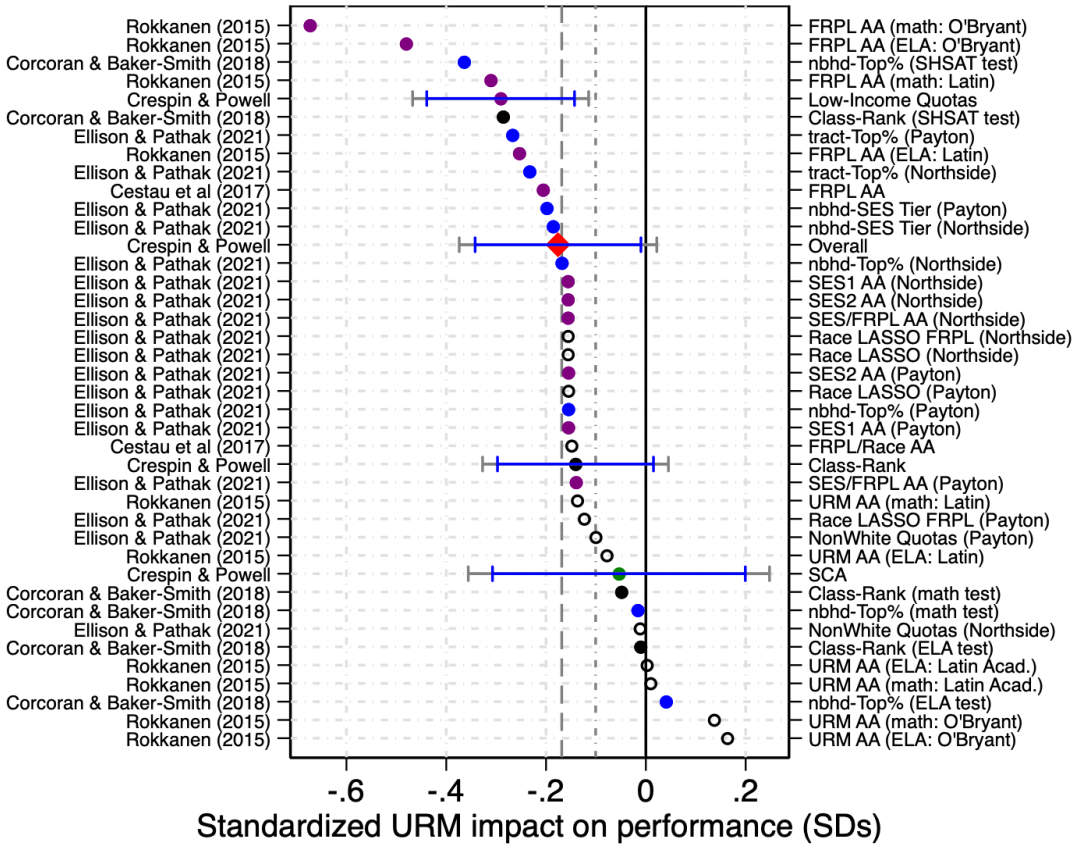
Notes: Based on sample of untreated ASHSs in 2015-16 and treated ASHSs in the year prior to treatment. Each starting point represents each specific ASHS's actual average ACT score for the corresponding group, while the arrow is the simulated average score after matching district URM share.

Figure A18: Forest Plot, FL Enrollment Impacts on Academic Achievement (by specific policy type)



Notes: Each estimate represents the impact of a 10 percentage point increase in FL enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate.

Figure A19: Forest Plot, URM Enrollment Impacts on Academic Achievement (by specific policy type)



Notes: Each estimate represents the impact of a 10 percentage point increase in URM enrollment shares on average academic achievement in terms of percentiles. The right-side y-axis categorizes the policy type each estimate is based on. The error bars are based on the 95% confidence interval for each estimate.

Table A1: Summary of K–12 Affirmative Action Simulation Studies

Study	District	Grades	Years	AA Policies	Outcome(s)
Cestau et al. (2017)	Anonymous (midsize urban)	1st	2003–2005	No AA SES-based AA SES- & Race-based AA	SES Composition URM Composition IQ Score
Corcoran and Baker-Smith (2018)	NYC	9th	2009	No AA Middle School Class Rank (10%) Within-Borough Rank (proportional % by borough)	SES Composition URM Composition 7th grade ELA test Avg. 7th grade Math test Avg. ASHS exam Avg.
Ellison and Pathak (2021)	Chicago	9th	2013	No AA Tract SES-Tier Rank Within-Tract Rank (10%) Within-Community Rank (10%) SES-based AA (weighted) SES-based AA (unweighted) SES-based AA (with FRPL) SES-based AA (predicting race) SES-based AA (with FRPL & predicting race) Race Quotas	SES Composition URM Composition ASHS application score
Idoux (2022)	Boston	7th	2004–2016	No AA Middle School Class Rank (5%) Middle School Class Rank (10%)	URM Composition
Rucinski and Goodman (2022)	Boston	7th	2007–2013	No AA Middle School Class Rank (7%) Middle School Class Rank (17%)	SES Composition URM Composition
Rokkanen (2015)	Boston	7th	2000–2004	No AA SES-based AA URM-based AA	SES Composition URM Composition 4th grade ELA test Avg. 4th grade Math test Avg.

Notes: Summarizes the available studies on the effects of race-neutral AA policies on the demographic and academic composition of ASHSs. To be included, a study had to: (a) use data rather than only theoretical models; (b) focus on academically selective K–12 schools; and (c) analyze AA policies that assign admissions priority based on student characteristics (as opposed to modifying application requirements such as GPA or test scores).

Table A2: ASHSs Treated with a Race-Neutral AA Admissions Policy Change in our Analysis Sample

Year	School Name	School District	State	Treated Periods
2019	Brooklyn Latin School	New York City Geographic District #14	NY	4
2019	High School Mathematics, Science, and Engineering at CCNY	New York City Geographic District #5	NY	4
2019	Manhattan/Hunter Science High School	New York City Geographic District #3	NY	4
2019	Queens High School for the Sciences at York College	New York City Geographic District #28	NY	4
2019	Staten Island Technical High School	New York City Geographic District #31	NY	4
2018	Bard High School Early College Queens	New York City Geographic District #24	NY	4
2017	Liberal Arts and Sciences Academy (LASA) High School	Austin Independent School District	TX	4
2016	Academic Magnet High	Charleston County School District	SC	4
2020	Leon M Goldstein High School for the Sciences	New York City Geographic District #22	NY	3
2016	Brooklyn Technical High School	New York City Geographic District #13	NY	3
2016	High School of American Studies at Lehman College	New York City Geographic District #10	NY	3
2021	Beacon High School	New York City Geographic District #3	NY	2
2021	Thomas Jefferson High	Fairfax County Public Schools	VA	2
2021	Boston Latin	Boston Public Schools	MA	2
2021	Boston Latin Academy	Boston Public Schools	MA	2
2021	O'Bryant School Math/Science	Boston Public Schools	MA	2
2019	Lowell High	San Francisco Unified	CA	2
2019	Millennium High School	New York City Geographic District #2	NY	2
2017	Bronx High School of Science	New York City Geographic District #10	NY	2
2022	Townsend Harris High School	New York City Geographic District #25	NY	1
2021	Eleanor Roosevelt High School	New York City Geographic District #2	NY	1
2021	School of the Future High School	New York City Geographic District #2	NY	1
2020	Baruch College Campus High School	New York City Geographic District #2	NY	1
2020	NYC Lab High School for Collaborative Studies	New York City Geographic District #2	NY	1
2018	Stuyvesant High School	New York City Geographic District #2	NY	1
2023	New Explorations into Science, Tech, and Math High School	New York City Geographic District #1	NY	0
2023	Young Women's Leadership School	New York City Geographic District #4	NY	0
2023	Baccalaureate School for Global Education	New York City Geographic District #30	NY	0

Table A3: Summary statistics of ASHSs, subsample with ACT/SAT data (2015–16 SY)

	Full Sample			Final Analysis Sample		
	ASHSs	Untreated	SES Treated	ASHSs	Control	SES Treated
	(1)	(2)	(3)	(4)	(5)	(6)
Average Enrollment Shares by Economic Disadvantage						
Free Lunch	0.47	0.40	0.46	0.38	0.40	0.33
	(0.31)	(0.28)	(0.27)	(0.26)	(0.29)	(0.16)
Average Enrollment Shares by Race/Ethnicity						
Asian	0.19	0.17	0.29	0.22	0.17	0.42
	(0.21)	(0.22)	(0.20)	(0.24)	(0.22)	(0.22)
Black	0.29	0.26	0.19	0.20	0.24	0.06
	(0.28)	(0.25)	(0.25)	(0.22)	(0.23)	(0.05)
Hispanic	0.19	0.23	0.16	0.22	0.24	0.12
	(0.20)	(0.24)	(0.12)	(0.23)	(0.25)	(0.10)
URM	0.48	0.50	0.35	0.42	0.49	0.17
	(0.31)	(0.30)	(0.28)	(0.30)	(0.30)	(0.14)
White	0.29	0.30	0.31	0.32	0.31	0.35
	(0.23)	(0.24)	(0.21)	(0.24)	(0.24)	(0.22)
Other Race	0.04	0.03	0.05	0.03	0.03	0.05
	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)	(0.04)
Average Cohort/School Size						
9th Graders	252.04	242.68	311.15	256.67	230.51	359.31
	(214.81)	(172.38)	(300.08)	(242.93)	(173.21)	(412.93)
Total Students	1056.82	1019.94	1317.21	1058.66	966.78	1419.08
	(811.64)	(645.77)	(1115.95)	(910.86)	(640.49)	(1571.57)
Testing						
ACT/SAT	24.87	24.40	27.18	25.83	24.47	31.05
	(5.29)	(4.99)	(4.43)	(5.35)	(5.08)	(2.24)
N	117	53	39	64	51	13

Note: Summary statistics are based on sample described in section 2.5. University Preparatory High School, University High School, and Dr. TJ Owens Gilroy Early College Academy are included as control units in the testing analysis; however, they have missing ACT/SAT data for the 2015-16 school year, so they are omitted from the ACT score average. Central Magnet School, Hume - Fogg High, and Martin Luther King Jr School, and Merrol Hyde Magnet School are missing Free Lunch share data in 2015-16. Merrol Hyde Magnet School and Dr. TJ Owens Gilroy Early College are control units for the testing analysis, so in column 4, Free Lunch and ACT/SAT averages are based on 63 observations and in column 5 they are based on 50 observations. Standard deviations are shown in parentheses.

Table A4: Effects of SES-based Admissions Policy on Demographic Enrollment Shares, for balanced sample (initial post-shock period)

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
SES Policy × Post	0.156	0.055***	-0.044	0.006	-0.019
	(0.099)	(0.016)	(0.030)	(0.024)	(0.013)
Percent Effect	46.1%	23.6%	-12.9%	1.7%	-28.6%
Observations	774	774	774	774	774

Notes: Estimation is based on equation 1. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A5: Effects on Demographic Enrollment Shares, by policy type

	Low-Income Quotas	Class Rank	Geographic Rank	SCA
	(1)	(2)	(3)	(4)
Panel A: Share Free Lunch				
SES Policy × Post	0.220***	0.079	0.186**	0.074
	(0.055)	(0.055)	(0.089)	(0.055)
Percent Effect	62.1%	50.0%	53.6%	19.6%
Observations	776	714	729	758
Panel B: Share URM				
SES Policy × Post	0.099***	0.068***	0.054	0.041***
	(0.020)	(0.016)	(0.061)	(0.012)
Percent Effect	34.3%	54.4%	16.0%	33.3%
Observations	776	714	729	758
Panel C: Share White				
SES Policy × Post	-0.094**	-0.008	-0.072	0.013
	(0.036)	(0.036)	(0.059)	(0.030)
Percent Effect	-25.8%	-2.0%	-21.4%	4.6%
Observations	776	714	729	758
Panel D: Share Asian				
SES Policy × Post	0.009	-0.040	0.021	-0.013
	(0.035)	(0.050)	(0.015)	(0.047)
Percent Effect	3.4%	-9.7%	7.7%	-2.5%
Observations	776	714	729	758
Panel E: Share Other Race				
SES Policy × Post	-0.014	-0.023*	-0.006	-0.044
	(0.017)	(0.013)	(0.023)	(0.033)
Percent Effect	-19.2%	-41.8%	-11.1%	-64.7%
Observations	776	714	729	758

Notes: Estimation is based on equation 1. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A6: Effects of SES-based Admissions Policy on Enrollment Shares, excluding or including NYC schools

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
Panel A: Treated units only include non-NYC academically selective high schools					
SES Policy Change X Post	0.158***	0.050	-0.020	-0.013	-0.020
	(0.056)	(0.036)	(0.019)	(0.026)	(0.015)
Percent Effect	79.1%	18.9%	-5.5%	-4.2%	-36.0%
Observations	748	748	748	748	748
Panel B: Treated units only include NYC academically selective high schools					
SES Policy Change X Post	0.105**	0.068***	-0.045	0.001	-0.025
	(0.051)	(0.016)	(0.033)	(0.030)	(0.021)
Percent Effect	27.3%	30.7%	-13.5%	0.3%	-35.6%
Observations	843	843	843	843	843

Notes: Estimation is based on equation 1. Panel A excludes treated New York City ASHSs, while Panel B excludes treated non-New York City ASHSs. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A7: Robustness to Sample Choices

	Main Estimates	Pre-2020	Pre-2021	No 100% FL Years	No 100% FL Schools	Only Top ASHSs	No Treated District Controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Share FL	0.135**	0.088	0.124**	0.104*	0.078	0.163**	0.147**
	(0.056)	(0.060)	(0.060)	(0.056)	(0.053)	(0.066)	(0.059)
Share URM	0.065***	0.057***	0.061***	0.064***	0.061***	0.070**	0.066***
	(0.017)	(0.013)	(0.015)	(0.017)	(0.017)	(0.035)	(0.018)
Share White	-0.042	-0.028	-0.047**	-0.043	-0.038	-0.028	-0.040
	(0.026)	(0.022)	(0.023)	(0.026)	(0.026)	(0.023)	(0.026)
Share Asian	-0.001	-0.005	0.006	0.001	-0.000	-0.017	-0.002
	(0.022)	(0.017)	(0.020)	(0.022)	(0.022)	(0.032)	(0.022)
Share Other	-0.024	-0.027	-0.024	-0.024	-0.026	-0.029*	-0.025
	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.017)	(0.019)
N	898	620	703	878	827	749	838

Notes: Estimation is based on equation 1. Column (1) presents the main estimates for reference. The sample in column (2) excludes all cohorts after 2019. The sample in column (3) excludes all cohorts after 2020. The sample in column (4) drops any observations with 100 percent share of students eligible for Free Lunch. The sample in column (5) drops any school that has at least one year with 100 percent share of students eligible for Free Lunch. The sample in column (6) keeps only the highest-ranking ASHS within each district when multiple ASHSs are treated within the district (effectively drops non-Stuyvesant ASHSs in NYC and non-Boston Latin ASHSs in Boston). The sample in column (7) drops all control ASHSs in districts with treated ASHSs from the analysis sample. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A8: Robustness to Additional Controls

	Main Estimates	Race Controls	Income Controls
	(1)	(2)	(3)
Share FL	0.135**	0.149***	0.133**
	(0.056)	(0.054)	(0.056)
Share URM	0.065***	0.063***	0.064***
	(0.017)	(0.017)	(0.017)
Share White	-0.042	-0.045*	-0.042
	(0.026)	(0.025)	(0.026)
Share Asian	-0.001	-0.001	-0.001
	(0.022)	(0.022)	(0.022)
Share Other	-0.024	-0.018	-0.024
	(0.018)	(0.015)	(0.018)
N	898	875	890

Notes: Estimation is based on equation 1. Column (1) presents the main estimates for reference. Column (2) includes controls for yearly changes in 9th-grade racial shares for the entire district. Column (3) includes controls for annual changes in 9th-grade low-income shares for the entire district. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table A9: Robustness to Alternative Specifications

	Main Estimates	Stacked DiD	No De-Trending
	(1)	(2)	(3)
Share FL	0.135**	0.081***	0.045
	(0.056)	(0.023)	(0.029)
Share URM	0.065***	0.054***	0.028**
	(0.017)	(0.015)	(0.012)
Share White	-0.042	-0.039	-0.025
	(0.026)	(0.024)	(0.019)
Share Asian	-0.001	-0.003	-0.005
	(0.022)	(0.022)	(0.015)
Share Other	-0.024	-0.010	0.002
	(0.018)	(0.013)	(0.004)
N	898	3615	898

Notes: Estimation is based on equation 1. Column (1) presents the main estimates for reference. The model used for column (2) is a stacked difference-in-difference approach that creates a sub-experimental dataset with centered control observations for each treatment year, then estimates a weighted regression with the main model (Wing et al., 2024). The model for column (3) does not include school-level pre-period de-trending. Standard errors clustered at the school district level are in parentheses. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$

Table A10: p-values from Permutation Tests—100 iterations

	Share Free Lunch	Share URM	Share White	Share Asian	Share Other
	(1)	(2)	(3)	(4)	(5)
Percent larger than baseline	0.06	0.04	0.05	0.73	0.12

Notes: We randomize placebo treatment dates 100 times for each outcome and calculate the percent of cases that generated policy impact estimates larger than our baseline estimates. Refer to section 4.4.4 for details.

Table A11: Effects on Academic Performance, by policy type

	Avg. ACT Score		
	Low-Income Quotas	ELC	SCA
	(1)	(2)	(3)
Panel A: Reduced form estimates (academic performance)			
Admissions Change	-1.425***	-0.360	-0.110
	(0.482)	(0.238)	(0.313)
Percent Effect	-5.0%	-1.2%	-0.3%
Standardized Effect	-0.28sd	-0.07sd	-0.02sd
N	439	421	465
Panel B: IV estimates (SES composition)			
FL Enrollment Shares	-1.570**	-0.647	-0.270
	(0.708)	(0.482)	(0.705)
Percent Effect	-5.6%	-2.1%	-0.8%
Standardized Effect	-0.31sd	-0.13sd	-0.05sd
N	439	421	465
Panel C: IV estimates (racial composition)			
URM Enrollment Shares	-1.456***	-0.721	-0.294
	(0.451)	(0.483)	(0.841)
Percent Effect	-5.2%	-2.3%	-0.9%
Standardized Effect	-0.29sd	-0.14sd	-0.05sd
N	439	421	465

Notes: Estimation in panel A is based on equation 1. Panels B and C report 2SLS coefficients where SES Policy implementation is the instrument and either FL Enrollment Shares or URM Enrollment Shares is the endogenous variable, respectively. IV estimates are multiplied by 10 for ease of presentation. Average ACT Scores are based on the average between ACT and the converted SAT scores, when available, as described in section 2. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Standardized effect shows how the average ACT score changed relative to the standard deviation of the treated ASHSs in the period before treatment. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table A12: ACT Impacts - Robustness to Sample Choice and Specification

	Main Estimate	Non-NYC Treatments Only	NYC Treatments Only	Stacked DiD
	(1)	(2)	(3)	(4)
SES Policy Change X Post	-0.529	-0.913*	-0.415	-0.999***
	(0.326)	(0.505)	(0.353)	(0.243)
Percent Effect	-1.8%	-3.2%	-1.4%	-3.5%
Observations	501	429	484	1449

Notes: Estimation is based on equation 1, using an unbalanced sample. Percent effect shows the effect relative to the mean of all treated ASHSs in the period before treatment. Average ACT Scores are based on the average between ACT and the converted SAT scores, when available, as described in Section 2. Column (1) replicates the main ACT estimates. In column (2) the treated units only include NYC ASHSs. In column (3) the treated units only include non-NYC ASHSs. Estimates in Column (4) are from a stacked difference-in-difference approach that creates a sub-experimental dataset with centered control observations for each treatment year, then estimates a weighted regression based on Wing et al., 2024. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Appendix B Policy Changes

B.1 Low-Income Student Quota Admissions Policy

B.1.1 New York City Diversity in Admissions Program (New York, NY)

New York City has sought to increase diversity in its non-specialized schools through prioritization of students with free or reduced-price lunch status. In 2012, PS 133 William A. Butler School in Brooklyn became the first to use this approach, setting aside 35 percent of its seats for FRPL and English language learner (ELL) students.[46]

In the years following the implementation of this policy at PS 133, pressure mounted on the city to make efforts to integrate its public school system. A report from the UCLA Civil Rights Project found that the state of New York had the most segregated school system in the United States.[47] In 2014, a coalition of students founded IntegrateNYC—a non-profit which has since filed a lawsuit against New York City's admissions process.[48] Additionally, the state education department enacted the Socioeconomic Integration Pilot Program (SIPP) which offered up to $10 million from 2015-2018 to support progress toward integration plans and implementation at Title I focus schools in the state.[49]

Beginning with the 2016-17 school year, New York City launched the Diversity in Admissions pilot program at seven schools. Under this program, these schools set aside seats for priority groups such as those eligible for free or reduced-price lunch and English language learners.[50] District leaders were not allowed to explicitly target race in admissions because of the 2007 Parents Involved SCOTUS case. The NYC Education Department followed up the Diversity in Admissions pilot with a diversity plan[51] in June 2017 that set out the city's first publicized diversity goals.[52] One tenet of the city's diversity plan was to expand the Diversity in Admissions program. It proposed outreach to school leadership and communities to spread awareness of the program and its application process.

Building on the pilot program of 2016-17, the Diversity in Admissions program grew from seven to 19 schools for the 2017-18 school year. An additional 11 schools joined the

[46] https://www.dnainfo.com/new-york/20141210/park-slope/how-park-slope-school-is-fighting-stay-diverse/

[47] https://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/ny-norflet-report-placeholder/Kucsera-New-York-Extreme-Segregation-2014.pdf

[48] https://ny.chalkbeat.org/2021/3/9/22321193/nyc-students-file-lawsuit-taking-aim-at-admissions-screens-school-segrega

[49] https://www.p12.nysed.gov/funding/2015-18-title-1-ses-integration-grant/ses-integration-grant.pdf

[50] https://static1.squarespace.com/static/53ee4f0be4b015b9c3690d84/t/5c4f765088251bb7f8e4fd6c/1548711510681/PromisingOutcomes.pdf

[51] https://web.archive.org/web/20170611214002/http://schools.nyc.gov/NR/rdonlyres/D0799D8E-D4CD-45EF-A0D5-8F1DB246C2BA/0/diversity_final.pdf

[52] https://www.nytimes.com/2017/06/06/nyregion/long-awaited-plan-for-integrating-schools-proves-mostly-small-bore.html

program in 2018-19, including the first academically-selective high school in NYC to set aside seats for FRPL-eligible students—Bard High School Early College Queens. The program has continued to grow, with 43 high schools participating in 2023-24.[53] This includes 13 schools that are documented as ASHSs by Finn and Hockett (2012).

Mader et al. (2018) provide a descriptive analysis of the changes in enrollment demographics in participating schools in the first two years of the Diversity in Admissions Initiative (none of the schools during this time were ASHSs). They showed significant increases in low-income enrollment, but no statistically significant change in any race/ethnicity category on average across the pilot schools.

B.1.2 Liberal Arts and Science Academy (LASA) High School (Austin, TX) – Diversity Process

Until 2017, LASA, located in Austin, Texas, had admitted students based solely on their performance on the school's admissions exam, past grades, an application essay, and teacher recommendations. For the cohort applying for fall 2017 admission, the school considered applicants' race and socioeconomic status for the first time. At this point, LASA shifted to an admissions model in which 80 percent of students were admitted based on their holistic application score. The remaining 20 percent of admitted students are chosen based on their application scores, contingent on meeting at least three out of the five special circumstances defined by the school. These circumstances are being a future first-generation college student, being zoned to a neighborhood middle school with at least 70 percent economically disadvantaged students, belonging to an underrepresented racial or ethnic group, demonstrating academic achievement in the face of difficult circumstances, and embracing academic challenges.[54]

In addition to amending the admissions criteria, LASA also sent out a recruitment team made up of parents of LASA students during the 2017 admissions season. This group met with students at all Title I middle schools in the district and provided guidance to those seeking admission to LASA.[55]

B.2 Middle School Class Rank Priority

Another type of policy mechanically increases the geographic diversity of admitted students by automatically granting admission to a specified number of students from each middle

[53] https://www.schools.nyc.gov/enrollment/enrollment-help/meeting-student-needs/diversity-in-admissions

[54] https://lasa.austinschools.org/acceptance-procedures

[55] https://www.kxan.com/news/lasa-aims-to-diversify-campus-with-new-recruiting-application-process/

school in a district. This approach dictates that students from every school have an opportunity to be admitted to the selective high school.

B.2.1 Academic Magnet High School (North Charleston, SC) – Top Two Program

Academic Magnet High School in North Charleston, South Carolina implemented this approach through a "Top Two Program" that first applied to the fall 2016 cohort.[56] In this admission system, Academic Magnet High accepts the top two students, by grade point average, from every public middle school in the district. The offer of admission is conditional on the students achieving a minimum score on the school's application rubric which includes grades, standardized test scores, a math task, a writing task, and a teacher recommendation.

B.2.2 Thomas Jefferson High School (Alexandria, VA) – Top 1.5% Policy

In December 2020, the Fairfax County School Board voted to change the admissions policy for Thomas Jefferson High School for Science and Technology. At this meeting, two alternative methods for selecting students were proposed. First, a proposal to enact a Hybrid Merit Lottery—which would have admitted the top 100 scoring applications, while allocating the remaining 450 seats by lottery among qualified applicants—was rejected. The Board then approved an admissions policy that admits the top 1.5 percent of students from each middle school based on their application scores. This leads to an allocation of approximately 450 seats, with the remaining 100 seats filled by the highest-scoring applicants who have yet to be admitted. This policy has been subject to several recent legal challenges, with the U.S. Court of Appeals for the Fourth Circuit finding in favor of the School Board in May 2023.[57] There have been calls for the case to be heard by the Supreme Court.[58]

B.3 Geographic Rank Priority

B.3.1 Lowell High School (San Francisco, CA) – Priority Middle School

Lowell High School began its current practice of setting aside 15 percent of seats for students from underrepresented middle schools in 2001. In October 2018, the San Francisco Unified School District Board of Education approved an update to the eligibility criteria under this third band of their admissions system.

[56] https://www.postandcourier.com/news/academic-magnet-continues-offering-admissions-to-top-2-graduates-of-middle-s article_65c57dbe-e250-11e6-9a39-171418065599.html

[57] https://www.ca4.uscourts.gov/opinions/221280.P.pdf

[58] https://www.nytimes.com/2023/08/21/us/politics/supreme-court-thomas-jefferson-high-school-admissions.html

Lowell analyzes the enrollment totals at each middle school in San Francisco, total enrollment in the district, and the number of applicants and admitted students from each school over the three preceding years. With these numbers, Lowell identifies a subset of middle schools that are underrepresented within Lowell High School. Under this framework, if an underrepresented school becomes especially proficient at placing students at Lowell, they can lose their underrepresented status. This update identified Willie Brown Middle School as a Band 3 school, regardless of its students' representation within Lowell. The change went into effect for the cohort that entered Lowell for the 2019-20 school year.[59]

It was estimated that the policy would affect 11 current 8th graders at Willie Brown when it was enacted in October 2018.[60] The policy could also impact the composition of Willie Brown as middle school parents respond to the policy. For example, from 2020-2022, the share of Black students at Willie Brown has fallen from 33 percent to 21.9 percent and the share of white students has risen from 10.5 to 18.0 percent. The share of students qualifying for Free or Reduced-price Lunch fell from 66.5 to 52.0 percent.[61]

Lowell instituted lottery admissions which affected the 2021-22 and 2022-23 entering cohorts before returning to selective admissions for 2023-24. Therefore, we only estimate the effects of the 2019 policy change on the 2019 and 2020 cohorts.

B.3.2 Goldstein High School for the Sciences (New York, NY) – Priority for Underrepresented Zip Codes

In 2019, Leon M. Goldstein High School for the Sciences planned to start a Diversity in Admissions Initiative, which would be implemented starting in 2020.[62] Their initiative was to prioritize students from underrepresented zip codes. To do this, they rank order students within each zip code and admit the top ranking students in each zip code, improving the odds of admissions for students who were historically less likely to be admitted.[63]

B.3.3 Boston Public Schools – Socioeconomic Tiers Based on Address

Boston Public Schools operates three academically selective high schools: Boston Latin, Boston Latin Academy, and the O'Bryant School of Mathematics and Science. In 1997, these schools were legally mandated to remove their racial quotas for incoming cohorts. Starting with the fall 2021 cohort, BPS instituted a new admissions system that divides the city's

[59] https://go.boarddocs.com/ca/sfusd/Board.nsf/goto?open&id=AZS28M7BDEA9

[60] https://www.sfchronicle.com/education/article/SF-school-board-changes-Lowell-admission-policy-13294817.php

[61] https://drive.google.com/drive/folders/1MQg9Lrdw0xKxc8FZ8eD2No67kB5P4rYq

[62] https://www.nydailynews.com/2019/10/02/four-nyc-high-schools-adopt-new-admissions-methods-to-increase-diversity

[63] https://www.myschools.nyc/en/schools/high-school/22K535/

Census tracts into eight socioeconomic tiers. Applicants continue to be assessed based on their middle school grades and test scores; however, they are only compared to other students within their tier. An equal number of seats are allocated to each SES tier.[64,65]. Additionally, extra application score points are assigned to students experiencing homelessness, foster care, or public housing, as well as students that in the prior year attended a school that serves at least 40% low-income students. For fall 2021 only, the entrance exam was not used due to the COVID-19 pandemic, but the test was reinstituted, along with the socioeconomic tier system beginning with the fall 2022 cohort.[66] For the 2024 cohort, the extra points for students who attended a school with a high proportion of low-income students will depend on the tier the student lives in.[67]

B.4 Submarginal Conditional Acceptance (SCA)

B.4.1 New York City Specialized High School Discovery Program

The NYC Specialized High School Discovery Program is an alternative admissions option for disadvantaged students who scored within a certain range below the admissions cutoff on the Specialized High School Admissions Test (SHSAT). The main goal of this program is to increase diversity in New York's specialized high schools. Historically, specialized school participation in this program has been optional. A provision in the 1971 New York State law that required admission to the city's specialized high schools to be based solely on the admissions test allowed specialized high schools to sponsor a Discovery Program [68]. By 2015, the majority (5/8) of specialized high schools were still admitting some students through the Discovery Program. The exceptions were Stuyvesant, Bronx Science, and the High School for American Studies at Lehman College.[69,70] At that time, students were eligible for the program as long as they scored within a certain range below the admissions cutoff, were recommended by their middle school, attended and passed a summer preparatory program administered by the specialized high school, and were certified as disadvantaged, based on household income, family receipt of government aid, foster care status, or being recent immigrants who are classified as English language learners.

[64] https://www.bostonpublicschools.org/tier-map

[65] https://www.wbur.org/news/2021/07/15/new-exam-school-admissions-policy

[66] https://www.the74million.org/article/scotus-passes-on-hearing-affirmative-action-case-involving-elite-boston-schools/

[67] https://web.archive.org/web/20241008025235/https://drive.google.com/file/d/1i8R0AJO2o1wtlt7TrTRVMZX8UTyK24xs/view

[68] https://web.archive.org/web/20140623060948/http://nypress.com/benign-neglect/

[69] https://www.nytimes.com/2018/08/13/nyregion/discovery-program-specialized-schools-nyc.html

[70] Since the early 2000s, Stuyvesant and Bronx Science had terminated their participation in the Discovery Program https://web.archive.org/web/20140623060948/http://nypress.com/benign-neglect/.

During the 2016-2017 school year, the NYC DOE released a diversity plan that included their goals and existing strategies to increase diversity in NYC schools. Part of this plan was to expand the Discovery plan. This included restarting the Discovery program at the High School for American Studies at Lehman College in the summer 2016, at Bronx Science in summer 2017, and at Stuyvesant in summer 2018. Additionally, in summer 2016, they expanded the number of Discovery Program students admitted at Brooklyn Tech. Beginning in summer 2019, Discovery Program eligibility was restricted to require that all students have attended a high-poverty middle school in addition to being disadvantaged (low-income family, ELL student recently moved to NYC, and/or student in temporary housing). By 2019-20, 20 percent of seats at each specialized high school were reserved for Discovery Program participants.[71]

[71] https://www.schools.nyc.gov/docs/default-source/default-document-library/specialized-high-schools-proposal

Appendix C Addressing Multiple Policy Changes

To identify the effects of specific SES-based admissions policies, we must ensure that we are capturing "clean" policy changes. This mean that if many policy changes are happening at once, it would be difficult to identify which of those policy changes led to the outcome of interest. Therefore, we must deal with issues of this type when we're creating our analytic dataset of policy changes.

Nine ASHSs in Chicago were removed due to transitioning from a race-based affirmative action admissions policy to a place-based admissions policy. While we would be interested in the impacts of the SES place-based policy in Chicago, we believe the termination of the race-based admissions policy contemporaneously impacts admissions in a way that obscures the impact of the SES-based policy.

Additionally, before compiling our data to create Table 1, we removed schools with policy changes in consecutive school years that are impossible to disentangle from each other, as each new policy change may have a new intention/goal. This decision leads us to drop: Academy at Palumbo, Carver High School, Central High School, Julia R Masterman Secondary School, Parkway Center City Middle College, Paul Laurence Dunbar High School, Western High School, and Maggie L. Walker Governor's School.

In some cases, consecutive policy changes have the same intention/goal. We merge policy changes that occur in consecutive school years to be represented as a single policy change in the first year of incidence because all the follow-up changes were made to strengthen the adjusted policy using the same strategy as the initial change. These schools are: McKinley Technology High School, School Without Walls High School, Science Park High School, Baruch College Campus High School, Beacon High School, NYC Lab High School for Collaborative Studies, and School of the Future High School.

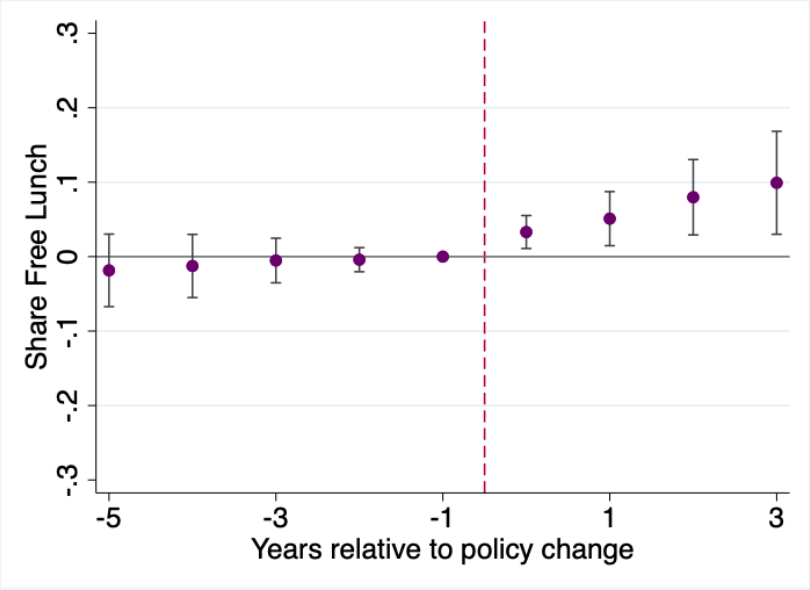
As described in section 2, we further restrict our analytic sample of treated ASHSs to those that are not contaminated by multiple policy changes. We allow ASHSs that had multiple admissions policy changes, as long as they meet our requirement that the treatment timing does not overlap with a previous or future admissions policy change at the same school. That is, we require that the policy change must be at least six years after any previous policy change and at least one year prior to any future policy change. In practice, no school appears more than once in our analysis sample, though many schools have multiple policy changes.

Appendix D Dose-Response vs. Traditional DID and Event Study Estimates

Appendix Figure D1 shows an event study that is analogous to that of figure 4, but without dose-response. The event study shows a small, but immediate, deviation at period zero that increases over time. However, each post-shock coefficient has a different interpretation since only a portion of the school is treated and, hence, must be rescaled for comparability.

Appendix Table D1 shows that DID estimates that do not incorporate a dose-response are mathematically equivalent to those calculated by the dose-response DID approach as long as they are rescaled appropriately. Column (1) shows dynamic DID coefficients estimated in each post-shock period without accounting for dosage. Column (2) shows the proportion of the typical high school that is treated with the additional cohort entering under the new admissions policy. In column (3), we rescale the corresponding estimate by the proportion of cohorts treated, which provides the estimated impact if the whole unit of observation had been treated. In column (4), we show the estimated dynamic impacts from our main approach that incorporates the dose-response. Column (1) and (4) show that both approaches provide estimates with the same level of statistical significance.

Figure D1: Event Study Estimates (without dose-response) – Free-Lunch Enrollment Shares



Notes: Event study is based on equation 2, but without dose-response, therefore $AdmissionsChange_{st}$ is always set to 1 for all treated units. Pre-period trend is estimated and removed from all observations for each school prior to the event study estimation, as described in section 3. For treated schools, this is estimated using all available pre-shock years for each school. For control schools, we use the 2010 to 2015 time-span to estimate the trend.

Table D1: Equivalence between dynamic DID estimates, without and with dose-response

	Estimates w/o Dose-Response	Proportion Treated	Rescaled Estimates	Estimates w/ Dose-Response
	(1)	(2)	(3)	(4)
Panel A: Impact on Free Lunch Enrollment Share				
SES Policy Change X Post0	0.028*	1/4	0.112	0.111*
	(0.014)			(0.057)
SES Policy Change X Post1	0.048**	2/4	0.096	0.095**
	(0.021)			(0.042)
SES Policy Change X Post2	0.077***	3/4	0.103	0.103***
	(0.027)			(0.036)
SES Policy Change X Post3	0.102***	4/4	0.102	0.102***
	(0.038)			(0.038)
Observations	898			898

Notes: Estimates are based on dynamic DID specification, which allows for a separate estimate in each post-shock period. We keep observations with relative time up to six periods after treatment (or after 2016 for control units) in the analysis sample, and we bin relative time periods greater than or equal to four. Refer to section D for details. Standard errors clustered at the school district level are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$