The Short-Run Effects of Boston's Housing Based School Assignment Policy on Academic Outcomes

By Skyler Lee and Summit Sarkar

We exploit differences in treatment across time, geography, and grade to estimate the effects of Boston's HBAP school assignment policy on academic outcomes. Our DDD results show that HBAP had positive, yet statistically insignificant total effects on both ELA and math test outcomes. While most estimates are small and insignificant, we find that black 3rd graders see a significant 14.29pp and 9.95pp increase in "meeting" or "exceeding expectations" on MCAS math and ELA exams, respectively. Our results suggest that HBAP, which sought to shorten commutes and ensure equitable access to high-quality schools, disproportionately benefited young black students.

Introduction:

Public schools are public goods. Allocating such resources among communities has always remained a contentious affair, especially due to histories of segregation and the outsized impact of *desegregation* upon life outcomes (Johnson, 2015). But as the bitterest aspects of segregation fade from memory, some school districts have revised their complex, cross-district busing school assignment policies, instead focusing on proximity and convenience. Yet secondary questions exist as well: to what extent must parents be involved in choosing schools for their children? To what extent do concerns regarding equity relate to school choice? And importantly, to what extent do school assignment policies play a role on student's outcomes?

The gold standard of previous literature on school assignment policies involves North Carolina's public school districts, such as the Wake County Public School random school assignment study (Hill, 2023) that found evidence that when low income students are assigned to high-achieving schools, all students benefit in test scores. But there are two shortcomings with these studies in North Carolina. First, results from suburban North Carolina may not be generalizable to other regions of the country, like dense urban areas and demographically different populations. More importantly, Charlotte-Mecklenburg Schools (CMS) from Billings (2014) followed a restrictive assignment system, with administrative barriers for parental application to non-neighborhood schools. In contrast, Boston Public Schools designed a more nuanced algorithm with a basket of school options, presenting ranked choice, followed by lottery, to parents. Does such a medium ameliorate the racial inequality gap found in CMS?

Furthermore, there is a substantial gap in the literature for Boston. The only evaluation of the HBAP was conducted in 2018 by Northeastern's Boston Area Research Initiative (BARI), which found that HBAP shortened commutes but increased racial segregation. Yet BARI did not analyze academic performance, making it our outcome of focus.

This paper uses a triple difference (DDD) approach, exploiting variation in treatment across time, geography, and within-Boston grades (due to staggered rollout) to estimate the causal impact of Boston's HBAP policy on the academic outcomes of elementary and middle school students in Boston. For our three differences, we have pre and post treatment, Boston versus the untreated control cities (Springfield, Worcester, and Lowell), and grades 3-4, 6-8 versus the untreated 5th grade. From the Massachusetts Department of Elementary and Secondary Education (DESE), we leverage comprehensive school-grade-year MCAS standardized test data from 2010-2019 for every public school in Massachusetts. We also have access to outcomes by race, sex, and income status for heterogeneity analysis. Our outcome of interest is the percentage of students who are "meeting" or "exceeding expectations" on MCAS exams. We find no significant effect on overall academic achievement, but observe some significant positive effects for black students and younger students. This provides an alternative reference point in the literature and discussion of racial academic gaps in the wake of neighborhood-centric school assignment policies. Our results contribute to the literature as we are the first to study the short-run effects of Boston's 2014 unique school assignment policy on academic outcomes. We also use a defensible DDD methodology that leverages within-Boston differences in treatment among grades to validate our results.

Background:

The Racial Imbalance Act of 1965, passed by the Massachusetts State Legislature, required any public school in the state with a majority minority population (>50% non-white) to desegregate to retain state funding. By 1988, students enrolled in Boston Public Schools (BPS) dropped from 100,000 to 57,000, the product of "white flight." To illustrate: at the start of forced, inter-zone busing and integration in 1974, BPS was composed of 57% white students; today, the figure is closer to 14%.

Up until 2013, BPS assigned schools via the "3-zone system" (3Z), which entitled children to apply to any school in one of three large geographic zones to which they belonged (see Figure 1). Then, with concerns about inequitable access to 'high-quality' schools and the long commutes that many students faced, BPS did away 3Z, approving the Home-Based Assignment Plan (HBAP), a novel elementary and middle school assignment plan with two goals in mind: equal opportunity to high-quality education and greater access to schools close to home. Since BPS was already overwhelmingly minority-enrolled, the focus of the policy was not on segregation but on equity and efficiency. This paper examines the extent that HBAP increased student overall academic outcomes.

HBAP uses an algorithm to assign each family a customized basket of 10-18 schools, determined by their home address. Each basket is guaranteed to contain the closest: two Tier 1 schools (top 25% school), two Tier 1/Tier 2 schools (top 50%), and two Tier 1/Tier 2/Tier 3 schools (top 75%). Tiers are determined largely by Massachusetts Comprehensive Assessment System (MCAS) standardized test scores, with a Tier 1 school having the top 25% of MCAS scores, Tier 2 having top 50% scores, etc. After ranking the schools by preference, students are then assigned schools with oversubscribed schools determined by a lottery system. Unlike the old 3-zone system, there are no sub-districts.

Under HBAP, each family is guaranteed the chance to attend the closest Tier 1 school. On face value, these individualized baskets of schools give children equal access to quality local schools. The algorithm did not consider race, gender, or income.

HBAP was rolled out to all of Boston public schools starting in the 2014-15 academic year, affecting the school assignment of kindergarteners and 6th graders that year (affecting approximately 9,000 students annually). Each subsequent academic year we see the treatment of two additional cohorts: the incoming kindergarteners and 6th graders. Figure 2 shows which grades were impacted by HBAP at a given school year, which grew to include opt-in internal and external transfer students over time. Students already enrolled in a school before HBAP implementation were guaranteed that they could remain at that school until they 'graduate' into middle school. For example, a 2nd grader in 2015 would not be subject to the HBAP assignment policy until they reach 6th grade. Importantly, our data sample includes years up to 2020, which is when we would see 5th graders be treated. 5th graders are subsequently never treated in our sample, and serve as our within-Boston control group.

Literature Review

Our paper examines the impact of the Boston Public School's Home Based School Assignment Program (HBAP) on student test outcomes.The only evaluation of the HBAP was conducted in 2018 by Northeastern's Boston Area Research Initiative (BARI), which found that HBAP shortened commutes but increased racial segregation. O'Brien (2018) had access to a confidential dataset, with access to each student's choice basket, ranked list, and eventual enrollment. Student home addresses were also matched, allowing researchers to construct the counterfactual attended schools. Findings of note include “neighborhoods that saw the greatest proportional reductions [in open seats] were in the southern urban core,” which also “had the fewest Tier 1 schools and seats on average.” Our research seeks to advance on this point: how does the heterogeneous competition for higher-quality schools induced by HBAP’s geography-centric design affect overall academic achievement in Boston?

In the sub-field of school reassignment and desegregation/segregation research, the Charlotte-Mecklenburg school (CMS) district may be the most popular due to its size and number of significant policy changes. In Fall 2002, CMS ended its thirty-year busing program, switching to a neighborhood-based school choice plan. It was similar to HBAP since a neighborhood school option was guaranteed, but rather than generating a customized basket per family, the burden was upon parents to individually apply to other schools. The identification strategy that Billings (2014) used compared students within neighborhoods yet placed in different schools due to boundary lines. The researchers found that resegregation led to an increase in racial inequality, with the racial gap in math scores increasing by 0.025 SD, and nonminority students assigned to students with more minority students were significantly less likely to graduate or attend a four-year college.

In another county in North Carolina, Hill (2022) studies the Wake County Public School System’s random school reassignment policy that sought to improve education outcomes by (1) balancing across schools the proportion of economically disadvantaged and academically struggling students and also (2) alleviating overcrowding in schools. Students within a specified geographic area were randomly assigned to a new school. Hill conducts an RCT and shows that students that remain in their school see increases in test scores when peers’ achievement level increases (peer demographic effects). But students that moved to a new school see little to no effect on test scores (disruption effects), even while holding constant the achievement level of peers. This study shows that school reassignment policies focused on increasing equity, like HBAP, may also be effective in improving education outcomes. However, there are fundamental differences between the school assignment policies of Wake County and Boston, as well as differences in characteristics of students (with more low income and racially diverse students in Boston). Thus, our paper seeks to study if, despite these differences, HBAP would increase student outcomes.

Campos (2023) studies the early-2010s Los Angeles Unified School District Zone of Choice (ZOC) policy that expanded the options of schools that students could choose from. The ZOC policy was enacted in select areas, targeting socioeconomically disadvantaged communities. Similar to HBAP, students in ZOC areas submit a ranked list of schools, where oversubscribed schools determine enrollment by lottery. Leveraging variation over time and across geographic zones (DD model), the authors estimated that ZOC had large effects on academic and college outcomes. Like HBAP, ZOC’s goal was to increase access to high-quality schools. However, ZOC increased the number of options available to students while HBAP redefined the options.

Abdulkadiroglu (2020), using structural models, studies the links between parent/student preferences to “value added” treatment effects on test scores, graduation, college attendance, and college choice of approximately 250,000 applicants in NYC public high school ranked-choice admissions. The paper concludes that applicants successfully identify peer quality, with the average math score of the first choice being 0.2 SD over the city average, and scores monotonically declining with lower rank. After conditioning on peer quality, parents' choices are unrelated to most measures of school effectiveness. The results provide further ambiguity in evaluating whether public school choice is preferable to administrative constrainment.

Similarly, Rothstein (2006) empirically tests structural models using the National Educational Longitudinal Survey 1988 and 330,000 metropolitan SAT test-takers from the high school class of 1994 to identify a potential relationship between inter-district choice and peer effects, finding no evidence that peer effects vary systematically with a district's degree of school choice–the performance gap between the least and most desired schools in a district are not meaningfully larger in districts with less parental choice, and vice versa. This finding is aligned with Abdulkadiroglu (2020)–parents may not need to be so involved if the assessing metric of a policy is centered around academic outcomes, potentially assuaging the fears of 3Z (HBAP's predecessor) advocates.

In conclusion, there is a substantial gap in the literature for Boston. The only evaluation of HBAP, apart from BARI, only used data through 2017 and did not analyze academic performance—our study focuses on academic outcomes through 2020. Even if not the main goal of HBAP, academic performance is critical when evaluating education policies. Even among school reassignment policies, HBAP is novel as it reduced school choice in the pursuit of equality, but not to an onerous degree: it presented roughly a dozen options for applicants to rank–unlike CMS. Also, HBAP affected a student population (around 50,000) greater than most studies, though not as large as CMS. Though, we do not improve on the internal validity of the Wake County random assignment experiment, our triple difference methodology (leveraging variation across time, geography, and grades) improves on the internal validity of the LA DD study.

<u>Data:</u>

This paper seeks to estimate the causal impact of Boston Public Schools' (BPS) HBAP school reassignment policy on math and english test scores of 3rd, 4th, 6th, 7th, and 8th graders through 2019. From Massachusetts' Department of Education and Secondary Education (SEDE) website, we have school-grade-year level data of the statewide Massachusetts Comprehensive Assessment System (MCAS) standardized test scores for each student demographic. With this data, we will conduct a DDD to study the academic effects of HBAP. Springfield, Worcester, and Lowell (the three largest districts after Boston) schools are used as geographic controls and untreated cohorts (5th graders) are within-Boston controls. The "On Grades" column of Figure 2 summarizes which grades in a given year are treated.

The SEDE website contains comprehensive school level data for every public school in Massachusetts, going back as far as the 2000s. Information includes annual reports on student demographics, teacher descriptors, test scores, etc. With MCAS achievement level data, our outcome variable is the percent of each school-year-grade "meeting expectations" and "exceeding expectations" on a given MCAS exam. Figure 3 shows how these achievement levels are defined, in terms of actual MCAS scores. SEDE also reports this MCAS data for each student group, which includes: economically disadvantaged, Black/White, Female, etc.. Economically disadvantaged is measured by participation in state-administered welfare programs, such as SNAP, TAFDC, etc. SEDE also gives school-by-grade-by-year characteristics like student-teacher ratios, annual city unemployment rates, and annual per-pupil district expenditures, which we use as our covariates.

The PARCC exam, administered in 2015-2016 in lieu of MCAS, severely limits usage of the SEDE data. PARCC reports achievement levels on a different scale than MCAS, and does not provide demographic subgroup data. We are unable to convert between MCAS and PARCC outcomes due to fundamental differences in scaling and lack of data. Similarly, takeup of PARCC vs MCAS differs (likely nonrandomly) between schools and cities in these years. For example, according to Figure 4, several city-years experienced approximately a 2:1 or 1:1 ratio of MCAS/PARCC takeup and vice versa. It is likely that schools with more capable administrators managed this transition, as opposed to the less informed ones remaining on the MCAS, or fearing that an unfamiliar exam could harm academic performance. Thus, they are omitted from our dataset. Pre-periods and post-periods vary by grade. For 6th grade, pre-period is 2010-2014, and 2015-2019 serve as post-periods. For 7th grade, pre-period is 2010-2015 and post-period is 2016-2019. See figure 2 for a visual on which grades are treated when. Again, 5th grade is untreated between 2010-2019.

Unfortunately, the SEDE website does not have readily downloadable data—the data is contained in tables on various pages. To circumvent this, we wrote a simple Python script to scrape the MCAS by student group tables for every grade of every school in the Boston, Springfield, Worcester, and Lowell school districts for the years 2010-2019. Our webscraper extracted excessive information, so much of the data cleaning process involved removing extra variables, renaming columns, and splitting data up. For example, each MCAS table for a given school and year is titled “GRADE-LEVEL-03-ENGLISH-LANGUAGE-ARTS”, which we split into a grade and test category component. Similarly, some variables were titled differently depending on the year and city, such as “Ever EL” and “Ever ELL”, or “Afr. Amer./Black” and “African Amer./Black.”

We also drop unreported SYG subgroup scores, as SEDE does not report MCAS outcomes for student groups with fewer than 10 individuals for data privacy. This does not affect our overall estimates, where all students, regardless of subgroup data omission, are included in the “Total” data. However, such omission would bias our subgroup heterogeneity estimates if underrepresented students performed differently from their counterparts more represented in other SYGs.

After cleaning, we obtain achievement level percentages for every test subject, school, year, grade, and student group. We use student demographic MCAS data for subgroup heterogeneity analysis of female, Hispanic, economically disadvantaged, Black, and disabled students. We create our outcome variable, which is the percentage of students that are either “meeting” or “exceeding” expectations, as opposed to “partially meeting” or “not meeting” expectations.

Table 1 reports summary statistics for grades 3-8 in BPS and control districts between 2010-19. The demographics section describes the *weighted* mean student composition, accounting for differently sized cohorts, as our unit of observation is at the school-year-grade level. The differential number of observations, i.e., for Asians, indicates school-year-grades with zero Asian students. In Table 1, the Hispanic/Latino, Asian, Economically Disadvantaged, and

Male/Female proportions are remarkably balanced. However, BPS is home to considerably fewer white students and considerably more black students relative to the control districts.

Methodology:

This paper seeks to estimate the causal impact of the HBAP school reassignment policy on the MCAS scores of grades 3-4 and 6-8, via DDD. In the first year of HBAP, all kindergartners and 6th graders (in Boston) were treated while every other grade was not (they remain in their schools from the old assignment policy). The next year, 1st/7th graders are still treated in addition to the incoming K/6th grade. Figure 2 is a handy reference for which grades are treated in which years. In our entire 2010-19 sample, 3rd graders are treated in 2018-19 (the year students take the MCAS), 4th graders in 2019, 6th graders from 2015-19, etc. Each grade has different pre- and post-intervention periods. The variation in treatment—for a particular grade—by year is our first difference in our DDD model. 5th graders are never treated in our Boston sample and we use them for the second difference. The third difference comes from our geographic controls. We use schools and grades in other major Massachusetts cities, such as Worcester, Springfield, and Lowell, as our comparison group.

There are some strengths to our DDD design. A naive comparison of outcomes before and after HBAP could be biased by, for example, an upward trend in academic performance due to increased school funding within Massachusetts. Our geographic controls containing other cities in Massachusetts addresses this issue. But then a further concern would be a school funding shock within Boston. Perhaps, the city government decided to cut Boston public school funding after HBAP, and if increased school funding is predictive of increased academic outcomes, would downward bias our results. But our within-Boston control group, 5th grade, addresses Boston specific shocks that would otherwise bias a typical DD model that leverages time and geography.

Next, we explore the possibility of city specific shocks that would further violate our identifying assumption. First, other school districts could have implemented large-scale changes *à la* HBAP, either as a consequence of Boston's landmark implementation (spillover effects) or not. We have no reason to believe so, after repeatedly (for each city) Google searching "[City] public school assignment policy change," restricting each search to a one-year timespan so as to not miss any minor changes, between January 01 2009 and 2020. We confirmed this with deep research features of various LLMs. Second, it is possible that city or county-level economic shocks occurred in one or more cities and not others. The degree of adult employment could affect MCAS/PARCC performance. Thus, we control for city-level unemployment, collected in February for each year in the sample. Similarly, it is possible that cities had school budget windfalls or cuts in certain years, so we control for yearly per-pupil district expenditures.

Third, it is possible that HBAP directly results in sensitive incoming K/6 student families exiting Boston Public Schools, or such families relocating to wealthier neighborhoods with greater proximity to higher quality public schools. In the first instance, we can observe whether FRPL-eligible students significantly decreased following HBAP rollout. If we observe that BPS experienced a steeper decline in low-income students relative to the control cities after HBAP implementation in 2015, the internal validity of our results are threatened. Our estimates would be upwardly biased if low income students with poor academic outcomes moved out of Boston in response to HBAP.

Finally, it is possible that policy changes could affect younger or older students differently. For example, in the 2009 Fiscal Year in Boston, money was allocated to support English Language Learners, and in 2017, Massachusetts permitted bilingual education programs. If language learning is most effective for younger students (this Boston-specific shock affects grades differentially), then their academic growth rates would be greater, and our DDD would overestimate causal effects for grades below 5th grade and underestimate for grades above 5th grade. However, our estimates for the average total effect of HBAP across all grades would likely not be overly biased with this concern.

We pool analysis of all data and use OLS to estimate the following equation on the whole sample which contains grades 3-8 from 2010-2019 in the above mentioned cities. Our expanded equation is:

$$\begin{aligned} y_{syg} = \alpha_0 + \beta(Boston_s \times TreatedGrade_g \times Post_{yg}) \\ + \delta_{sy} + \gamma_{sg} + \phi_{yg} + \mathbf{X_{sgy}} + \varepsilon_{syg}, \end{aligned} \tag{1}$$

where s, y, g indexes school, grade, and year, respectively. The outcome variable is the percentage of students at or above “meeting expectations”. Boston and TreatedGrade are binary variables set to one if the school is in Boston or if the grade was ever treated by HBAP (grades 3-4 and 6-8 in our sample), respectively. Post is a dummy variable equal to 1 if the grade is 3rd grade and the year is at least 2018, or if grade is 4th and year is at least 2019, or if grade is 6th grade and year is at least 2015, etc. Delta, gamma, and phi are fixed effects that control for unobserved differences at the school-year, school-grade, and grade-year level. Our results would not be driven by, for example, transitory illnesses/local spikes in crime that disproportionately affected outcomes of one school over a period of time. X is a vector of school-by-grade-by-year characteristics which includes student-teacher ratios, annual city unemployment rates, and annual per-pupil district expenditures. Epsilon is a random error term. Our coefficient of interest is beta which we may interpret (with our identifying assumption) as the average (across all grades) percentage point change in students meeting or exceeding expectations, due solely to the implementation of HBAP.

Results

This paper seeks to estimate the causal impact of Boston’s HBAP school reassignment policy on Math and ELA test scores on elementary and middle school students. Our outcome variable is the percent of students whose MCAS achievement level is “meeting” or “exceeding expectations”. Incomplete MCAS administration years are removed (PARCC administered in these years), per Figure 4.

We conduct empirical tests to support our identifying assumption: that the relative change in achievement levels between grades in Boston and control cities, as well as between a given grade and 5th grade (5th grade is never treated in our sample), is driven by the implementation of HBAP alone. We show this by plotting by-grade event study coefficient estimates for our full sample (between 2010-19) in Figure 5. Coefficients that seem random or are zero before the treatment (represented by the vertical line) give support to the validity of our test. However, the pretrends show a potential violation of our identifying assumption—particularly 7th and 8th grade's very high pretreatment coefficients. Though, 3rd, 4th, and 6th graders do not egregiously violate our identifying assumption. These event study estimates suggest that our DDD may have some potential biases.

Another threat to identification is knowledge of the HBAP implementation differentially affecting the behavior of Boston families. We surmise that the only group this could meaningfully affect is academically-focused, poorer families–those who are skeptical of high-tier school guarantees when their own neighborhood schools are inferior, thus moving outside of Boston Public Schools, such as a private school or to a public school in the suburbs. If they relocated to another BPS school, this would not systematically bias our aggregate results. Since the share of low-income students post-HBAP (2015) in Boston fell relative to our control districts (see Figure 6), our DDD estimates may be downward biased (assuming it is low income, higher achieving students leaving Boston).

We hoped to add covariates such as unemployment rate, student-teacher ratio, and per-pupil district expenditures. However, our covariates were nearly perfectly collinear with our three sets of two-way fixed effects (further inspection of data shows that our covariates varied little to none in at least one of the school-grade-year dimensions). Thus, our main robustness check is to exclude each of the control cities, one at a time, to check how sensitive our results are to the geographic controls we picked. We use our heterogeneity (by-grade, by-demographic, by-grade-by-demographics) as further checks on our main results.

Our main results, total and by-grade estimates on Math and ELA outcomes, are reported in Table 2. BostonxGradexTreat is our treatment variable that represents whether a grade (in Boston) is treated by HBAP in a given year. Panel A shows that the total treatment effects (average change in percentage of students at or above "meeting expectations", relative to control groups) of HBAP are positive yet insignificant for both Math and ELA tests. Per Panel B, grade-specific effects are more varied, with 3rd, 4th, and 6th graders having slightly positive outcomes, and 7th, and 8th graders having slightly negative outcomes–all insignificant barring 4th graders who see positive and significant Math effects at the 10% level.

Table 3 reports by-demographic group total effect coefficients on Math outcomes. There was an 8.76pp increase (significant at 10% level) in Black students in the meeting/exceeding expectations category for Math, relative to Black students in untreated grades and control cities.

Table 4 reports by-grade-by-demographic effects on math outcomes. We observe larger treatment effect magnitudes when observing specific demographic groups at the grade level. Again, we observe 14.292pp more Black 3rd graders in Boston attain meeting/exceeding expectations Math scores (significant at the 5% level), relative to Black students in untreated grades and control cities. We also see 5.058 and 6.505pp more meeting/exceeding expectations for 3rd and 4th grade girls, respectively. Though, Asian 8th graders experienced a 12.957pp decrease in the upper achievement levels.

Table 5 reports total by-demographic effects on ELA outcomes, which were all statistically insignificant. In Table 6, we see several positive, significant by-grade-by-demographic treatment effects. Black 3rd graders improved their ELA outcomes (% meeting expectations or above) by 9.951pp, relative to control groups. Hispanic/Latino, Male, and Asian 6th graders improved their ELA outcomes by 5.874pp, 5.603pp, and 6.237pp respectively.

Finally, in Table 7 and Table 8, we run robustness checks by dropping each control city one at a time. We see that relative to our baseline DDD where we do not drop any cities, the coefficients are not significantly different. They remain statistically zero. This implies that our control cities do not drive much of our estimated causal effects of HBAP.

Conclusion:

Though HBAP's novel school assignment policy had clear goals in increasing equity towards accessing high quality education, it is not immediately obvious if the policy would improve student academic outcomes. Other economists have studied specific school reassignment policies and shown that all students benefit when low-income students move to high-achieving schools (Hill, 2023), but the implementation of the policy is much different from HBAP's addressed based policy. We thus sought to estimate the causal effect of HBAP on student outcomes using a DDD that utilized differences in treatment across years, cities, and grades. Using eight years of standardized exam achievement data at the school-grade-year level for elementary and middle schoolers in Massachusetts, our DDD approach yielded insignificant causal positive effects on academic outcomes from the Boston Home-Based Assignment Policy (HBAP). However, heterogeneity is important to recognize. Young black students are the beneficiaries of HBAP, with positive (and often significant) estimates across both ELA and Math test outcomes. At the same time, older black students–those in middle school–are the only racial demographic with negative (though insignificant) treatment effects across all three grades. This is in line with McDonough (2015), where "relative to white children, black children are less (more) likely to move up (down) through the distribution of test scores over time." HBAP suggests no sign of rectifying this issue, instead exacerbating it.

Yet black students are also the only racial group with significant positive effects as a whole (for Math). Altogether, this warrants discussion in the context of O'Brien (2018), which found that historically black neighborhoods still lacked high-quality schools relative to historically white neighborhoods despite a higher student density, and that "black students on average commute nearly 2 miles to attend a high-quality school—almost twice the distance traveled by white and Asian students." Despite the relative obstacles that black students in BPS faced as a result of HBAP, their overall MCAS achievement levels show no indication of statistically significant declines in academic outcomes. However, it is important to note we only possess one and two years of post-periods for fourth and third graders, respectively. Perhaps a longer-term timeframe would illustrate different results. Additionally, detailed student-level data, like that used in O'Brien (2018), would alleviate the internal validity concerns regarding unreported academic data for specific demographic groups.

Altogether, HBAP can be considered a qualified success. The goal was to reduce commute times and increase equity in access to high-quality schools, and we saw academic gains from young Black students. Our paper shows that school assignment policies can be a powerful policy lever in reducing education inequality. At worst, HBAP shortened commutes and did not differentially harm any racial group academically, unlike the expanding racial academic gap found in Billings (2014). Moreover, HBAP reduced annual administrative and bureaucratic burden, especially in the longer term. But our findings also reinforce the conclusion from Abdulkadiroglu (2020), where parents lack a unique capacity to identify particularly compatible schools for their children beyond superficiality. Our findings are in line with Rothstein (2006), as

HBAP's geographical constraining (relative to 3Z) did not significantly increase or decrease total academic effects. Our detailed empirical analysis of BPS, as opposed to Rothstein's broad nationwide approach, affords us greater internal validity, freeing us from the inhibiting assumptions that Rothstein's structural models assume.

All in all, the algorithmic nature of HBAP yields a fine medium between zero parental input and overbearing parental input, zero awareness of racial and economic histories and heavy-handed levels of social engineering–those that might otherwise continue fracturing neighborhoods and communities.

References

Abdulkadiroğlu, A., Pathak, P. A., Schellenberg, J., & Walters, C. R. (2020). Do Parents Value School Effectiveness? *American Economic Review*, *110*(5), 1502–1539. https://doi.org/10.1257/aer.20172040

Billings, S. B., Deming, D. J., & Rockoff, J. (2014). School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg. *The Quarterly Journal of Economics*, *129*(1), 435–476. https://doi.org/10.1093/qje/qjt026

Campos, C., & Kearns, C. (2024). The Impact of Public School Choice: Evidence from Los Angeles's Zones of Choice. *The Quarterly Journal of Economics*, *139*(2), 1051–1093. https://doi.org/10.1093/qje/qjad052

"Grandfathering" Included in Proposed Boston Public School Choice Plans. (2012, October 10). South End, MA Patch. https://patch.com/massachusetts/southend/grandfathering-included-in-proposed-boston-public-schb333d43660

Hill, D. V., Hughes, R. P., Lenard, M. A., Liebowitz, D. D., & Page, L. C. (2023). New schools and new classmates: The disruption and peer group effects of school reassignment. *Economics of Education Review*, *92*, 102316. https://doi.org/10.1016/j.econedurev.2022.102316

Johnson, R. C. (2015). Long-Run Impacts of School Desegregation & School Quality on Adult Attainments. NBER Working Paper Series No. 16664. https://www.nber.org/papers/w16664

Levinson, M. (2015). The ethics of pandering in Boston Public Schools' school assignment plan. *Theory and Research in Education*, *13*(1), 38–55. https://doi.org/10.1177/1477878515569034

News, S. (2018, July 24). *The "Home-Based" Assignment report: What should be done about it?* Medium. https://schoolyardnews.com/the-home-based-assignment-report-what-should-be-done-about-it-6211c61d9ee5

O'Brien, D., & Hill, N. (n.d.). *Evaluation of Boston Public Schools' Home-Based Assignment Plan*. https://www.boston.gov/sites/default/files/file/2020/05/Janey%20-%20BPS%20C3-4%20-%20Home-Based%20Assignment%20Plan.pdf

O'Brien, T., Hill, N., Contreras, M., Phillips, N., & Sidoni, G. (2018). *An Evaluation of Equity in the Boston Public Schools' Home-Based Assignment Policy*. The Boston Area Research Initiative. https://news.northeastern.edu/wp-content/uploads/2018/07/BPSHBAP.pdf

Rothstein, Jesse, M. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." American Economic Review 96 (4): 1333–1350. https://doi.org/10.1257/aer.96.4.1333

Tables and Figures

Table 1: Summary Statistics

City	Boston					Control				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
A. Demographics										
White	4053	0.170	0.165	0.008	0.778	5203	0.269	0.169	0.011	0.850
Black	4613	0.321	0.205	0.010	1.000	5209	0.143	0.082	0.011	1.000
HispanicLatino	4817	0.412	0.229	0.038	1.000	5289	0.439	0.189	0.023	1.000
Asian	2883	0.131	0.150	0.005	0.716	4152	0.136	0.143	0.005	1.000
EconDis	4831	0.715	0.197	0.125	1.000	5305	0.748	0.175	0.107	1.000
Disabled	4808	0.218	0.127	0.002	1.000	5286	0.197	0.083	0.014	1.000
Female	4810	0.489	0.076	0.037	1.000	5253	0.483	0.066	0.045	1.000
Male	4825	0.512	0.077	0.091	1.000	5305	0.517	0.067	0.137	1.000
B. ELA Outcomes										
% Exceeding Expectations	2418	18.901	13.782	0.000	100.000	2654	18.538	10.761	0.000	92.000
% Meeting Expectations	2418	39.198	15.440	0.000	85.000	2654	40.608	11.948	0.000	83.000
% Partially Meeting Expectations	2418	36.056	18.973	0.000	92.000	2654	35.416	13.853	0.000	80.000
% Not Meeting Expectations	2418	5.757	7.160	0.000	50.000	2654	5.432	6.012	0.000	61.000
C. Math Outcomes										
% Exceeding Expectations	2415	23.525	16.617	0.000	100.000	2653	25.160	14.691	0.000	100.000
% Meeting Expectations	2415	37.494	14.605	0.000	84.000	2653	39.473	12.602	0.000	84.000
% Partially Meeting Expectations	2415	28.736	15.592	0.000	80.000	2653	26.833	11.652	0.000	74.000
% Not Meeting Expectations	2415	10.148	12.442	0.000	69.000	2653	8.497	9.141	0.000	68.000

Notes: Sample consists of school-grade-year observations of third through eigth grades from 2010 to 2019. We have 191 schools total across Boston and the control cities (Worcester, Springfield, Lowell). All summary statistic variables are weighted against the number of students in a given school, grade, and year. ELA and Math outcomes are reported using the four MCAS achievement levels.

Table 2: DDD Estimates of Treatment Effects on MCAS Outcomes

	Math	**ELA**
Panel A: Total Effect		
Boston × Grade × Treat	1.888 (1.825)	1.252 (1.390)
Panel B: Grade-Specific Effects		
Boston × Grade 3 × Treat	3.526 (2.145)	1.220 (1.764)
Boston × Grade 4 × Treat	4.455* (2.429)	0.187 (1.899)
Boston × Grade 6 × Treat	0.385 (2.847)	3.814 (2.458)
Boston × Grade 7 × Treat	−1.912 (3.134)	−0.169 (2.822)
Boston × Grade 8 × Treat	−3.278 (3.196)	−0.729 (2.466)
Observations	4,970	4,976

Notes: Clustered (at school level) standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3: DDD Total Effects on Math Outcomes by Demographic Group

Group	**Boston × Grade × Treat**
Black	8.762* (4.550)
White	0.439 (3.088)
Hispanic/Latino	2.244 (2.215)
Asian	−0.558 (4.386)
Male	0.570 (2.102)
Female	2.741 (2.143)
Econ. Disadv.	1.856 (1.882)
Disabled	1.409 (1.971)
Observations (N)	
Black	2,170
White	1,928
Hispanic/Latino	3,934
Asian	759
Male	4,735
Female	4,628
Econ. Disadv.	4,846
Disabled	2,524

Notes: Clustered (at school level) standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: DDD Grade-Specific Effects on Math Outcomes by Demographic Group

Group	Grade 3	Grade 4	Grade 6	Grade 7	Grade 8
Black	14.292** (5.596)	7.149 (5.865)	−1.027 (7.778)	−2.009 (7.144)	−3.412 (7.562)
White	0.643 (4.365)	3.666 (5.731)	−1.620 (3.264)	4.350 (9.025)	0.133 (9.349)
Hispanic/Latino	1.633 (2.761)	3.987 (2.773)	2.373 (3.430)	1.135 (3.241)	1.716 (3.419)
Asian	6.714 (7.277)	−4.774 (10.144)	0.451 (4.154)	−6.478 (5.826)	−12.957* (6.929)
Male	2.063 (2.583)	1.853 (2.925)	−0.784 (3.119)	−1.386 (3.556)	−3.441 (3.637)
Female	5.058** (2.473)	6.505** (2.997)	0.108 (3.137)	−2.958 (3.569)	−3.831 (3.665)
Econ. Disadv.	3.100 (2.364)	2.725 (2.517)	0.679 (2.931)	0.207 (2.904)	−1.291 (2.995)
Disabled	0.321 (2.849)	2.636 (2.352)	0.558 (3.072)	1.738 (3.203)	3.868 (3.493)

Notes: Clustered (at school level) standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: DDD Total Effects on ELA Outcomes by Demographic Group

Group	**Boston × Grade × Treat**
Black	6.796 (4.321)
White	0.254 (2.894)
Hispanic/Latino	1.650 (1.953)
Asian	1.977 (3.507)
Male	0.808 (1.799)
Female	1.674 (1.713)
Econ. Disadv.	1.581 (1.606)
Disabled	−0.731 (1.747)
Observations (N)	
Black	2,173
White	1,933
Hispanic/Latino	3,949
Asian	763
Male	4,741
Female	4,634
Econ. Disadv.	4,850
Disabled	2,528

Notes: Clustered (at school level) standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: DDD Grade-Specific Effects on ELA Outcomes by Demographic Group

Group	Grade 3	Grade 4	Grade 6	Grade 7	Grade 8
Black	9.951* (5.293)	5.442 (4.842)	1.123 (6.402)	3.014 (6.874)	0.989 (6.387)
White	−0.417 (3.689)	−2.450 (3.724)	2.226 (4.163)	−1.220 (7.778)	1.599 (7.523)
Hispanic/Latino	0.042 (2.566)	−0.176 (2.556)	5.874** (2.886)	2.769 (3.065)	3.712 (2.730)
Asian	3.165 (5.637)	−5.312 (6.397)	6.237* (3.412)	0.616 (5.228)	−2.109 (4.505)
Male	−0.238 (2.060)	−2.894 (2.641)	5.603** (2.809)	1.401 (3.280)	1.745 (3.006)
Female	2.994 (2.374)	2.555 (2.333)	1.827 (2.949)	−1.685 (3.259)	−3.090 (3.050)
Econ. Disadv.	0.981 (2.111)	0.051 (2.045)	3.706 (2.808)	3.320 (2.888)	1.130 (2.520)
Disabled	1.240 (2.221)	−2.874 (2.427)	−0.082 (3.260)	−1.245 (3.425)	−3.647 (3.696)

Notes: Clustered (at school level) standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Robustness of DDD Estimates on Math Outcomes: Dropping Each Control City

Specification	**Boston × Grade × Treat**	**Std. Error**	**N**
Main Estimate (Table 2)	1.888	(1.825)	4,970
Drop Worcester	2.078	(2.093)	3,818
Drop Lowell	2.582	(1.930)	4,480
Drop Springfield	1.217	(1.962)	3,999

Notes: Standard errors clustered at the school level.
"Main Estimate" refers to full sample result reported in Table 2.

Table 8: Robustness of DDD Estimates on ELA Outcomes: Dropping Each Control City

Specification	**Boston × Grade × Treat**	**Std. Error**	**N**
Main Estimate (Table 2)	1.252	(1.390)	4,976
Drop Worcester	1.599	(1.615)	3,824
Drop Lowell	1.929	(1.520)	4,486
Drop Springfield	0.475	(1.505)	4,003

Notes: Standard errors clustered at the school level.
"Main Estimate" refers to full sample result reported in Table 2.

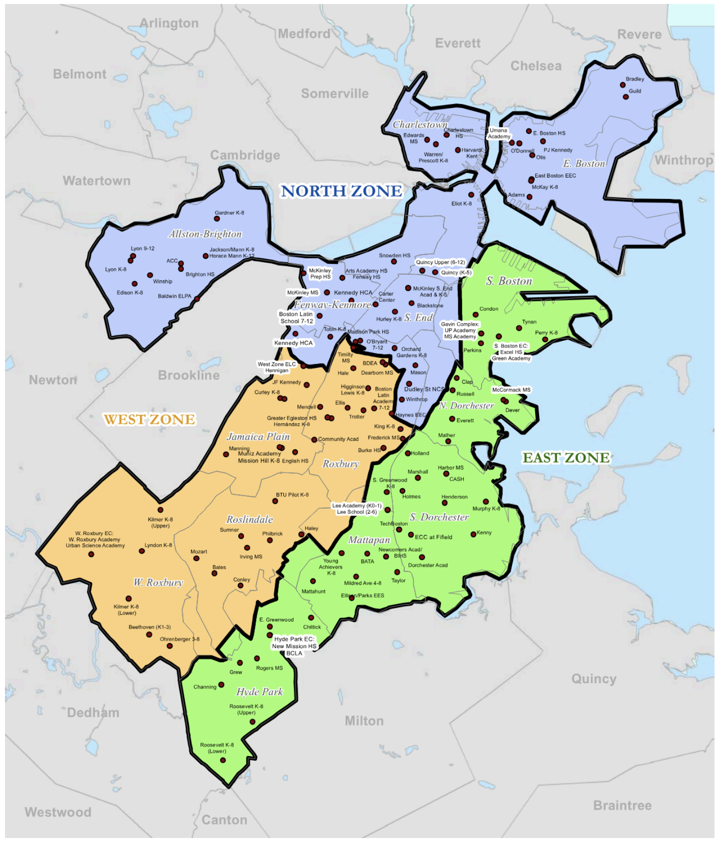

Figure 1: Former 3-zone system

On/Off Grades by School Year		
School Year	On Grades (Incoming Students)	Off Grades
2014/2015	Kindergarten and 6^{th}	1^{st} - 5^{th}, 7^{th} and 8^{th}
2015/2016	Kindergarten, 1^{st}, 6^{th} and 7^{th}	2^{nd} - 5^{th} and 8^{th}
2016/2017	Kindergarten, 1^{st}, 2^{nd}, 6^{th} – 8^{th}	3^{rd} – 5^{th}
2017/2018	Kindergarten, 1^{st} - 3^{rd}, 6^{th} - 8^{th}	4^{th} and 5^{th}
2018/2019	Kindergarten, 1^{st} - 4^{th}, 6^{th} – 8^{th}	5^{th}
2019/2020	All	None

Figure 2: HBAP by school year. “On Grades” show the grades affected by the policy.

Achievement Level	Scaled MCAS Score
Exceeding Expectations	530-560
Meeting Expectations	500-529
Partially Meeting Expectations	470-499
Not Meeting Expectations	440-469

Figure 3: Scaled MCAS score ranges for each achievement level

City	Year	% MCAS	% PARCC
Boston	2010–2014	100%	0%
	2015	1.9%	98.1%
	2016	0%	100%
	2017–2019	100%	0%
Lowell	2010–2015	100%	0%
	2016	0%	100%
	2017–2019	100%	0%
Springfield	2010–2014	100%	0%
	2015	36.9%	63.1%
	2016	0%	100%
	2017–2019	100%	0%
Worcester	2010–2014	100%	0%
	2015	45.9%	54.1%
	2016	32.6%	67.4%
	2017–2019	100%	0%

Note: Utilization of MCAS vs PARCC, per city per year. Results reflect percentages of school-year-grades administering each exam, not weighted for unit size.

Figure 4: MCAS vs PARCC utilization by year

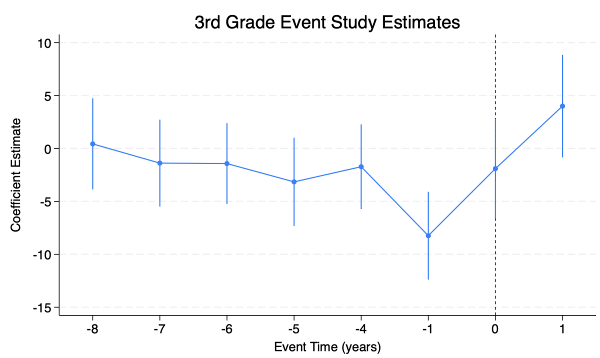

Figure 5a

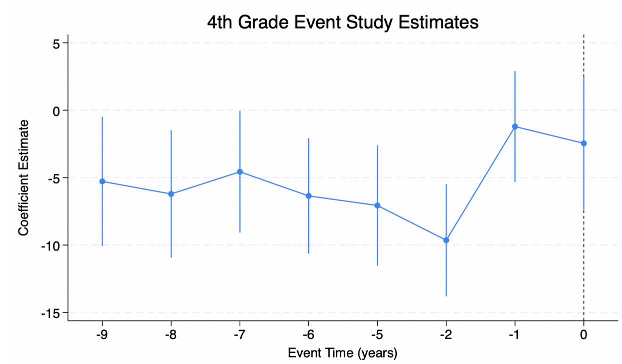

Figure 5b

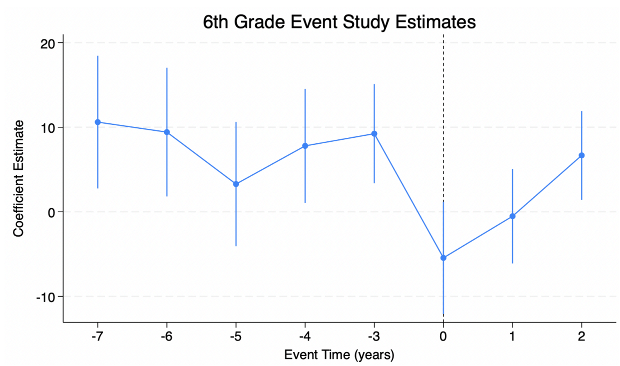

Figure 5c

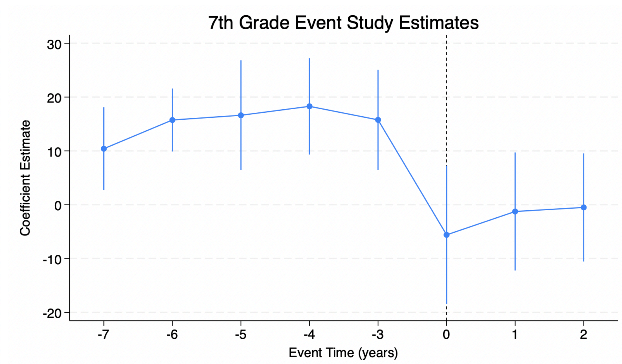

Figure 5d

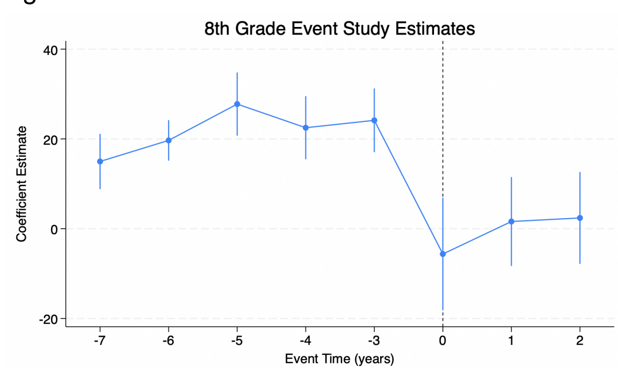

Figure 5e

Figure 5*:* Event study DDD coefficients plotted for each grade with years ranging from 2010-2019. For example, the 3rd grade plot represents the DDD with 5th grade and non-Boston cities as control groups. The vertical line reflects the first year of treatment for a grade. Discontinuities in the x-axis reflect the fact that 2015-16 are dropped due to incomplete non-random takeup of the PARCC exam, which we lack serviceable data for.

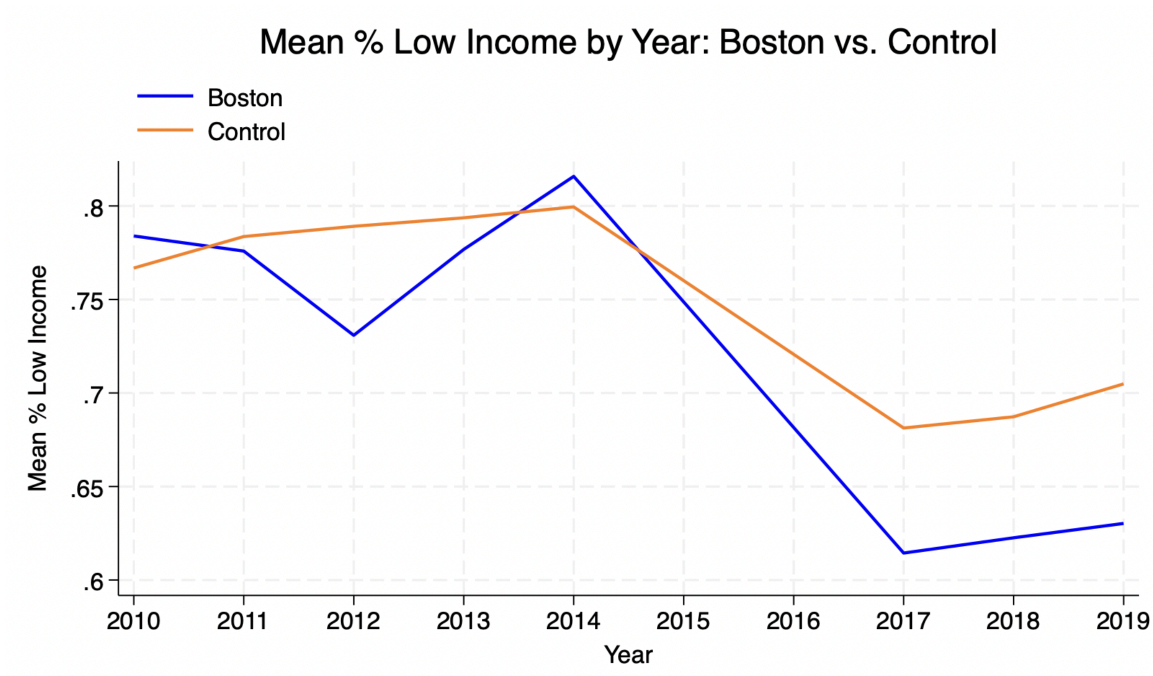

Figure 6: Trends in the percentage of low income students in Boston and control cities between 2010-2019