Boston Area Research Initiative
Boston Area Research Initiative • Northeastern & Harvard Universities
E-Mail: bari@northeastern.edu • Web: www.bostonarearesearchinitiative.net

The Opportunity Index: A Data-Driven Tool for Countering Inequities in Boston Public Schools

FINAL REPORT

Executive Summary

The Boston Area Research Initiative (BARI) and Boston Public Schools (BPS) have partnered to construct an Opportunity Index (OI)—*a place-based metric that captures inequities in academic achievement that arise from factors that are outside of the control of schools*—intended to inform funding and programming. This has entailed: 1) measuring dimensions of neighborhood context that theoretically could impact the performance of resident students for a Neighborhood Opportunity Index (NOI); 2) quantifying the impact these dimensions have on academic achievement on MCAS tests; 3) re-assessing the impact of individual-level characteristics that already receive programmatic attention (i.e., poverty, special education needs, English Language Learner) *plus* race and gender (which are not explicitly considered in programming) and thus broadening the NOI to the fuller OI; and 4) aligning these insights with our ability to predict academic performance for high school students using earlier experiences, like course failures, suspensions, and absenteeism.

The final product is three OIs—one each for elementary, middle, and high school students—that use individual- and neighborhood-level characteristics to quantify for each school the assets and deficits its students carry with them into the classroom; and an additional OI for high school students based on "risk" indicators from previous performance. This tool will help BPS to organize funding and programming to counteract inequities and assist schools to target the specific needs their students.

This document describes results and implications of the work, with detail on methods in Appendices. Here we briefly summarize the major findings.

Highlights

- Neighborhoods varied by as much as 20 pts. in their expected test scores when accounting for resident characteristics already considered by BPS programming (i.e., poverty, disability, English Language Learner status). This variation was driven largely by academic attainment of adults, public safety, and socioeconomic status.
- Individual-level characteristics considered in programming, including poverty, special education status, and English Language Learner status were also important factors in predicting academic outcomes, accounting for 25-35% of the variation.

- Race accounted for an additional 5% of variance, and 14% of the information in the full model, indicating that other individual and neighborhood factors *are not sufficient* to account for racial inequities in academic achievement.
- Unsurprisingly, previous course failures and suspensions were the strongest predictors of high school MCAS scores, but some individual characteristics remained informative.

Table of Contents

1. Overview

The Boston Area Research Initiative (BARI) and Boston Public Schools (BPS) partnered to construct an Opportunity Index (OI)—*a place-based metric that captures inequities in academic achievement that arise from factors that are outside of the control of schools*. The goal was to develop a data-driven tool for directing funding and programming to counteract the inequities that naturally arise in Boston because its residential neighborhoods are heavily segregated along racial and socioeconomic lines; it could also act as a model for any urban district facing similar challenges. Further, it was a pioneering attempt to develop policies that account for "neighborhood effects," or the impacts that residential context can have on individual residents.

The project unfolded in four stages, which form the organization of this report.

1) We identified and measured dimensions of neighborhood context that theoretically could impact the performance of students living there (the Neighborhood Opportunity Index, or NOI).
2) We "validated" these dimensions by quantifying their impact on academic achievement, measured through MCAS performance. This told us which dimensions mattered for academic outcomes and to what extent. In doing so, we discovered that neighborhood effects are real and persistent for BPS students across years and age groups, but that a substantial amount of variation is better understood in terms of the characteristics of individual students themselves
3) We identified individual-level characteristics that also contribute systematically to inequities and quantified their effects (expanding from an NOI to a more comprehensive OI). This focused on elements already considered by the district, like poverty (i.e., qualification for free-or-reduced-price lunch, FRPL) and English Language Learner (ELL) status, but we also recognized that current funding and programming models do not explicitly account for race. We included this in an additional set of models, finding that it further explained differences in academic achievement.
4) For high school students, we built a final set models incorporating previous outcomes thought of as "risk" indicators, like course failures and suspensions, that might be even stronger in *predicting* performance. This is based on a set of risk indicators already used by BPS.

Throughout, we found differences in the extent to which the neighborhood- and individual-level components of the OI predicted academic achievement for elementary, middle, and high school students. For this reason, the final product is not a single OI but three OIs, one each for elementary, middle, and high school students, each with its own characteristic set of individual- and neighborhood-level factors that are included in the final calculation of the strengths and vulnerabilities of a given school's population, weighted according to their association with MCAS scores at that level. We also produced a fourth OI for high school students that incorporated the "risk" indicators based on previous performance.

1.1 A Note on Methodology

In order to select and weight the proposed components of the OI we evaluated their relationship with MCAS scores using a statistical tool referred to as multilevel models (also known as hierarchical linear models). Multilevel models are ideal for this question in that they disentangle the effects of individual-level characteristics (e.g., an individual's race, poverty level) from the characteristics of a shared environment (e.g., concentrated poverty), in this case those of a residential neighborhood. Put colloquially, a major early goal of research with multilevel models was to answer the question, "Do poor neighborhoods have lower outcomes solely because the individuals who live there have less individual access to resources, or because the neighborhood has an overall dearth of resources that impacts everyone who lives there?"

We ran a series of multilevel models that used both individual- and neighborhood-level factors to predict MCAS Math and MCAS ELA scores for elementary, middle, and high school students across three different school years (2011-2012, 2012-2013, and 2013-2014). This amounted to 18 separate models (2 tests x 3 age groups x 3 school years). We synthesize the results of these models in the text and report them in full in Appendix B, where we also provide a more detailed description of how to interpret the results and their implications.

There are two main ways to translate regression results into implications. The first is to state "the percentage of variance explained," also known as R^2. The other is to transform the parameters into the units of the outcome variable—in this case, the predicted difference in MCAS scores.[1] Essentially, these are two different ways of expressing the same statistical relationship between a predictor and an outcome. The first describes the predictive strength of a model, but the latter

[1] This depends on the arithmetic translation of a regression coefficient into the units of the outcome variable. The parameters we report throughout are standardized, meaning they reflect the number of standard deviations that the outcome would be expected to change if the predictor increased by one standard deviation. We then multiply this by the standard deviation of the outcome variable for our estimate of the tangible implications. For example, the standard deviation of the MCAS Math test for elementary school students for the 2013-2014 academic year was 19.16 points. We found that an increase in the public safety of a neighborhood by one standard deviation is equal to an expected increase of .05 standard deviations in the MCAS Math test score in middle school students, amounting to .06*19.16 pts. = 0.96 pts. per standard deviations. MCAS tests have a standard deviation of 15-19 pts, and most variables have an approximate range of 4.5-5 standard deviations. This would translate to a difference of 5.74 pts. or somewhat higher (based on an actual 6 standard deviation range for public safety) between the neighborhoods with the highest and lowest levels of crime. *Importantly, it does not account for the effects of* all *proposed OI constructs.*

quantifies the tangible implications, for which reason we more often use the latter in the discussion that follows.

2. Neighborhood Context and Academic Achievement

To create a *place-based metric* that captures inequities among students at BPS, we first had to identify the aspects of neighborhood context that could potentially influence academic achievement, which we refer to as *constructs* (e.g., public safety). We then measured them using *indicators* that reflect those constructs (e.g., incidences of gun-related violence), which we gathered from various data sources, including the American Community Survey, the Boston Public Health Commission, and custom measures developed by BARI from the City of Boston's 311 and 911 call records.

We subjected the constructs and their constituent indicators to an increasingly stringent series of statistical tests evaluating: statistical consistency from year-to-year (i.e., reliability), insuring the stability of the OI; and whether combined indicators in fact are correlated in the manner necessary to justify their treatment as an overarching construct (i.e., convergent validity). This culminated in a tailored list of constructs and indicators, including: 1) socioeconomic status (e.g., median household income), 2) public safety (e.g., public violence), 3) public health (e.g., premature mortality), 4) physical disorder (e.g., graffiti), 5) academic attainment of adults (e.g., % with bachelor's degree), 6) "newcomer" (e.g., % foreign-born), 7) residential stability (e.g., median years in home), and 8) custodianship (e.g., use of 311).

Once we had a full set of constructs describing the neighborhoods of Boston, we quantified the extent to which each was predictive of the academic achievement of the students living in a given neighborhood (using multilevel models; see Section 1.1). This section further describes each of the constructs and the ways in which they might theoretically influence student achievement. They are presented in order of the strength with which they were found to predict MCAS scores in BPS students. We then enumerate the full set of neighborhood constructs to be included in the NOI and their relative weighting (reported in Table 1). For more detail, Appendix A reports the full list of indicators that we considered, the process for inclusion/exclusion, and the final mathematical steps for calculating constructs from indicators.

2.1. Neighborhood Constructs

Initial models found that neighborhoods differed by as much as 20 pts. on the expected MCAS scores of resident students, holding individual-level factors constant (i.e., poverty, special needs, English Language Learner status). This is a sizeable inequity based solely on neighborhood of residence. Here we describe the eight neighborhood constructs we considered as part of the OI, and the extent to which they predicted these differences in academic achievement. The first five were predictive of outcomes for at least one age group (elementary, middle, or high school)—academic attainment of adults, public safety, socioeconomic status, physical disorder, and custodianship; the latter three were not—"newcomer," residential stability, and public health.

These neighborhood characteristics fit into larger models and theories that outline how neighborhood factors shape achievement.

- *Social and cultural isolation theories* suggest that living in a disadvantaged neighborhood isolates children from influential adults who value education and model educational achievement and discourage risky behaviors (Jencks and Mayer 1990; Wilson 1987).
- *Social disorganization theory* suggest that greater disadvantage and residential turnover in a neighborhood prevents the development of social ties, making it difficult for the community to establish and pursue shared goals and expectations, including the prevention of crime and socialization of youth (Shaw and McKay 1942/1969; Sharkey 2010; Sharkey et al. 2012; Morenoff and Sampson 1997).
- *Environmental theories* focus on the disparate exposure to hazards to both physical and mental health, including environmental toxins (Crowder and Downey 2010), disorder (i.e., broken windows) and crime (Massey 2004), all of which is associated with diminished achievement (Sharkey and Faber 2014); and differential access to amenities like health care, high quality childcare, and healthy foods, which are associated with academic readiness and achievement especially in the earlier grades (Wodtke and Parbst 2017).

We were able to capture and assess many of the variables implicated by theories on neighborhood context using data regularly collected by BPS, BARI, and the Boston Public Health Commission (BPHC). Notably, the different theories listed here sometimes overlap in the variables they would predict to impact student achievement, so rather than organize our constructs according to these theories, we list them in a more general way.

<u>Academic Attainment of Adults</u> has been found to significantly predict school dropout and achievement. This is because neighborhood adults apart from one's

parents can model the benefits of achievement for local youth and help them to navigate school contexts in ways that explicitly and implicitly mentor youth (Rendon 2014). *This was the most predictive construct in the NOI. It predicted higher test scores for both test subjects in all three age groups, with a particularly strong effect on Math tests and high school students in general.*

Public Safety. Exposure to high crime in a neighborhood disrupts engagement in school and ultimately academic performance among the students living there (Sharkey 2010; Sharkey et al. 2012; Bowen and Bowen 1999). Research on this subject has largely found that such effects are mediated by elevated levels of disruptive behavior and depressive symptoms in young people who have been exposed to community violence (Hardaway, Larkby, and Cornelius 2014; Borofsky et al. 2013). *Measured through rates of 911 calls for various types of crime and social disorder, was notably consistent in its predictive strength, having a similar impact across test subjects and age groups.*

Socioeconomic Status of a neighborhood, separate from family-level SES, has regularly been shown to predict academic achievement (Leventhal and Brooks-Gunn 2000; Burdick-Will et al. 2011) and dropout rates (Card and Rothstein 2006; Harding 2003). This is largely due to differences in access to social capital (Noguera 2003), including a lack of adult models for achievement and mentorship (Jencks and Mayer 1990; Galster et al. 2016). *Neighborhood socioeconomic status was the third strongest component of the NOI, consistently predicting test scores across subjects and age groups.*

Physical Disorder, or denigration and deterioration of public spaces and infrastructure (also referred to colloquially as "broken windows"), is closely related to public safety in that it reflects a lack of care and respect for the neighborhood (Wilson and Kelling 1982). Numerous studies have found that people use it as an easily accessible cue to assess neighborhood safety, and it is thus associated with fear of crime in residents (O'Brien and Wilson 2011; Gau and Pratt 2008; Skogan 1992). In theory, this can impact student social and emotional functioning and in turn undermine academic performance. *We found that physical disorder, measured through reports received by Boston's 311 system, indeed predicted lower academic performance in elementary school students.*

Custodianship is related to physical disorder in that it captures the extent to which residents proactively engage in the upkeep and improvement of their neighborhood. This concept is rather new, proposed by O'Brien (co-director of BARI and co-author of this report; O'Brien 2015, 2016, 2016; O'Brien, Gordon, and Baldwin-Philippi 2014; O'Brien et al. 2016) through work with the City of Boston's 311 system, and thus has not previously been applied directly to student academic achievement. It is possible, however, that neighborhoods whose adult residents more often act as "custodians" provide a more nurturing environment, either through their beneficent treatment of local spaces or through related acts of caretaking for their neighbors and institutions. *We found that elementary school students in neighborhoods with higher levels of custodianship had higher MCAS scores.*

Proportion of "Newcomers" (i.e., immigrants) in the neighborhood has a complex association with academic achievement (Hibel and Hall 2014). On the one hand, the density of immigrants in a community correlates positively with poverty. Thus, immigrants tend to live in neighborhoods with fewer resources and adults with a bachelor's degree. On the other hand, immigrant communities often have stronger the social capital thanks to a common set of cultural values, which can positively related to achievement (Kao and Tienda 1998; Oropesa and Gorman 2000). Last, especially when immigrant communities speak a language other than English, students may have less opportunity to practice ELA-related skills when they leave school. These various effects may counter each other, and may also be explained by other correlates (e.g., socioeconomic status, individual-level English Language Learner status). *We found that "newcomer" was not independently predictive of performance at any level.*

Residential Stability is often considered in models of neighborhood effects on achievement because of its relation to social capital and school transitions. For the former, consistent turnover in communities can undermine social ties and thereby limit collective socialization of children (Morenoff and Sampson 1997; Shaw and McKay 1942/1969), as well as shared learning between students and their neighbors. However, districts with comprehensive choice policies (like BPS) may eliminate this effect because school assignment has been disconnected from neighborhood residence, meaning that residential movement is not correlated with changing schools. *We found that residential stability was not independently predictive of performance at any level.*

Public Health, which we assessed with rates of 911 calls for medical emergencies and BPHC estimates of premature death, captures both access to resources that promote health—including healthcare facilities, and healthy food—and diminished exposure to environmental toxins. *We found that these factors were not independently predictive of performance at any level.*

Although we have examined these as separate constructs, it is important to keep in mind that many are correlated within neighborhoods and that the most advantaged communities in Boston tend to lack many of these assets. Neighborhoods with higher levels of SES also tend to have higher levels of public safety, custodianship, order, and academic achievement in adults. Students residing in these neighborhoods have access to better healthcare and other resources. That is to say that the construct-by-construct results often combine to create magnified differences between advantaged and disadvantaged neighborhoods.

2.2. Summary

There were five neighborhood constructs that were predictive of MCAS scores for at least one of the three age groups, in order of consistency and strength: 1) academic attainment of adults (e.g., % with bachelor's degree), 2) public safety (e.g., gun-related violence), 3) socioeconomic status (e.g., median household income), 4) physical disorder (e.g., graffiti), and 5) custodianship (e.g., use of 311). The weighting of these (or their omission) in the NOI for elementary, middle, and high school is reported in Table 1. Note that the effect of neighborhood was most salient for elementary school students, as reflected by an NOI with more neighborhood factors and stronger weightings than for the other age groups.

A difference of up to 20 pts. in MCAS scores across neighborhoods is notable, but it is far from the entirety of variance in BPS. Indeed, this tangibly important finding only accounted for 6-10% of variation in test scores. This is not entirely surprising as individuals are clearly subject to many other effects, from their own capacities and inclinations, to family context, to relationships with teachers, etc., that are also responsible for their academic outcomes. Thus, we decided it was necessary to further consider the individual-level factors of import.

As we think about the distribution of resources to schools based on the place-based needs of their student populations, it begs the question of how those dollars are best spent. Research based on over 500 urban school districts suggests that neighborhood effects matter because they decrease the socioemotional learning and safety in schools and, to a lesser degree, their academic rigor (McCoy, Roy, & Sirkman, 2013). Similarly, randomized trials that examine the impact of students moving from low SES to high SES neighborhoods is largely through increased school safety and time spent on schoolwork (Leventhal & Brooks-Gunn, 2004). To compensate for neighborhood or out-of-school factors, monies might be directed toward better mental health services and increasing students' feelings of safety and security at school.

Table 1. Weighting of Neighborhood Constructs in the Opportunity Index for Elementary, Middle, and High School Students.

	Elementary	Middle	High
Acad. Attainment	.095	.106	.139
Custodianship	.075	--	--
Public Safety	.058	.073	.081
Physical Disorder	-.094	--	--
SES	.059	.056	.045

3. Individual-Level Characteristics and Academic Achievement

As described, the analysis evaluating the relationship between neighborhood constructs and academic achievement took account of individual-level factors as well—including special education (SPED) status, free-or-reduced-price lunch (FRPL) status, and English Language Learner (ELL) status (Current or Former, in comparison to Never). All of these features are already accounted for in BPS models of funding and programming. However, one of the most important predictors in all research on inequities—whether academic achievement or otherwise—is not yet considered by such models: race. It is well-known that there are racial inequities in academic achievement, but that there are political challenges to dealing with it directly. The implicit hope is that other variables, like poverty and neighborhood context, will account for all of the differences across race, thereby eliminating the need to address it directly. In this same spirit, we also incorporate gender, which is a standard control variable known to be associated with certain test outcomes.

We run the models twice, once without race and gender, and once with them added (with race broken out into White, Asian, Hispanic, Native American, and Multi-/Other, with Black as the reference group, as it is the most common in BPS), to evaluate the assumption that differences between groups, especially races, can be fully accounted for by other factors. We present both and compare the results.

3.1. Individual-Level Factors that Matter

Nearly all individual-level factors included in the initial models (i.e., without race) strongly predicted academic achievement, even when accounting for neighborhood context. As expected, these effects were considerably stronger than the effect of neighborhood context, accounting for 25-35% of the variance across models.

The strongest and most consistent predictor was SPED status, which equated to ~10-13 pts. on both MCAS Math and ELA across all three age groups. Current ELL students also tended to score lower than their peers, especially on ELA tests. This difference became especially exaggerated in high school students at ~17 pts. Interestingly, those classified as Former ELL tended to perform *better* than those never classified as ELL. Those with FRPL status tended to score ~5 pts. lower on tests than their peers across age groups.

Table 2. Weighting of Individual-Level Characteristics in the Opportunity Index for Elementary, Middle, and High School Students.

	Elementary	Middle	High
SPED	-0.72	-0.75	-0.85
FRP Lunch	-0.39	-0.33	-0.24
LEP Status (Ref. Never)			
Current ELL	-0.35	-0.61	-0.88
Former ELL	0.44	0.31	0.21

3.2. Adding Race (and Gender) to the Models

When we added race to the models, we found that it strongly predicted student performance, even when taking into account other individual- and neighborhood-level characteristics associated with disadvantage. Asian students and White students heavily out-performed Black and Latino students, and Latino students, as well, scored higher than Black students (once ELL was accounted for). This is a striking finding for two reasons. First, this means that there are persistent racial inequities *apart from* those produced and reinforced by residential context, poverty, and differential diagnosis with learning disabilities. Second, this is not something that BPS already accounts for in its funding and programming, whereas the other variables are included in such considerations. Sex also had a measureable effect on performance, albeit far smaller; females performed better on ELA tests and males performed better on Math tests in elementary school only.

Comparing the with- and without-race models, it is clear that in the without-race models some, but not all, of the variance associated with race was re-attributed to other variables correlated with race. Most notably, the amount of variance associated with neighborhoods decreased from 6-10% to 1.5-2%. That is to say much of the neighborhood effects described in the previous section were actually an artifact of the uneven distribution of racial groups across neighborhoods. The effect of academic attainment of adults in particular was diminished by 2/3rds. It is important to note that this *does not* mean that the academic attainment of parents or neighbors explains racial differences; it indicates instead that academic attainment of the adults in a neighborhood was most strongly correlated with the distribution of racial groups across the city and in turn acted as a proxy for it. Similar though lesser effects were seen for public safety and socioeconomic status of the neighborhood.

Race substantially strengthened the predictive power of the OI, allowing us to explain ~5% more variance. Put another way, given that the models on the whole explain ~40% of the total variance, race accounted for 1/8th of our overall ability to predict test scores, and about 2.5 times as much information as explained uniquely by neighborhood context. This suggests that an OI with race would be more effective in countering inequities than one that omits it.

Table 3. Weighting of Neighborhood Constructs in the Opportunity Index when Race is Included in the Model, for Elementary, Middle, and High School Students.

	Elementary	Middle	High
Acad. Attainment	.031	.041	.027
Custodianship	.063	--	--
Public Safety	.044	.044	.036
Physical Disorder	-.069	--	--
SES	.034	.039	.039

3.3. Summary

We found, indeed, that a comprehensive OI considering both individual- and neighborhood-level characteristics was the most effective tool to quantify and counteract inequities across the district. At the individual level, SPED status, FRPL status, and ELL status were all important factors, accounting for ~30% of variance in test scores across subjects and age groups. That said, we also found that, even when considering all of the other individual- and neighborhood-level factors, one of the strongest single predictors was race. This is something not yet considered in BPS' current models of funding and programming, likely because of the political and legal challenges such a policy would pose. We therefore have in the accompanying documents provided guidelines for how to calculate the OI without race, based on the same model. We recommend this strategy because if accounting for race directly is not an option, it would be important to fully account for all of the individual- and neighborhood-level characteristics that are associated with race that are also predictive of test outcomes.

Table 4. Weighting of Individual-Level Characteristics in the Opportunity Index when Race is Included in the Model, for Elementary, Middle, and High School Students.

	Elementary	Middle	High
Female	.05	.17	.07
Race (Ref. Black)			
White	.50	.42	.44
Asian	.69	.68	.71
Latino	.11	.11	.06
Nat. Amer.	.08	.12	.17
Multi-/Other	.33	.24	.25
SPED	-.70	-.73	-.81
FRP Lunch	-.31	-.26	-.19
LEP Status (Ref. Never)			
Current ELL	-.38	-.64	-.90
Former ELL	.36	.23	.16

4. Predicting High School Performance with Leading Indicators

As a final stage of the development of the OI, we turned to a set of "risk indicators" that BPS already uses to anticipate student academic performance. These indicators include: the proportion of Math or ELA courses failed; failing a Math or ELA MCAS test (yes/no); chronic absenteeism (<90% attendance; yes/no); and one or more suspensions (yes/no).[2] All are measured as dichotomous (i.e., yes/no) variables. They are all available for middle school students; only absenteeism and suspensions are available for elementary school students. By using such indicators in 8th grade to predict 10th grade performance, or in 5th grade to predict 6th grade performance, we are still identifying factors that are outside the control of the high

---

[2] To confirm that these were the best way to measure each of these variables, we examined: proportion of failed courses and MCAS tests as well as whether one had failed any such course or test; chronic absenteeism at 95%, 90%, 85%, and 80%; and number of suspension incidents, suspension days, and whether one had ever been suspended. The versions of the variables that we use were the most predictive of future MCAS performance.

schools that these students attend, and yet are likely predicting performance more precisely than individual demographics and neighborhood characteristics would on their own. Because of the need for multiple years of previous data, we only conduct the risk indicator analysis for students who took MCAS tests in 2013-2014.

4.1. Predictive Power of Risk Indicators

An initial set of analyses found that all 8th grade risk indicators significantly predicted test scores in 10th graders *except* that failing an MCAS ELA test did not predict future MCAS Math performance. When incorporated into our analysis, the risk indicators became some of the strongest predictors in the models. Performance on previous MCAS tests was by far the strongest predictor, which stands to reason as the tests build on previous material over time. Failing a higher proportion of earlier courses was the next most important predictor, followed by suspensions and absenteeism. This differs somewhat with the current weighting of risk indicators utilized by BPS as it suggests a much wider range of weights (the current system uses 1, 1.8, and 1.9 as the primary weights; our results would suggest instead a range of 1 to 6 across predictors).[3] Similarly, 6th graders who were chronically absent or suspended at least once during middle school had lower performance on MCAS.

The risk indicators accounted largely, but not completely, for the other factors in the OI. SPED, ELL, and FRPL remained important predictors, albeit to a lesser extent. Even neighborhood remained somewhat relevant, particularly in terms of the academic attainment of adults in the neighborhood; public safety and socioeconomic status of the neighborhood also were predictive for middle school students.

These findings demonstrate that leading risk indicators are effective tools for forecasting academic outcomes, but they do not tell the whole story. Individual- and neighborhood-level characteristics provide additional information that help us to more precisely predict inequities in MCAS scores. Also, as we noted above, the goal of the OI project is not only to *predict* academic performance, but to offer guidance in *counteracting* inequities in that performance. The risk indicators are effective for the former, but do not necessarily tell us why an individual will perform better or worse. For this reason, we have provided mechanisms for calculating a risk indicator-based OI for high schools that can be used for funding

[3] We also assessed how this new weighting compared with the old system of "risk points." When the models were compared, risk points accounted for 27% of the variance not explained by neighborhood or other individual-level factors, whereas the separate entry of the risk indicators accounted for 54%–exactly double the accuracy.

allocation, as well as the neighborhood- and individual-level characteristics-based OI presented in Sections 2 and 3, which can be used to communicate to schools the strengths and vulnerabilities of their student population. Examples of how to visually communicate such information is provided in Appendix C. With this information school leaders can utilize their resources most effectively to counteract inequities that are manifest in their school.

4.2. Caveats to the Use of Risk Indicators in the OI

Risk indicators are clearly an effective way of forecasting academic achievement for individual students, offering information not available through demographics and home neighborhood alone. That said, there are a number of weaknesses of this approach, some interpretive and others methodological. Most important is the extent to which risk indicators are "outside the control of the school." This interpretation is true for transitions between schools structured according to the traditional divisions between elementary (K-5th grade), middle (6th grade-8th grade) and high school (9th grade-12th grade), but many schools in BPS do not fit this criteria. In turn, previous risk indicators are in part a function of the school itself. To probe this further, we found that all risk indicators varied more by school than by neighborhood, in some cases finding that up to 30% of variation was clustered at the school level. Even if we allow that 5th and 8th grades are important transition points at which we should evaluate how the next school will best serve its population, this ignores students in all other middle and high school grades. If we calculate the OI for a school based on risk indicators, we will inevitably include events occurring in that school. If funding is then linked to the OI, we create perverse incentives for principals, especially in terms of suspension rates (i.e., schools could choose to use suspensions more often, raising the risk indicators of their students).

Methodologically, there were some limits to our analysis and its application. First, a substantial number of students (~25% in the case of high school students) were entering the district for the first time, and thus had no information regarding risk indicators. There may be differences between them and current BPS students that are not accounted for in this analysis, but more concerning, we are not able to assess their needs and in turn the overall needs of their schools fairly through and OI based on risk indicators. Second, for high school students, we were unable to incorporate into the analysis individuals who were left back in 9th grade because they had not yet reached the point of completing the 10th grade MCAS tests (~15% of students). This might have had some impact on our interpretations, but should not be an issue for the calculation of the OI as these individuals still have risk indicators associated with them and would be included.

5. Conclusion

Here we have presented a four-step process for calculating an OI for BPS, and one that acts as a model for other urban districts trying to counteract inequities that arise naturally from factors outside of the control of schools. We have concluded that both neighborhoods and individual-level characteristics matter. The former reflects a novel opportunity for BPS to implement funding and programming policies that directly address “neighborhood effects.” The latter has presented another distinctive opportunity in that it revealed racial differences in academic outcomes *even when considering individual- and neighborhood-level hardships*. This is more challenging to deal with directly, but it does highlight the fact that correlates alone are not sufficient to account for racial inequities. We have also shown that for high school students the best way to predict performance is a mixture of lead risk indicators and individual-level characteristics. The Appendices and supplementary materials associated with this report provide: the process for developing the NOI constructs (Appendix A); the methodological details and full analytic results for selecting and weighting the components of the OI (Appendix B); forms that might be used to communicate how each school scores on the OI and its individual components (Appendix C); and programming syntax for calculating the OI based on current data (with and without leading risk indicators for high school students; separate set of documents, see cover page). We also provide.

What this report does not do, though we hint at it throughout, is answer the question of how resources would be best utilized to counteract inequities across the district. Because the factors considered are so numerous, it would seem that each school should target the particular strengths and vulnerabilities of their student population. For neighborhood context, the overwhelming evidence points to impacts on student socioemotional learning, suggesting that schools with a low NOI would benefit from investment in mental health services and mechanisms that will bolster student feelings of safety and security at school. In contrast, at the individual level SPED and ELL already have dedicated programs. As in other places throughout this report, race remains the unanswered question—even if a district chooses to target race directly, what are the experiences and contexts responsible for these differences? How does a district address them effectively? Here we have recommended that the, given political and legal challenges, the best approach will be to use the parameters from models that do not account for race, permitting the district to maximally address those negative contexts that are associated with academic performance, but we trust that such debate will continue. This strikes us as a deeply important set of questions that arise from this work, the answers to which will ensure that the information contained in the OI has meaningful benefits for public school students in Boston.

6. References

Borofsky, Larissa A., Ilana Kellerman, Brian Baucom, Pamella H. Oliver, and Gayla Margolin. 2013. Community violence exposure and adolescents' school engagement and academic achievement over time. *Psychology of Violence* 3 (4):381-395.

Bowen, Natasha K., and Gary L. Bowen. 1999. Effects of crime and violence in neighborhoods and schools on the school behavior and performance of adolescents. *Journal of Adolescent Research* 14 (3):319-342.

Burdick-Will, Julia, Jens Ludwig, Stephen W. Raudenbush, Robert J. Sampson, Lisa Sanbonmatsu, and Patrick Sharkey. 2011. Converging evidence for neighborhood effects on children's test scores: An experimental, quasi-experimental, and observational comparison. In *Whither Opportunity*, edited by G. J. Duncan and R. J. Murnane. New York, NY: Russell Sage Foundation.

Card, David, and Jesse Rothstein. 2006. Racial segregation and the black-white test score gap. *Journal of Public Economics* 91 (11-12):2158-2184.

Crowder, Kyle, and Liam Downey. 2010. Inter-neighborhood migration, race, and environmental hazards: Modeling micro-level processes of environmental inequality. *American Journal of Sociology* 115 (4):1110-1149.

Galster, George, Anna Santiago, Lisa Stack, and Jackie Cutsinger. 2016. Neighborhood effects on secondary school performance of Latino and Afraican American youth: Evidence from a natural experiment in Denver. *Journal of Urban Economics* 93:30-48.

Gau, Jacinta M., and Travis C. Pratt. 2008. Broken windows or window dressing? Citizens' (in)ability to tell the difference between disorder and crime. *Criminology and Public Policy* 7 (2):163-194.

Hardaway, Cecily R., Cynthia A. Larkby, and Marie D. Cornelius. 2014. Socioemotional adjustment as a mediator of the association between exposure to community violence and academic performance in low-income adolescents. *Psychology of Violence* 4 (3):281-293.

Harding, David J. 2003. Counterfactual models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology* 109 (3):676-719.

Hibel, Jacob, and Matthew Hall. 2014. Neighborhood coethnic immigrant concentrations and Mexican American children's early academic trajectories. *Population Research and Policy Review* 33 (3):365-391.

Jencks, C., and S. Mayer. 1990. The social consequences of growing up in a poor neighborhood. In *Inner-City Poverty in the United States*, edited by J. Lynn, L. Washington, DC: National Academy Press.

Kao, Grace, and Marta Tienda. 1998. Educational aspirations of minority youth. *American Journal of Education* 106 (3):349-384.

Leventhal, Tama, and Jeanne Brooks-Gunn. 2000. The neighborhoods they live in: The effect of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin* 126 (2):309-337.

Massey, Douglas S. 2004. Segregation and stratification: A biosocial perspective. *Du Bois Review* 1:7-25.

Morenoff, Jeffrey, and Robert J. Sampson. 1997. Violent Crime and the Spatial Dynamics of Neighborhood Transition: Chicago, 1970-1990. *Social Forces* 76:31-64.

Noguera, Pedro A. 2003. Social capital and the education of immigrant students: Categories and generalizations. *Sociology of Education* 77 (2):180-183.

O'Brien, Daniel Tumminelli. 2015. Custodians and custodianship in urban neighborhoods: A methodology using reports of public issues received by a city's 311 hotline. *Environment and Behavior* 47:304-327.

———. 2016. 311 hotlines, territoriality, and the collaborative maintenance of the urban commons: Examining the intersection of a coproduction policy and evolved human behavior. *Evolutionary Behavioral Sciences* 10 (2):123-141.

———. 2016. Using small data to interpret big data: 311 reports as individual contributions to informal social control in urban neighborhoods. *Social Science Research* 59:83-96.

O'Brien, Daniel Tumminelli, Eric Gordon, and Jesse Baldwin-Philippi. 2014. Caring about the community, counteracting disorder: 311 reports of public issues as expressions of territoriality. *Journal of Environmental Psychology* 40:320-330.

O'Brien, Daniel Tumminelli, Dietmar Offenhuber, Jesse Baldwin-Philippi, Melissa Sands, and Eric Gordon. 2016. Uncharted territoriality in coproduction: The motivations for 311 reporting. *Journal of Public Adminstration Research and Theory* In press.

O'Brien, Daniel Tumminelli, and David S. Wilson. 2011. Community Perception: The Ability to Assess the Safety of Unfamiliar Neighborhoods and Respond Adaptively. *Journal of Personality and Social Psychology* 100 (4):606-620.

Oropesa, R.S., and Bridget K. Gorman. 2000. Ethnicity, immigration, and beliefs about marriage as a "tie that binds". In *The Ties that Bing: Perspectives on Marriage and Cohabitation*, edited by L. J. Waite and C. Bachrach. Piscataway, NJ: Transaction Publishers.

Rendon, Maria G. 2014. Drop out and "disconnected" young adults: examining the impact of neighborhood and school contexts. *The Urban Review* 46 (2):169-196.

Sharkey, Patrick. 2010. The acute effect of local homicides on children's cognitive performance. *Proceedings of the National Academy of Sciences* 107 (26):11733-11738.

Sharkey, Patrick, and Jacob W. Faber. 2014. Where, when, why, and for whom do residential contexts matter? Moving away from the dichotomous understanding of neighborhood effects. *Annual Review of Sociology* 40:559-579.

Sharkey, Patrick T., Nicole Tirado-Strayer, Andrew V. Papachristos, and C. Cybele Raver. 2012. The effect of local violence on children's attention and impulse control. *American Journal of Public Health* 102 (12):2287-2293.

Shaw, Clifford, and Henry McKay. 1942/1969. *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.

Skogan, Wesley G. 1992. *Disorder and Decline*. Berkeley, CA: University of California Press.

Wilson, James Q., and George Kelling. 1982. Broken Windows: The Police and Neighborhood Safety. *Atlantic* 127:29-38.

Wilson, William J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

Wodtke, Geoffrey T., and Matthew Parbst. 2017. Neighborhoods, schools, and academic achievement: A formal mediation analysis of contextual effects on reading and mathematics abilities. *Demography* 54 (5):1653-1676.

Appendix A. Construction of the Neighborhood Opportunity Index

A1. Developing a List of Neighborhood Opportunity Index Indicators

The process of developing the neighborhood portion of the Opportunity Index (OI) occurred in five steps:

1) *Proposing constructs* that describe the overarching characteristics of neighborhoods that might be associated with academic achievement, for example, the safety of a neighborhood may influence how much time or when a student is allowed out-of-home for enrichment activities;

2) *Identifying indicators* that reflect each of those constructs, for example, the rate at which residents report violence through 911 might reflect neighborhood safety;

3) *Gathering indicator data* from various sources, and excluding those for which data could not be accessed, for example some proposed indicators (e.g., number of Facebook friends) could not be accessed;

4) *Establishing trackability* for an annually-updating OI, based on the anticipated consistent accessibility and nature of the source, for example many of the proposed health indicators were not guaranteed for future use;

5) *Evaluating statistical properties* of indicators for two types of validity: face validity and construct validity. Face validity regards whether an indicator measures what it is hoped to. For our purposes, this included questions of year-on-year volatility (i.e., reliability) and screening for data quality and completeness that would insure the likelihood that a measure's content is consistent with its definition. Construct validity determines if and how a set of indicators can be statistically combined to constitute a single score for an intended construct. An example includes testing whether all indicators of Public Health, drawn from two different sources, were sufficiently correlated to state that they were describing the same thing. This was done through confirmatory factor analyses that analyzed the covariance structures across indicators. We excluded any indicator that did not correlate sufficiently with the other indicators for that construct.

Steps 1-2 generated an extensive list of potential neighborhood indicators, and Steps 3-5 trimmed this list to those indicators that could feasibly be included in an annually-updating OI. Table A1 contains the complete list of neighborhood constructs and indicators we considered for the OI. Each column reflects a step in the process, with a checkmark meaning that an indicator was maintained through that stage, and a negative sign showing the point at which it was removed from consideration. For indicators that failed face validity we describe the nature of the issue. In a few select cases we included an indicator with unclear future accessibility (noted with a '?') in latter tests. The sources, definitions of constructs that were tested against student MCAS scores and their constituent indicators are described in greater detail in Section A2.

Table A1. Complete List of Candidate Indicators, Organized by Construct.

	Practical Considerations			Validity Criterion	
***Initial Construct* / Candidate Indicator**	**Possible Source**	**Future Accessibility**	**Alignment to Census**	**Face**	**Construct**
Neighborhood Amenities					
Neighborhood Assets (e.g. community center)	-	-	-	-	-
Physical Disorder (311 reporting)	BARI[4]	✓	✓	✓	✓
Ease of access to public transportation	-	-	-	-	-
Ambulance rates	City[5]	-	-	-	-
Density of public housing	City	-	-	-	-
Youth-serving institutions	-	-	-	-	-
Street Score	MIT Media Lab	?	✓	✓	-
Neighborhood Public Health					
% of babies born preterm	BPHC[6]	?	✓	✓	-
% of active Boston TB cases	-	-	-	-	-

[4] BARI: Boston Area Research Initiative.
[5] City: City of Boston Dashboard.
[6] BPHC: Boston Public Health Commission.

Access to Healthy Foods	-	-	-	-	-
Substance abuse	-	-	-	-	-
Body Mass Index	BPS	✓	✓	Data collection unreliable	-
Late-Life Medical Emergencies	BARI	✓	✓	✓	✓
Asthma	BPS	✓	✓	Data collection unreliable	-
Lead poisoning	BPHC	-	-	-	-
Infant mortality	BPHC	-	-	-	-
Premature Mortality	BPHC	-	-	-	-
# of health centers / # of well-child visits	BPHC	-	-	-	-
Youth Health Emergencies	BARI	✓	✓	✓	✓
Neighborhood Public Safety					
# of Residential Neighborhood Watch Groups	City	-	-	-	-
Walking and bicycle beat patrols	-	-	-	-	-
Private Conflict	BARI	✓	✓	✓	✓
Homicides	BARI	✓	✓	Unstable annually	-
Public Violence	BARI	✓	✓	✓	✓

Public Social Disorder	BARI	✓	✓	✓	-
Prevalence of Guns	BARI	✓	✓	✓	✓
Residential Stability of Neighborhood					
Median years in neighborhood	ACS[7]	✓	✓	✓	✓
Gentrification / Change in Income	BARI	✓	✓	Difficult to measure annually	-
% Homeownership	ACS	✓	✓	✓	✓
Resources Available for School Support/Preparedness					
Build BPS Data	-	-	-	-	-
Crime in the area surrounding the school	BARI	-	-	-	-
First Language Spoken at Home	BPS	-	-	-	-
Mobility rate for school	BPS	-	-	-	-
Mobility – Stability	BPS	-	-	-	-
Mobility – Intake	BPS	-	-	-	-
Rate of single parent households	ACS	✓	✓	-	-
Post-secondary education attainment	ACS	✓	✓	✓	✓

[7] 5-year estimates from the U.S. Census Bureau's American Community Survey

% Immigrant	ACS	✓	✓	✓	✓
Neighborhood Social Capital					
Credit Score	Federal Reserve	-	-	-	-
Collective Efficacy	-	-	-	-	-
Availability of formal and informal mentors	-	-	-	-	-
Per-capita parental involvement in schools	-	-	-	-	-
Extracurricular involvement	-	-	-	-	-
Undermatch to college	-	-	-	-	-
# of Facebook Friends.	-	-	-	-	-
Neighborhood Socioeconomic Status (SES)					
Delinquent real estate notices.	City	-	-	-	-
Median Household Income	ACS	✓	✓	✓	✓
Youth Employment rate	ACS	-	-	-	-
Unemployment Rate of 16+	ACS	✓	✓	✓	✓
Family Poverty Rate	ACS	✓	✓	✓	✓
SNAP Rate	ACS	✓	✓	✓	✓

A2. Data Sources

This section lists the data sources leveraged and the constructs to which they contribute and the next (Section A3) defines the indicators for each construct. Note that in Section A3 we include which data source each indicator is drawn from in order to facilitate cross-referencing with Section A2.

American Community Survey

The U.S. Census Bureau's American Community Survey (ACS) tracks socioeconomic, demographic, and housing characteristics of census tracts. A set of these indicators is curated by the Boston Area Research Initiative for the Commonwealth of Massachusetts. The ACS indicators are based on five-year estimates that renew yearly, beginning with 2005-2009 and finishing in 2011-2015. ACS's data contributed to the formation of the following constructs:

- Academic Attainment of Neighborhood
- Newcomer
- Residential Stability
- Socioeconomic Status

Boston Area Research Initiative

The Boston Area Research Initiative (BARI) has built and maintains the Boston Data Portal,[8] a key part of BARI's efforts to collect and disseminate information that fosters research-policy collaborations. Through the Data Portal, BARI constructs and curates *ecometrics* that describe the conditions or characteristics of a neighborhood at a given time, based on various digital data sources. Ecometrics used in this project are mainly derived from 911 call records (CAD) and requests for non-emergency government services received by 311. BARI's databases, contributed to the formation of the following constructs:

- Custodianship
- Physical Disorder
- Neighborhood Safety
- Neighborhood Health
- Socioeconomic Status

Boston Public Health Commission

[8] https://www.northeastern.edu/csshresearch/bostonarearesearchinitiative/boston-data-portal/

Boston Public Health Commission (BPHC) is an independent public agency providing a wide range of health services and programs. It is governed by a seven-member board of health appointed by the Mayor of Boston. Its mission is to "protect, preserve, and promote the health and well-being of all Boston residents, particularly those who are most vulnerable. BPHC collects data that contribute to the formation of the following construct:

- Public Health

A3. Indicators by Construct

This section provides the final organization of OI constructs as well as the description of each constituent indicator. We also report the average weight (*M weight*) of each indicator across 3 years (2011-2013) which specifies how strongly an indicator contributes to its construct's overall measurement relative to the other indicators. Conventionally, weights greater than +/-.40 are considered a robust contribution to a factor score.

Academic Attainment of Neighborhood

Indicator	*Source*	*Definition*	*M Weight*
Postsecondary attainment	ACS	Percent of census tract adults (25 years and older) who have earned a Bachelor's degree or higher	1.0

Custodianship

Indicator	*Source*	*Definition*	*M Weight*
Custodianship	BARI	The likelihood that residents will use 311 to call in an issue in the public domain (e.g., pothole)	1.0

Neighborhood Health

Indicator	*Source*	*Definition*	*M Weight*
Late-life medical emergencies	BARI	Prevalence of events that reflect major medical emergencies (e.g., stroke)	.61
Youth health emergencies	BARI	Prevalence of events that reflect medical emergencies surrounding birth and young children	.66
Premature mortality	BPHC	Rate of deaths occurring before projected life expectancy.	.45

Neighborhood Safety

Indicator	*Source*	*Definition*	*M Weight*
Gun use	BARI	Rate of events that involve the use of guns (e.g., shooting).	.89
Private crime	BARI	Rate of events that reflect interpersonal conflict in the neighborhood (e.g., domestic violence).	.88
Public crime	BARI	Rate of events that reflect interpersonal violence that do not involve a gun (e.g., fight).	.88

Newcomer

Indicator	*Definition*	*M Weight*
Foreign born	Percent of Census Tract born outside the United States or United States territory	1.0

Physical Disorder

Indicator	*Definition*	*M Weight*
Physical disorder	The deterioration to and denigration of neighborhood structures and spaces, a combination of two measures from 311 reports regarding *private neglect* and *public denigration*.	1.0

Socioeconomic Status

Indicator	*Definition*	*M Weight*
Family poverty	Rate of poverty at the Census Tract level	.88
Median household income	Median household income of Census Tract	-.85
Public assistance	Rate of receipt of SNAP benefits at the Census Tract level	.85
Unemployment	Rate of unemployed individuals 16 years of age and older at the Census Tract level	.84

Residential Stability

Indicator	*Definition*	*M Weight*
Homeownership	Percent of census Tract who owns the residence in which they live	.89
Tenure	Median number of years within census Tract that residents have lived in their residence as of 2010	.89

Appendix B. Weighting the Opportunity Index: Regression Results

In order to select and weight the proposed components of the OI we evaluated their relationship with MCAS scores using multilevel models (also known as hierarchical linear models). Multilevel models are ideal for this question in that they disentangle the effects of individual-level characteristics (e.g., an individual's race, poverty level) from the characteristics of a shared environment (e.g., concentrated poverty), in this case those of a residential neighborhood. Put colloquially, a major early goal of research with multilevel models was to answer the question, "Do poor neighborhoods have lower outcomes solely because the individuals who live there have less individual access to resources, or because the neighborhood has an overall dearth of resources that impacts everyone who lives there?"

We ran a series of multilevel models that used both individual- and neighborhood-level factors to predict MCAS Math and MCAS ELA scores for elementary, middle, and high school students across three different school years (2011-2012, 2012-2013, and 2013-2014). This amounted to 18 separate models (2 tests x 3 age groups x 3 school years). Here we report the results of these models (Tables B1-B6), though see main text for interpretation and implications for the OI. At the individual level, these models accounted for sex, race (broken out into White, Asian, Hispanic, Native American, and Multi-/Other, with Black as the reference group), special needs status, free-or-reduced-price lunch (FRPL) status, and, for 2013-2014 only, English Language Learner (ELL) status (Current, Former, or Never). At the neighborhood (i.e., census tract) level, the models included the constructs of academic attainment of adults, custodianship, public health, public safety, "newcomer," physical disorder, socioeconomic status, and residential stability (see Appendix A for more on their measurement). For each school year, the tract measures associated with the earlier year were used as predictors to best model causality (i.e., measures for 2013 were used to predict academic performance in 2013-2014).

The models were run in five stages. First, a *null model* with no predictors was run to assess raw variance across levels. Second, a *conditional null model* with only individual-level predictors was run, to reveal the true neighborhood-level contribution to scores controlling for features that might cluster across neighborhoods. Third, a *multivariate model* with all neighborhood predictors was run to assess the independent effect of each on academic outcomes. Fourth, a *parsed multivariate model* was run with only the significant neighborhood-level factors included, isolating the most important elements of a neighborhood for academic outcomes; this was done because of strong correlations between many of the neighborhood constructs. Fifth and last, we ran *common models* that included the same set of neighborhood constructs for all tests and school years within an age group. Neighborhood constructs were included in an age group's common models if they were found to be significant predictors in two or more school years for at least one MCAS test subject in that age group. For example, academic attainment of adults in the neighborhood was significant for two of the school years for elementary school performance on MCAS math, but only one for ELA. Based on this, we included academic attainment of adults for all final models for elementary school. Tables B1-B3 report the final results of these common models without race and gender included; Tables B4-B6 report the final results of these common models with race and gender included; and Table B7 reports the results of the risk indicator analysis for high school students.

Before reporting the tables, this Appendix also describes how to interpret the regression results and their tangible implications.

Interpretation of Models

Our multilevel models disentangled individual- and neighborhood-level effects on a given outcome by running two simultaneous sets of regressions.[9] The first assesses the relevance of individual-level characteristics by comparing students to others in the same neighborhood (approximated here with census tracts), and then combining the results from all neighborhoods to estimate the effect of a given characteristic. To illustrate, one can interpret the final parameter estimate for eligibility for free and reduced price lunch as the extent to which someone with that designation does better or worse on a particular MCAS test (depending on the outcome variable) than another student living in the same census tract who is not eligible for free and reduced price lunch. These parameters also hold constant each

[9] The models take the form of:

$Y_{jk} = \beta_{0k} + \beta_{1k} * x_{1jk} + \cdots + \beta_{1k} * x_{njk} + r_{jk}$ Individual-Level Model

$\beta_{0k} = \gamma_{00} + \gamma_{1k} * x_{1k} + \cdots + \gamma_{nk} * x_{nk} + \mu_{0k}$ Neighborhood-Level Model

$r_{jk} \sim N(0, \sigma^2)$ Individual-Level Variance

$\mu_{0k} \sim N(0, \tau_0)$ Neighborhood-Level Variance

where βs are the parameter for the relationship between each variable (x) and the outcome. In the individual-level model, x is the value for each individual (j) in each tract (k). In the neighborhood-level model, x is the value for each tract (k). σ^2 is the measure of variation at the individual level (i.e., within neighborhoods) and τ_0 is the measure of variation in the outcome measure between census tracts. The percentage of variation attributable to census tracts is then calculated as the intraclass correlation coefficient (ICC):

$$ICC = \frac{\tau_0}{\sigma^2 + \tau_0}$$

of the other characteristics included in the individual-level model, communicating an "independent" or "unique" effect.

The second regression is at the neighborhood level, estimating the extent to which each of the proposed OI constructs predicts the mean outcome of all students living in a neighborhood. Importantly, these two sets of regressions are calculated at the same time, so the latter is not the raw mean, but an approximate mean adjusted for the individual-level characteristics of the people who live there. This is critical because individual- and neighborhood-level characteristics are often correlated. For example, students who are eligible for free or reduced price lunch are clustered in certain neighborhoods that have lower overall socioeconomic status. If individuals with this eligibility perform lower on MCAS tests regardless of their neighborhood of residence, we will need to account for this, otherwise we will inflate the relevance of the neighborhood construct. The individual-level model identifies such differences between neighbors and accounts for them, permitting a truer observation of the effect of neighborhood characteristics.

Table B1. Parameter estimates from multilevel models predicting MCAS scores in elementary school students, with race excluded. Average parameters are used in the calculation of the Opportunity Index.

	Parameter | notation		ELA				Math			
			SY1112	SY1213	SY1314	**AVG**	SY112	SY1213	SY1314	**AVG**
Student level	Disability	γ_{70}	-.799***	-.782***	-.798***	**-.793**	-.667***	-.654***	-.682***	**-.668**
	FRP Lunch	γ_{80}	-.337***	-.395***	-.431***	**-.388**	-.370***	-.375***	-.460***	**-.402**
	Current ELL^	γ_{90}	-.445***	-.492***	-.490***	**-.476**	-.199***	-.226***	-.247***	**-.224**
	Former ELL	$\gamma_{10.0}$	.508***	.386***	.311***	**.402**	.612***	.486***	.363***	**.487**
Neigh. level	Academic Attain.	$\gamma_{11.0}$	.077**	.106***	.051*	**.078**	.135***	.118***	.082**	**.112**
	Custodianship	$\gamma_{12.0}$	.071*	.078*	.080**	**.076**	.078*	.080*	.063*	**.074**
	Health	$\gamma_{13.0}$								
	Safety	$\gamma_{14.0}$	-.043*	-.068**	-.066**	**-.059**	-.038~	-.085***	-.048*	**-.057**
	Newcomer	$\gamma_{15.0}$								
	Physical Disorder	$\gamma_{16.0}$	-.119**	-.086*	-.077*	**-.094**	-.131**	-.078*	-.074*	**-.094**
	SES	$\gamma_{17.0}$	-.083**	-.044~	-.070**	**-.066**	-.047~	-.040	-.067**	**-.051**
	Residential Stab.	$\gamma_{18.0}$								
Variance components										
Level-1:	Within-person	σ_{ε}^2	.664***	.662***	.694***		.713***	.733***	.756***	
Level-2:	In initial status	σ_0^2	.015***	.017***	.015***		.026***	.024***	.023***	

^ELL Status=Never ELL is the reference group.

~ = $p < .10$; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; [ns] = *variance* parameter is not statistically significant.

Table B2. Parameter estimates from multilevel models predicting MCAS scores in middle school students, with race excluded. Average parameters are used in the calculation of the Opportunity Index.

	Parameter | notation		**ELA**				**Math**			
			SY1112	SY1213	SY1314	**AVG**	SY112	SY1213	SY1314	**AVG**
Student level	Disability	γ_{70}	-.845***	-.845***	-.819***	**-.836**	-.719***	-.724***	-.633***	**-.692**
	FRP Lunch	γ_{80}	-.238***	-.320***	-.337***	**-.298**	-.289***	-.366***	-.416***	**-.357**
	Current ELL^	γ_{90}	-.864***	-.758***	-.736***	**-.786**	-.523***	-.403***	-.391***	**-.439**
	Former ELL	$\gamma_{10.0}$	.200***	.268***	.303***	**.257**	.273***	.404***	.439***	**.372**
Neigh. level	Academic Attain.	$\gamma_{11.0}$	.082***	.057**	.082***	**.074**	.141***	.117***	.189***	**.149**
	Custodianship	$\gamma_{12.0}$								
	Health	$\gamma_{13.0}$								
	Safety	$\gamma_{14.0}$	-.093***	-.096***	-.042**	**-.077**	-.089***	-.085***	-.026	**-.067**
	Newcomer	$\gamma_{15.0}$								
	Physical Disorder	$\gamma_{16.0}$								
	SES	$\gamma_{17.0}$	-.093***	-.096***	-.042**	**-.077**	-.089***	-.085***	-.026	**-.067**
	Residential Stab.	$\gamma_{18.0}$								
Variance components										
Level-1:	Within-person	σ_{ε}^2	.614***	.618***	.623***		.696***	.693***	.697***	
Level-2:	In initial status	σ_0^2	.010***	.005**	.007**		.030***	.024***	.024***	

*Race/ethnicity=Black is the reference group; ^ELL Status=Never ELL is the reference group.
~ = $p < .10$; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; ns = *variance* parameter is not statistically significant.

Table B3. Parameter estimates from multilevel models predicting MCAS scores in high school students, with race excluded. Average parameters are used in the calculation of the Opportunity Index.

	Parameter	**notation**	**ELA**				**Math**			
			SY1112	SY1213	SY1314	**AVG**	SY112	SY1213	SY1314	**AVG**
Student level	Disability	γ_{70}	-.890***	-.836***	-.770***	**-.832**	-.939***	-.879***	-.871***	**-.896**
	FRP Lunch	γ_{80}	-.219***	-.234***	-.279***	**-.244**	-.165***	-.202***	-.315***	**-.227**
	Current ELL^	γ_{90}	-1.12***	-1.07***	-1.25***	**-1.147**	-.572***	-.660***	-.609***	**-.614**
	Former ELL	$\gamma_{10.0}$	.129***	.226***	.084**	**.146**	.311***	.273***	.252***	**.279**
Neigh. level	Academic Attain.	$\gamma_{11.0}$	.137***	.028	.093**	**.086**	.148***	.091*	.116**	**.118**
	Custodianship	$\gamma_{12.0}$								
	Health	$\gamma_{13.0}$								
	Safety	$\gamma_{14.0}$	-.085***	-.067**	-.069**	**-.074**	-.085**	-.083**	-.054*	**-.074**
	Newcomer	$\gamma_{15.0}$								
	Physical Disorder	$\gamma_{16.0}$								
	SES	$\gamma_{17.0}$	-.027	-.085**	-.049~	**-.054**	-.022	-.090*	-.036	**-.049**
	Residential Stab.	$\gamma_{18.0}$								
Variance components										
Level-1:	Within-person	σ_{ε}^2	.590***	.596***	.565***		.704***	.656***	.693***	
Level-2:	In initial status	σ_0^2	.013**	.018**	.015**		.021**	.035***	.029***	

*Race/ethnicity=Black is the reference group; ^ELL Status=Never ELL is the reference group.
~ = $p < .10$; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; [ns] = *variance* parameter is not statistically significant.

Table B4. Parameter estimates from multilevel models predicting MCAS scores in elementary school students, with race included. Average parameters are used in the calculation of the Opportunity Index.

	Parameter | notation		**ELA**				**Math**			
			SY1112	SY1213	SY1314	**AVG**	SY112	SY1213	SY1314	**AVG**
Student level	Female	γ_{10}	.147***	.152***	.170***	**.156**	-.060***	-.023***	-.057**	**-.047**
	White*	γ_{20}	.488***	.489***	.414***	**.464**	.577***	.579***	.463***	**.540**
	Asian	γ_{30}	.456***	.589***	.522***	**.522**	.854***	.869***	.834***	**.852**
	Hispanic	γ_{40}	.102***	.113***	.057**	**.091**	.144***	.130***	.128***	**.134**
	Native American	γ_{50}	.150	-.085	.133	**.066**	.196	-.100	.164	**.087**
	Multi-/Other	γ_{60}	.356***	.314***	.344***	**.338**	.278***	.293***	.400***	**.324**
	Disability	γ_{70}	-.765***	-.741***	-.746***	**-.751**	-.661***	-.632***	-.659***	**-.651**
	FRP Lunch	γ_{80}	-.272***	-.303***	-.339***	**-.305**	-.296***	-.269***	-.365***	**-.310**
	Current ELL^	γ_{90}	-.448***	-.516***	-.505***	**-.490**	-.241***	-.268***	-.306***	**-.272**
	Former ELL	$\gamma_{10.0}$	.455***	.307***	.263***	**.342**	.511***	.379***	.269***	**.386**
Neigh. level	Academic Attain.	$\gamma_{11.0}$	.020	.051*	-.000	**.024**	.048*	.047*	.020	**.038**
	Custodianship	$\gamma_{12.0}$	.057*	.073**	.066**	**.065**	.063**	.073**	.043*	**.060**
	Health	$\gamma_{13.0}$								
	Safety	$\gamma_{14.0}$	-.032*	-.040*	-.062***	**-.045**	-.021	-.054**	-.054**	**-.043**
	Newcomer	$\gamma_{15.0}$								
	Physical Disorder	$\gamma_{16.0}$	-.087**	-.080**	-.051~	**-.073**	-.093**	-.071*	-.032	**-.065**
	SES	$\gamma_{17.0}$	-.058**	-.025	-.046*	**-.043**	-.026	-.016	-.034~	**-.025**
	Residential Stab.	$\gamma_{18.0}$								
Variance components										
Level-1:	Within-person	σ_ε^2	.641***	.632***	.669***		.671***	.688***	.719***	
Level-2:	In initial status	σ_0^2	.007**	.010***	.008**		.007**	.008***	.006**	

*Race/ethnicity=Black is the reference group; ^ELL Status=Never ELL is the reference group.

~ = $p < .10$; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; [ns] = *variance* parameter is not statistically significant.

Table B5. Parameter estimates from multilevel models predicting MCAS scores in middle school students, with race included. Average parameters are used in the calculation of the Opportunity Index.

	Parameter | notation		ELA				Math			
			SY1112	SY1213	SY1314	**AVG**	SY112	SY1213	SY1314	**AVG**
Student level	Female	γ_{10}	.209***	.251***	.252***	**.237**	-.026***	.029~	-.015	**-.004**
	White*	γ_{20}	.376***	.346***	.373***	**.365**	.487***	.440***	.525***	**.484**
	Asian	γ_{30}	.472***	.427***	.377***	**.425**	.963***	.939***	.928***	**.943**
	Hispanic	γ_{40}	.117***	.111***	.056**	**.095**	.155***	.123***	.115***	**.131**
	Native American	γ_{50}	.007	.137	.335*	**.160**	-.107	.047	.291*	**.077**
	Multi-/Other	γ_{60}	.259***	.191***	.260***	**.237**	.229***	.192**	.312***	**.244**
	Disability	γ_{70}	-.800***	-.799***	-.766***	**-.788**	-.699***	-.697***	-.611***	**-.669**
	FRP Lunch	γ_{80}	-.199***	-.265***	-.258***	**-.241**	-.234***	-.297***	-.322***	**-.284**
	Current ELL^	γ_{90}	-.874***	-.771***	-.730***	**-.792**	-.571***	-.448***	-.419***	**-.479**
	Former ELL	$\gamma_{10.0}$	.142***	.209***	.264***	**.205**	.143***	.282***	.325***	**.250**
Neigh. level	Academic Attain.	$\gamma_{11.0}$	.024	.008	.026	**.019**	.058**	.038*	.089***	**.062**
	Custodianship	$\gamma_{12.0}$								
	Health	$\gamma_{13.0}$								
	Safety	$\gamma_{14.0}$	-.063***	-.067***	-.030*	**-.053**	-.050***	-.045**	-.007	**-.034**
	Newcomer	$\gamma_{15.0}$								
	Physical Disorder	$\gamma_{16.0}$								
	SES	$\gamma_{17.0}$	-.038*	-.046**	-.057***	**-.047**	-.023	-.045**	-.025~	**-.031**
	Residential Stab.	$\gamma_{18.0}$								
Variance components										
Level-1:	Within-person	σ_{ε}^2	.588***	.588***	.593***		.648***	.647***	.649***	
Level-2:	In initial status	σ_0^2	.004**	.004**	.005**		.006**	.006**	.004*	

*Race/ethnicity=Black is the reference group; ^ELL Status=Never ELL is the reference group.
~ = $p < .10$; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; ^{ns} = *variance* parameter is not statistically significant.

Table B6. Parameter estimates from multilevel models predicting MCAS scores in high school students, with race included. Average parameters are used in the calculation of the Opportunity Index.

	Parameter	**notation**	**ELA**				**Math**			
			SY1112	SY1213	SY1314	**AVG**	SY112	SY1213	SY1314	**AVG**
Student level	Female	γ_{10}	.125***	.195***	.180***	**.167**	.028	-.081**	-.011	**-.021**
	White*	γ_{20}	.406***	.420***	.489***	**.438**	.337***	.546***	.471***	**.451**
	Asian	γ_{30}	.548***	.448***	.642***	**.546**	.824***	.930***	.894***	**.883**
	Hispanic	γ_{40}	.062~	.013	.102**	**.059**	-.009	.106**	.083*	**.060**
	Native American	γ_{50}	.480	-.103	.262	**.213**	.097	.163	.142	**.134**
	Multi-/Other	γ_{60}	.231*	.214*	.321**	**.255**	.180	.274*	.268*	**.241**
	Disability	γ_{70}	-.828***	-.774***	-.700***	**-.767**	-.880***	-.849***	-.825***	**-.851**
	FRP Lunch	γ_{80}	-.175***	-.197***	-.208***	**-.193**	-.130***	-.174***	-.246***	**-.183**
	Current ELL^	γ_{90}	-1.142***	-1.050***	-1.254***	**-1.149**	-.604***	-.706***	-.643***	**-.651**
	Former ELL	$\gamma_{10.0}$	.068~	.206***	.100	**.125**	.222***	.188***	.138***	**.183**
Neigh. level	Academic Attain.	$\gamma_{11.0}$	.062*	-.014	.010	**.019**	.062*	.012	.031	**.035**
	Custodianship	$\gamma_{12.0}$								
	Health	$\gamma_{13.0}$								
	Safety	$\gamma_{14.0}$	-.055**	-.030	-.040*	**-.042**	-.049*	-.019	-.022	**-.030**
	Newcomer	$\gamma_{15.0}$								
	Physical Disorder	$\gamma_{16.0}$								
	SES	$\gamma_{17.0}$	-.019	-.072**	-.031	**-.041**	-.026	-.079**	-.009	**-.038**
	Residential Stab.	$\gamma_{18.0}$								
Variance components										
Level-1:	Within-person	σ_{ε}^{2}	.564***	.568***	.525***		.657***	.603***	.642***	
Level-2:	In initial status	σ_{0}^{2}	$.004^{ns}$	$.007^{ns}$	.008*		$.004^{ns}$	$.004^{ns}$	.009*	

*Race/ethnicity=Black is the reference group; ^ELL Status=Never ELL is the reference group.

~ = $p < .10$; * = $p < .05$; ** = $p < .01$; *** = $p < .001$; ^{ns} = *variance* parameter is not statistically significant.

Table B7. Parameter estimates from multilevel models predicting MCAS scores in high school students, with risk indicators included. Average parameters are used in the calculation of the Opportunity Index.

	Parameter | notation	ELA	Math	Avg.
Student level	Disability	-.378***	-.308***	**-.343**
	FRP Lunch	-.269***	-.181***	**-.225**
	Current ELL^	-.337***	-.093*	**-.216**
	Former ELL	-.068*	.037	**-.016**
	Prop. MCAS Math Failures	-.562***	-1.302***	**-.932**
	Prop. MCAS ELA Failures	-.809***		**-.404**
	Prop. Math Course Failures	-.187***	-.423***	**-.305**
	Prop. ELA Course Failures	-.416***	-.174***	**-.295**
	Ever Suspended?	-.145**	-.221***	**-.183**
	Chronic Absentee (<90%)	-.072*	-.225***	**-.148**
Neigh. level	Academic Attain.	.031	.054*	**.042**
	Safety	-.025	-.010	**-.018**
	SES	-.033	.012	**-.010**
Variance components				
Level-1:	Within-person	.345***	.314***	
Level-2:	In initial status	.005^{ns}	.006^{ns}	

Table B8. Parameter estimates from multilevel models predicting MCAS scores in middle school students, with risk indicators included. Average parameters are used in the calculation of the Opportunity Index.

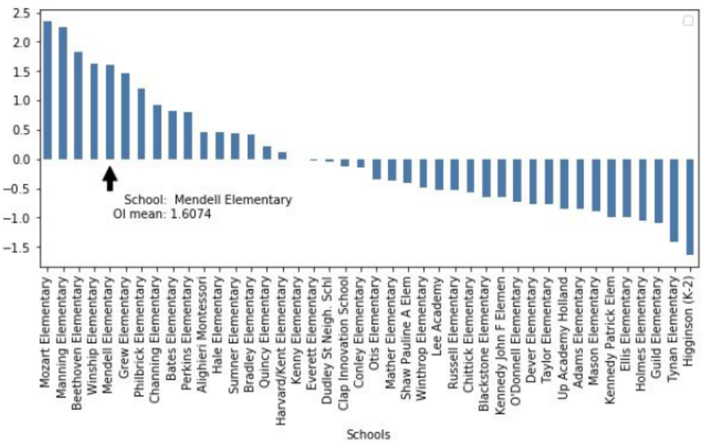
Parameter | notation		ELA	Math	Avg.
Student level	Disability	-.654***	-.524***	**-.585**
	FRP Lunch	-.412***	-.487***	**-.449**
	Current ELL^	-.738***	-.451***	**-.594**
	Former ELL	.302***	.420**	**.356**
	Ever Suspended?	-.223**	-.259***	**-.241**
	Chronic Absentee (<90%)	-.185***	-.325***	**-.255**
Neigh. level	Academic Attain.	.088**	.165***	**.126**
	Safety	-.041~	-.028	**-.034**
	SES	-.078**	-.047~	**-.062**
Variance components				
Level-1:	Within-person	.597***	.637***	
Level-2:	In initial status	.013*	.019*	

Appendix C. Example Visualizations for Communicating the OI

School Report for Mendell Elementary

Overview

The Opportunity Index (OI) is a data-based tool that uses a combination of individual and place-based factors to quantify the assets and deficits of students across Boston Public Schools that can impact their academic achievement. This report summarizes the OI scores of Mendell Elementary's student body.



Among all Elementary Schools in Boston, Mendell Elementary has the fifth highest Opportunity Index, with a mean student score of **1.61**.

Components of the OI Score

This graph shows how Mendell Elementary scores on each of the place-based components of the OI--that is, reflecting the average neighborhood environment of its students relative to other elementary schools. The numbers on the right are percentiles relative to other schools. For some variables, it is better to be higher (e.g., socioeconomic status) and for others it is better to be lower (e.g., crime).