Operator versus Partner: A Case Study of Blueprint School Network's Model for School Turnaround

John P. Papay
Matthew A. Kraft
Jessalynn K. James
*Brown University**

April 2020

Abstract – Numerous high-profile efforts have sought to "turn around" low-performing schools. Evidence of these programs' effectiveness, however, is mixed, and research offers little guidance on which types of turnaround models are more likely to succeed. We present a case study of turnaround efforts led by the Blueprint Schools Network in three schools in Boston. Using a difference-in-differences framework, we find that Blueprint raised student achievement in mathematics and ELA by at least a quarter of a standard deviation, on average. We document qualitatively how differential impacts across the three Blueprint schools relate to contextual and implementation factors. In particular, Blueprint's role as a turnaround partner (in two schools) versus school operator (in one school) shaped its ability to implement its model. As a partner, Blueprint provided expertise and guidance but had limited ability to fully implement its model. In its role as an operator, Blueprint had full authority to implement its turnaround model, but was also responsible for managing the day-to-day operations of the school, a role for which it had limited prior experience.

---

*The authors thank Boston Public Schools and Blueprint Schools Network for providing the necessary data, interview respondents who generously donated their time to speak with us, and Matt Spengler, Apryl Clarkson, Mary Dillman, Alan Melchior, and others who provided comments on this paper or earlier reports. The Blueprint Schools evaluation was funded by a grant from the federally-funded Social Innovation Fund (SIF), a program of the Corporation for National and Community Service (CNCS), through a grant to the GreenLight Fund via Blueprint Schools Network (for whom we served as an external third-party evaluator).

Introduction

In recent years, states and school districts have undertaken a range of high-profile efforts to "turn around" their lowest-performing and least-improving schools. The evidence of turnarounds' effectiveness, however, is mixed and provides little guidance on which types of models are more likely to succeed. While some turnaround efforts have had substantial success (e.g., Bonilla & Dee, 2020; Carlson & Lavertu, 2018; Dee, 2012; LiCalsi, Citkowicz, Friedman, & Brown, 2015; Papay & Hannon, 2019; Sun, Penner, & Loeb, 2017; Zimmer, Henry, & Kho, 2017), others have failed or even had negative effects on student achievement (e.g., de la Torre et al., 2013; Dee, 2012; Dee & Dizon-Ross, 2019; Dragoset et al., 2017; Heissel & Ladd, 2016; Hemelt & Jacob, 2017; Strunk, Marsh, Hashim, Bush-Mecenas, & Weinstein, 2016). Understanding the circumstances that permit some turnaround interventions to succeed while others flounder is vital to future school improvement efforts.

A common but understudied feature of many turnaround efforts involves hiring external partners to support school improvement efforts. These independent turnaround specialists typically either work as a partner with the school, district, or state agency to facilitate turnaround efforts, or lead these processes themselves as operators (VanGronigen & Meyers, 2017). Differences across these roles raise important questions. Are turnarounds more successful with an external partner or with an operator? Can turnaround specialists succeed in both roles? How do the differences in these roles shape the turnaround process?

We study the efforts of one such turnaround specialist, Blueprint Schools Network, to partner with two persistently underperforming schools in Boston and to directly operate a third. Blueprint aims to implement comprehensive, school-level reforms by building human capital,

providing intensive tutoring, increasing instructional time, using data to improve instruction, and setting a culture of high expectations.

This paper presents a case study of Blueprint's experience in Boston Public Schools (BPS), with evidence supported by the tools of causal inference. Our case study makes two key contributions. First, despite limited statistical power due to our case-study approach, we show that the Blueprint turnaround model improved student achievement in these three schools. Using a difference-in-differences framework, we find Blueprint raised student achievement in math and reading by at least a quarter of a standard deviation (SD) after three years, respectively. Understanding the effectiveness of specific turnaround models is imperative for disentangling the mixed findings across turnarounds as a whole, as well as for identifying rapid school improvement approaches that can succeed in particular settings.

Second, we use evidence from the implementation of the Blueprint model in these schools to explore what makes a turnaround effort succeed or fail by identifying the opportunities and constraints faced by Blueprint. Specifically, using interview data collected from a diverse range of stakeholders, we document how differential impacts across the three schools relate to contextual and implementation factors. We pay particular attention to differences that stemmed from the structure of its relationship as a partner verses operator in a given school. Prior evaluations of school turnarounds have rarely incorporated qualitative evidence (with the notable exception of Strunk et al., 2016). Doing so helps to elucidate the factors that might explain the successes (or failures) of a specific turnaround effort.

We show that the contextual features – including the interplay with other stakeholders such as state, district, and school officials – influenced Blueprint's efforts in these three schools. In particular, Blueprint's role as partner or operator was pivotal to its ability to implement its

model in each school. Each role created distinct opportunities and obstacles for Blueprint's scope of influence across these stakeholders.

Background

To date, states, districts, and schools have little guidance about what models or turnaround approaches might yield greater improvements. The most prominent advice comes from an Institute of Education Sciences (IES) guidebook for turning around low-performing schools (Herman et al., 2008), released as part of the US Department of Education's What Works Clearinghouse program. This guide recommends that turnaround efforts: 1) signal the need for dramatic change with strong leadership; 2) maintain a consistent focus on improving instruction; 3) make visible improvements early in the school turnaround process (i.e., quick wins); and 4) build a committed staff. Each of these recommendations, however, is made with the disclaimer that this guidance is based on a series of case studies of turnaround schools that were available at the time of the guide's writing, rather than through experimental or quasi-experimental evidence of their effectiveness.

More than a decade after IES issued its turnaround guide, a substantial evidence base around the *impact* of school turnaround efforts has developed, but its conclusions are mixed and typically do not speak to the implementation details or contexts that support successful turnarounds. This literature demonstrates that the effectiveness of turnarounds is variable, but it provides little new guidance for the implementation and design decisions that might have yielded a given outcome.

Evidence on School Turnarounds

Evidence from prior turnaround efforts across the country demonstrates that while these interventions can accomplish significant academic improvements, they often fail to meet or even

approach their ambitious goals. There have been documented successes for schools receiving school improvement (SIG) grants or accountability waivers in some states (e.g., Bonilla & Dee, 2020; Henry, Guthrie, & Townsend, 2015; LiCalsi et al., 2015; Papay & Hannon, 2019; Sun et al., 2017). Meanwhile, other turnaround efforts either resulted in a mixed record of success (e.g., de la Torre et al., 2013; Dee, 2012; Dougherty & Weiner, 2019; Henry & Guthrie, 2019; Henry, McNeill, & Harbatkin, 2019; Strunk et al., 2016; Zimmer, Henry, & Kho, 2017), or negative effects on students in affected schools (Heissel & Ladd, 2016).

Identifying the aspects of turnaround programs that might define success or failure is challenging, as programs typically encompass a package of reforms that vary from school to school. The variety of turnaround models implemented across schools can make it difficult to parse out whether certain approaches are more effective than others, as well as to identify contexts that might impede success for otherwise promising reform models. A handful of lessons have, however, recently emerged from other turnaround settings.

First, even among successful turnarounds, it can take multiple years before a program demonstrates positive effects on student achievement (de la Torre et al., 2013; Player & Katz, 2016; Sun et al., 2017). Second, efforts that are more substantial (e.g., replacing a larger share of the school staff) tend to have larger impacts than those that are more superficial (Hill, in press). Third, the ability of schools and other turnaround stakeholders to overcome critical obstacles can be pivotal for their success (Henry, Lam, Kho, & Zimmer, 2019). One central obstacle is staff turnover after the turnaround effort begins (Gill et al., 2007; Henry et al., 2015; Strunk et al., 2016), particularly in locales where there is not a ready supply of qualified replacement teachers (Le Floch et al., 2014; Scott et al., 2012). Fourth, several studies have found that the quality of supports for teachers are particularly important, as programs that embed targeted professional

development tend to demonstrate greater turnaround success (Bonilla & Dee, 2020; Sun, Liu, Zhu, & LeClair, 2019)

Finally, buy-in of school staff—and in particular of the school principal—is also key for the successful implementation of turnaround models; without core stakeholders adhering to the model, interventions can fail where they might otherwise have had a positive impact (Strunk et al., 2016). It is likewise important that the entity leading turnaround efforts builds relationships with the larger community—not just the families of students, but also larger political forces such as the teachers' unions and local departments of education (Glazer, Massell, & Malone, 2019; Meyers & Sadler, 2018; Schueler, 2019).

While recent literature has begun to explore key factors and dimensions for successful turnaround implementation, to date there has been scant attention to external organizations and, in particular, to the specific role such organizations play in designing and implementing the turnaround model. In Boston, Blueprint was engaged in two capacities: as a school turnaround partner and later as a turnaround school operator. As a partner, Blueprint assisted two schools with the implementation of their turnaround models—planning how the model might work in the school, supporting implementation, and providing monitoring and feedback support. As an operator, Blueprint held final responsibility for the same tasks and for running all aspects of the school. A priori, it is not clear how Blueprint's capacity in each role would facilitate or hinder its turnaround efforts. This paper supplements the emerging literature on turnaround mechanisms with a case study that highlights this important dimension of implementation.

School Turnaround in Massachusetts

In January of 2010, Massachusetts passed legislation giving the state wider latitude to intervene in low-performing districts and schools. These policy changes positioned the state well

to compete for federal SIG dollars; the state received $250 million in such funding in 2010. As part of the legislation, the state identified the lowest-performing and least-improving schools as Level 4 schools in need of substantial intervention and targeted resources for rapid improvement. The state did not require a single turnaround model. Instead, it offered districts and school leaders some flexibility over staffing and resource allocation and provided technical assistance. More recently, the system has placed Level 4 schools that have not improved into Level 5 status, making them subject to state takeover.

The Blueprint Model

Blueprint's turnaround model encompasses five core components: 1) Ensuring excellence in school leadership and instructional quality; 2) Increasing instructional time for students through extended school days and years; 3) Developing a culture of high expectations with an explicit focus on college-going culture; 4) Using data and regular formative assessments to track student performance and focus instruction; and 5) Providing daily, in-school, small-group tutoring (via the Math Fellows program) to support students in "critical growth years". Each of these elements is defined in detail in Table 1.

To implement its turnaround model in a given school, Blueprint goes through a three-phase process. The initial phase is focused on due diligence and strategic planning, during which Blueprint spends time with school and central office administrators to understand the local contexts. In the second phase, Blueprint provides technical assistance to support implementation of the program's five-point framework. Finally, Blueprint engages in ongoing monitoring, evaluation, and reflection, as needed. This process is intended to inform implementation over time and allows Blueprint to customize strategies and solutions for each school.

A precursor to Blueprint's turnaround framework was scaled in Houston, Texas under the Apollo 20 program. Apollo 20 was implemented in a subset of the lowest-performing schools in Houston in order to evaluate a set of practices for school turnaround. A randomized experiment found positive effects on participating schools (Fryer, 2013). Blueprint's turnaround efforts were grounded in similar principles to the Apollo 20 program, but the organization continued to evolve its practice and adapt its interventions to the local context.

Implementing Blueprint in Boston

In 2013, Blueprint was selected by the Boston Public Schools to serve as external lead turnaround partner in the turnaround of two "persistently underperforming" schools: English High School (EHS) and Elihu Greenwood Leadership Academy (EGLA). Table 2 describes the period and capacity of Blueprint's involvement with each school. Blueprint worked with EHS for the full three years of its contract (2013-14 through 2015-16), but the district closed EGLA following the 2014-15 school year. This closure decision was made before state assessment results for that year were available.

One year later, the Massachusetts Department of Elementary and Secondary Education (Mass DESE) selected Blueprint to serve as an operator for a third school, the Paul A. Dever Elementary School (Dever). Dever was one of only four schools in the state that had been rated "chronically underperforming" (Level 5) on the state's accountability system. It was thus subject to a different set of accountability mechanisms and contractual flexibilities than other BPS schools, and Blueprint reported directly to the state Commissioner rather than to district officials. As a result, while Dever remained a BPS school, Blueprint was named a Level 5 Receiver for the school in January 2014 and signed a three-party agreement with BPS and DESE to take control of operations at the start of the 20104-15 school year. At Dever, Blueprint

worked as a school operator akin to a Charter Management Organization (CMO), running the school's day-to-day operations. Blueprint worked with Dever for the duration of its initial three-year contract with DESE; in 2017-18, receivership transferred to the new superintendent of Boston Public Schools.

Data and Methods

The three Blueprint schools that we evaluate in this case study are broadly representative of other low-performing elementary and high schools in Boston (and, more generally, other urban school districts) in terms of performance and student population served. In Table 3, we present select demographic characteristics for our test-taking sample across the three years. We compare students in Blueprint Schools to those in all Level 4 schools in BPS and to all students in the district. While Blueprint's schools served a much lower-performing and less advantaged population than the city overall, they were roughly comparable to other Level 4 schools. For example, 89% of Blueprint students had low family income, compared to 88% of students in Level 4 schools and 77% of students in BPS as a whole. Similarly, in our sample, Blueprint students' average prior-year mathematics test scores were 0.33 SDs below the district mean, compared to 0.27 SD below the mean for Level 4 students.

Impact on Student Achievement

Data. To estimate Blueprint's effects on student achievement in these three schools, we rely on a panel of administrative data from BPS. Specifically, we use data from 2005-06 through 2016-17, when the last of these schools completed the Blueprint intervention. The data include detailed information about students, teachers, and schools in BPS. Our student achievement outcomes are derived from student scores on the state assessments in mathematics and English

language arts (ELA), which are available for students in grades 3 through 8 and 10.[1] We standardize test scores by subject, grade, and year. Thus, we can interpret our effects as representing SD differences in scores.

Our key predictor is whether a student attends a Blueprint School. For the relatively small number of students who enroll in multiple schools in a given year, we identify them as Blueprint students if they ever enrolled at one of the schools during the year. We discuss the implications of student mobility and attrition for the analysis alongside our results.

Sample. We focus on elementary and high school students in tested grades in Boston. For EGLA and Dever Elementary Schools, we focus on students in grades 3-5, the only tested grades. For English High School, we limit our sample to 10th grade. Our analyses compare Blueprint students to two different comparison groups: students in all other BPS schools and students in other Level 4 schools. The Level 4 schools make for a more credible comparison group, given that they are similarly low-performing, but with only 11 other schools in this category, these estimates may be less stable than those that use the full BPS sample. These two samples also rely on different conceptual counterfactuals, as other Level 4 schools were undergoing related turnaround interventions during the same period. Thus, these estimates compare Blueprint's model to other turnaround efforts, rather than to business as usual in the

[1] Through 2013-14, we use results from the Massachusetts Comprehensive Assessment System (MCAS) tests in mathematics and ELA. In 2014-15 and 2015-16, some schools in BPS used the MCAS tests while others used PARCC assessments. In 2016-17, the state transitioned to new "next generation" state MCAS examination. We use concordance data from the state Department of Education to link PARCC scores to the MCAS scale. Because Blueprint schools were differentially less likely to use PARCC rather than MCAS, one might be concerned about this biasing our point estimates if, for example, the transition to the new tests disrupts achievement. Our standardization approach means our estimates are defined relative to other students in the district in the same year; we simply examine at what point in the district's test-score distribution students fall. However, if the tests measure somewhat different constructs, we may conflate differences in true performance with differences in the test. Importantly, though, the tests have become more rigorous and aligned with new college and career ready content standards (see http://www.doe.mass.edu/mcas/nextgen/ for details.); thus, any relative improvements for Blueprint schools would either reflect true performance improvements or reflect better alignment between instruction in these schools and the more demanding content standards. Regardless, we test for robustness of results for an indicator of whether the students took the PARCC versus CAS exam, and find our estimates virtually unchanged.

district. When considered together, these two comparison groups might therefore provide bounds of Blueprint's effects. Because the district had only one other Level 4 high school, we restrict our comparisons for EHS to all district high schools.[2]

Quasi-Experimental Approach. Our sample sizes are necessarily quite small, with only three Blueprint schools. Thus, we see our approach as a small-scale case study. However, quantitative analysis in such a case study can still rely on the tools of causal inference more commonly applied in settings with larger samples. Here, our preferred models take the introduction of Blueprint supports as an arguably exogenous shock to school performance. We estimate difference-in-differences models that describe changes over time in Blueprint schools, after Blueprint's involvement, to changes over time in other schools in the district. We fit different versions of these models, accounting for time trends more and less parametrically.

We first estimate a non-parametric event-study model, which accounts for differences between achievement outcomes of Blueprint and other BPS (or Level 4) students at each year of the intervention, holding year effects constant. This model takes the following form:

$$Y_{ist} = \beta_0 + \sum_{j=-7}^{j=3} \beta_{t_c} \mathbf{1}\{t_c = j\} + \kappa_{gt} + X_{it}'\delta + \lambda_s + \varepsilon_{ist} \quad (1)$$

for student i in grade g, school s, and year t. In this model, X_{it} is a vector of student characteristics, κ_{gt} represents grade-by-year fixed effects and λ_s represents school fixed effects. β_{t_c} is a vector of coefficients relative time centered on the year Blueprint engaged with a school. In other words, year $t_c = 0$ is the first year of Blueprint's involvement in the school and we omit the year preceding the Blueprint intervention ($t_c = -1$). The values of β_{t_c} for $t_c = [0, 3]$

[2] As a robustness check, we explore restricted EHS comparisons that omit magnet schools in BPS. These analyses, available upon request, yield results that are nearly identical to those that use the full BPS comparison sample.

thus represent the change in student achievement associated with Blueprint's involvement—i.e., the average treatment effect—in a given year of Blueprint involvement, relative to the last pre-treatment year. Student characteristics include gender, race and ethnicity, and whether the student has a disability, comes from a low-income family, or is an English learner. We cluster our standard errors at the school level to account for correlated errors among students in the same school.[3]

The equation above allows effects to vary by year. We also estimate a complementary, more parametric model, which takes the following form:

$$Y_{ist} = \beta_0 + \beta_1(t_c) + \beta_2(t_c \geq 0) + \beta_3 t_c * (t_c \geq 0) + \kappa_{gt} + X_{it}'\delta + \lambda_s + \varepsilon_{ist} \quad (2)$$

Here, we can interpret parameter β_2 as an intercept shift capturing the immediate effect on student achievement in the first year of Blueprint implementation ($t_c = 0$) and β_3 as the incremental effect of an additional year of Blueprint implementation. The coefficient on the pre-treatment year relative time trend, β_1, provides a parsimonious test of the parallel trends assumption between Blueprint and non-Blueprint schools. As expected from our visual analysis of the event study figures, we fail to reject this null hypothesis in any of our models, suggesting that the parallel trends assumption holds.

The models above focus on the aggregate impact of Blueprint's involvement across the three schools, but we are also interested in exploring school-by-school variations in treatment effects and documenting how these relate to features of implementation. Thus, we also fit models

[3] We report traditional standard errors in our results tables, given these are the standard approach in the turnaround literature even in cases with few treated schools. However, because cluster robust standard errors can perform poorly and lead to over-rejection of the null hypothesis when there is a small number of treated clusters (see Cameron, Gelbach, and Miller, 2008), we also estimate wild cluster bootstrap-t tests. These tests demonstrate insufficient power to detect significant effects across most years and models, but the strongest effects remain statistically significant after these adjustments.

that estimate the treatment effect for each individual school, excluding the other Blueprint schools.

Understanding Implementation and Context Across the Blueprint Intervention

The second component of our case study analysis comes from interviews with key stakeholders and implementation data collected from the schools themselves. These analyses offer a more nuanced exploration of the intervention's rollout across the three schools. Given the limited research base about what makes some turnarounds succeed while others fail, examining the contextual factors that supported or constrained Blueprint's model implementation illuminates lessons that might inform other turnaround efforts. We seek to put our school-level impact estimates in context with insights from the implementation process. In other words, we explore *why* might see the achievement results we do for each of these three schools.

We use two general approaches to build our understanding of the Blueprint rollout in BPS. First, we interviewed a variety of stakeholders to understand how state and district policies and systems influenced school-level implementation. Second, we collected data about the implementation of each component of Blueprint's core framework in the schools.

Semi-structured interviews. We conducted seven interviews with a range of important stakeholders including officials from the Massachusetts Department of Elementary and Secondary Education, BPS central office staff, building administrators, and Blueprint staff. We sought a sample with diverse perspectives on Blueprint's work in the schools and with different relationships with Blueprint. Each semi-structured interview lasted approximately 30 minutes. Interview questions focused on the interplay between Blueprint and three levels of governance: the state, district, and school. We also probed the influence of Blueprint's role as operator vs.

partner in program implementation (see Appendix B for the interview protocol). We took detailed field notes that we then organized into thematic summaries (Maxwell, 2005) describing key insights from each interview. We exposed themes across interviews by writing analytic memos and checking those against the understanding of the researchers on the team.

Document review. We complement the findings from these interviews with reviews of internal documents that provide evidence on implementation fidelity in each participating school. Each year, we collected a series of documents to evaluate the extent to which Blueprint's model had been implemented in each school (see Appendix Table 1). Evidence sources included: site-visit agendas and executive reports; school calendars and weekly schedules; materials, schedules, and flyers from staff recruitment and selection efforts; and other operational information including the staffing, training, and supervision of Math Fellows. These data complement and allow us to cross-check and validate findings from our semi-structured interviews.

Overall Results

Impact estimates. Overall, we find positive effects of Blueprint's involvement on student achievement across the three schools. Figure 1 illustrates the trends in achievement in Blueprint schools relative to other schools over time, using estimates from our primary specification (Model 1). The estimates in these figures are relative to achievement in Blueprint schools in the year before Blueprint's involvement. We see, regardless of the comparison group or subject, no clear trend in achievement before Blueprint engaged with the school. This suggests that pre-trends for Blueprint schools were similar to those of non-Blueprint schools and supports the parallel trends assumption—a conclusion that is supported by the test for differences in pre-trends we describe above (Model 2). These event-study results reveal initially relatively small or null impacts in the first year, with subsequent steady and increasing returns to student

achievement in years two and three.

We present the point estimates from each of our models in Table 4. Estimates from Model 1 represent the β_{t_c} coefficients, capturing the impact on achievement in each year of Blueprint's involvement. Estimates from Model 2 reflect linear combinations of coefficients β_2 and β_3 for each specific year. While these estimates derive from small samples, they provide insight into the overall pattern of Blueprint's impact. Again, we present estimates from both comparison groups. In math, both models indicate null effects in the initial year, with larger but still imprecise gains in the second year. By the third year, however, Blueprint students were scoring at least a quarter of a SD (Model 1) higher than those in other BPS schools. These effects are robust to alternative modeling decisions, with the more parametric model (Model 2) yielding effects of roughly a third of a SD by year three, though this estimate is not statistically significant. ELA effects are of similar magnitude, growing to at least 0.25 SD by the third year. Unlike in math, we see evidence of gains in the first year in ELA. These effects are similarly large in the more parametric model (Model 2), equivalent to at least a quarter of a SD.[4]

Threats to validity. The central threat to validity is student mobility. We might worry about student mobility if, for example, Blueprint's test score gains reflect the schools pushing out the lowest-performing students or attracting higher-performing students. Three pieces of evidence suggest that this is not the case.

First, student mobility did not change substantially at Blueprint schools across the intervention period relative to non-Blueprint schools. Rates of school entries in tested grades were consistent across our samples; 17 percent of all BPS and Level 4 students were new to their schools during the pre-intervention period, compared to 18 percent in Blueprint schools. In the

[4] Estimates in ELA are more precise than those in math, reflecting larger variation in math treatment effects across the three schools; we discuss this in more detail in the School-Level Results section of the paper.

post-intervention period, mobility fell equally in the treatment and comparison groups. The share of students who were new to their schools was again similar across all BPS (13 percent), Level 4 (12 percent), and Blueprint (12 percent) students. We find that 11 percent of all BPS students (10 percent of Level 4 students) in tested grades exited their schools in the pre-intervention period, compared to 9 percent of Blueprint students. During the intervention period, 4 percent of Blueprint students exited, compared to 10 and 9 percent of all BPS and Level 4 students, respectively. Thus, there is little evidence that Blueprint's involvement drove students away from these schools or that they actively recruited a substantially different student body.[5]

Second, we explore whether Blueprint's involvement with schools altered the characteristics of enrolling and exiting students. We test for such dynamic sorting using a version of Model 1, replacing the outcome with each student covariate in turn and excluding other student covariates. Our estimates suggest that Blueprint's involvement did not substantially change the composition of student enrollment in these schools over the course of the intervention (see Appendix Table 2). Across the seven characteristics we test, we see consistent evidence of only one difference: Blueprint schools saw an increasing number of English learners relative to other Level 4 schools. We believe this pattern suggests that, if anything, our estimates would understate Blueprint's effects on student achievement.

Third, we re-specify our primary models to include student fixed effects, explicitly focusing on variation within student over time. Our estimates are of similar magnitude and are statistically significant. We do not feature this model because of the somewhat limited sample of

[5] Depending on the direction of sorting, changes to the share of entering and exiting students might reflect the turnaround's effectiveness if the schools attracted stronger students than they had previously. Regardless, changes to the student body could bias estimates of the effects on student achievement. To distinguish between these types of sorting, we also explore rates of entry and exit within student (not shown). Patterns of within-student mobility are consistent with the estimates that do not employ student fixed effects.

students with repeated observations, but the consistency in estimated effects indicates that selection is not driving our results.

Qualitative data. Our impact estimates show that Blueprint had little effect on student achievement in its first year in each of these BPS schools, but that student achievement substantially improved by its second and third year. These patterns are consistent with reports from interviewees about initial limitations on Blueprint's authority within each school, based on the circumstances under which these relationships began. Both in its role as an external turnaround partner at EHS and EGLA, and as an operator at Dever, the decision about Blueprint's involvement was made at the state level, with little involvement from the school or district. Our interviews suggested that this dynamic led to initially turbulent relationships in some of the schools. For example, multiple interviewees at EHS described long-standing residual "baggage" surrounding the way in which Blueprint was introduced to the school and the way these relationships were framed by BPS to school staff. As one interviewee noted, "if you want someone to partner with you in a collaborative fashion, you can't have it be your boss and tell you that you're being taken over." Stakeholders reported that Blueprint, therefore, spent its initial year at each of these schools building relationships with key partners, including the Mass DESE, BPS administrators, and school-level staff.

Broader contextual factors also played a role. Blueprint entered BPS at a time of substantial superintendent turnover. The system as a whole was unstable, as key players waited for the new regime to take over. Delays and contested negotiations in this relationship-building process led to somewhat sporadic implementation of the Blueprint model in the first year. These narrative accounts of the difficulties in the first year of the turnaround process accord with the lack of impacts on student achievement in the first year.

Over time, however, Blueprint made efforts to build stronger partnerships, to use these partnerships to leverage additional control over staffing, to articulate more clearly its model of classroom instruction, and to adapt to the specific needs and contexts of each school. These efforts led to the Blueprint model being incorporated more directly into the school culture and practices. Multiple interviewees commented that, despite initial challenges, Blueprint was responsive to its mistakes and improved substantially over the years—a perspective that is supported by substantial gains to student achievement in Blueprint's second and third years in these schools.

School-level Results

Because aggregate results obscure important school-level heterogeneity, we conduct exploratory analyses to examine how differences implementation might be related to differences in impact in each school individually. Figure 2 illustrates school-specific event study effects from our primary specification (Model 1). Appendix Figure 1 illustrates raw (unadjusted) test-score trajectories over time for each of the three Blueprint schools relative to BPS as a whole, as well as other Level 4 schools. Echoing our overall results above, in nearly all cases, test scores in Blueprint Schools were substantially higher at the end of Blueprint's involvement than they had been before Blueprint engaged with the school. We observe a general pattern of increasing test scores across the years Blueprint supported or led these turnaround efforts.

We also find considerable variation across schools and over time (Figure 2), likely reflecting the distinct experience of each school as it attempted to implement and sustain the model over the planned three-year intervention. The patterns across these findings, paired with evidence from stakeholder interviews, suggest that differences in implementation directly relate to differences in test-score trajectories.

In particular, our implementation and interview data point to three core lessons about how contextual and structural factors can impede or support turnaround efforts. First, the role that Blueprint played in each school mattered a great deal. Blueprint had distinct advantages and challenges at Dever (where it served as an operator) compared to the other two schools. As an external turnaround partner in EGLA and EHS, Blueprint lacked official authority over school operations, but engaged in work with which it had more substantial experience. By contrast, Blueprint had more authority but less experience to inform its work as a turnaround operator at Dever. Second, given the top-down nature in which Blueprint entered these schools, it needed to build and develop relationships with stakeholders over the course of the intervention. Each school's (and school leader's) openness to Blueprint and its model influenced the feasibility of developing this buy-in. Third, Blueprint's success depended on its ability to hire and retain effective staff, including principals who, in some cases, turned over rapidly. These and other factors affected Blueprint's ability to implement its model in each of these schools. We describe experiences, and corresponding achievement outcomes, for each school in turn below and summarize our implementation data in Appendix Table 1.

English High School. Interviews suggest that when Blueprint first entered EHS as a partner in the fall of 2013, it struggled to implement its model. In addition to entering EHS without prior buy-in from the school, Blueprint was limited in terms of its ability to screen and select new staff at EHS. Instead, the principal, an experienced and well-respected administrator in BPS, prioritized using her own network and connections in the district to facilitate the hiring process. Given the few open positions and the principal's preferred hiring approach, Blueprint did not implement its formal screening process at EHS, although it did advertise positions and refer candidates. Similarly, as a partner, Blueprint's access to and influence with teachers ran

directly through the principal; at EHS, the school administrator was highly experienced, had her own theory of action for leading change drawn from her past work in the district, and was less likely to prioritize Blueprint's recommendations and support. Several interviewees saw Blueprint as encroaching on the authority of the school's senior administrator. This lack of alignment on how to implement reforms at EHS was reflected with small achievement effects in year one at EHS.

Over time, however, Blueprint was gradually able to improve its relationships with relevant stakeholders, and by the second year, students were exhibiting large ELA achievement gains—effects that are robust to model specification—which grew in magnitude by year three. In math, on the other hand, Blueprint faced significant barriers to the successful implementation of its Math Fellows program. While, at the start of Blueprint's involvement in EHS, 9th grade students received the full intended dosage of math tutoring, in the following year (2014-15) Blueprint reduced the frequency of tutoring sessions by half due to scheduling changes that made it impossible to offer tutoring each day (see Appendix Table 1 for a summary of implementation fidelity). Tutoring was also spread among both 9th graders and 10th graders at the school. At the same time, the BPS teachers' union filed a lawsuit against the district for hiring non-union math fellows, a challenge that Blueprint attempted to resolve by transitioning the Math Fellows to Blueprint's, rather than BPS's, payroll. By this point (i.e., year two) the tested tenth graders were receiving tutoring for the second year. Each of our difference-in-difference models indicate year-two math effects that are correspondingly large, positive, and statistically significant relative to other BPS high schools.

By year three (2015-16), Blueprint and BPS decided mutually not to continue the Fellows program because of the union's ongoing lawsuit, the marginalized implementation of the Fellows

program in 2014-15, and the challenge of fitting the Fellows program into the schedule. Instead, they replaced time that had been dedicated to tutoring with an elective period. These struggles to implement and maintain Blueprint's math tutoring program may explain the negative math performance trajectories by its third year at EHS. Student achievement declined relative to the preceding years, with effect estimates that are statistically indistinguishable from the pre-treatment year.

Elihu Greenwood Leadership Academy. As with EHS, Blueprint was appointed as an external turnaround partner to EGLA by DESE in 2013. While relations with school leadership at EHS were strained, Blueprint worked with BPS to jointly hire a new principal for EGLA in year 1, and then replace this principal with a better match in year 2. These principals sought out more direct support and involvement from Blueprint around promoting staff morale and building a school-wide culture of excellence. At the same time, Blueprint relied on the school principals to help integrate Blueprint supports with the ongoing work of the school. This worked more effectively at EGLA, where the principals' vision tended to align more closely with Blueprint's work than at EHS, likely because Blueprint was a partner in the hiring decision and the principals entered the school knowing Blueprint would be involved.

Interviewees reported that by the second year Blueprint was better able to incorporate its model into the school's culture and practices as it built partnerships and gained control over staffing. Blueprint also exerted more influence over staffing decisions at EGLA, working directly with BPS officials to select the new principal and assistant principal, replacing 10 teachers, and promoting competency-based teacher hiring practices. Recruitment and selection-day materials suggest that Blueprint increased the rigor of the screening process to include both demonstration lessons and multi-part interviews, representing a clear change from existing

hiring practices.

Our quantitative evidence suggests that student achievement was increasing in both math and ELA during the two years Blueprint was active at EGLA before the school closed. By year two, math achievement had improved by nearly 0.5 SD, relative to other Level 4 schools, while ELA scores had increased by nearly 0.2 SD (Table 5). In early 2015, BPS decided to close the school. This decision was made well before the school's performance on state tests were available.

Dever Elementary School. Blueprint's experience with Dever differed considerably from EHS and EGLA. In January 2014, Blueprint was named the school's Level 5 Receiver and took over operations and control of the building on July 1. As a Level 5 school in Massachusetts, Dever was subject to a different set of accountability mechanisms and contractual flexibilities than other BPS schools, and Blueprint reported directly to the state Commissioner rather than to district officials. This provided Blueprint far greater latitude and authority for managing school operations, including more direct control over program implementation, staff compensation, budget, and hiring. Interviewees indicated that this level of autonomy and authority was freeing, giving Blueprint the flexibility to implement its model as desired, but also daunting. This was Blueprint's first attempt running a school, and interviewees reported that the limits of Blueprint's expertise were evident. First, Blueprint did not have the background in managing operational administrative tasks that more established charter management organizations do. While it staffed the school successfully—hiring 72 new staff members and a new administrative team—the timing of the formal Level 5 announcement provided only several months of lead time before school opened. Blueprint had to quickly create from scratch a system and pipeline to recruit large numbers of teachers and principals. More-experienced operators would likely have already had

such structures in place.

Connected to Blueprint's early operational struggles in Dever was the manner in which the receivership was granted. The Commissioner had named Blueprint the operator independently of BPS. While Blueprint and BPS maintained a productive working relationship, this arrangement led to some points of tension that produced inefficiencies at times. For example, Blueprint hoped to leverage BPS systems to provide some administrative support (e.g., budget, human resources, transportation, facilities and special education, etc.), particularly in areas where it did not have as much expertise, but BPS was less eager to engage in such cross-organization functional efficiencies with Blueprint given that the district had lost control of the school. As a result, some of Blueprint's collaborative efforts did not bear fruit early on.

Substantial school leadership turnover early in Blueprint's intervention at Dever served as a further impediment to success. For example, the original principal hired by Blueprint left during the first intervention year (2014-15). The second principal also left mid-year, requiring Blueprint's Network Director to then step in as acting principal. Blueprint struggled over this period to find a strong principal who was a good match for the school, as well as the broader turnaround context under which it was operating. These struggles were apparent in the student achievement outcomes documented in Table 6, with largely negative trends—roughly 0.14 SD decline—in math test scores for the first two years of Blueprint's receivership.

Interviews suggest that Blueprint learned directly from aspects of the first year of operation that did not go as well at Dever and changed its practices in future years. The organization became more aware of the tasks it needed to complete on its own and those for which it could rely on district support. This aligns with an upward trajectory of Dever's test-

scores. For example, Blueprint completely redesigned the school schedule and extended the school day, creating an additional hour for core math and reading instruction.

Blueprint's status as Dever's operator shielded it from the challenges it experienced trying to implement the Math Fellows program at EHS. Blueprint was able to implement the intended math tutoring dosage in Dever, and to add tutoring to additional grades as the need arose. Like EGLA, Blueprint was also able to increase the rigor of its screening process relative to prior hiring protocol. Blueprint successfully implemented competency-based best practices such as an intensive selection-day event where job candidates completed data analysis activities, participated in group discussions on instructional practices, had individual interviews, and toured the schools.

At the same time, stakeholders reported that Dever's Level 5 status had a substantial impact on hiring practices. Dever's hiring occurred outside of the traditional BPS system, and the school had a separate salary scale from other BPS schools. As a result, most new teachers came from outside of the district. The Boston Teachers Union (BTU) also fought against Dever's Level 5 turnaround plan, which fundamentally changed the working conditions and compensation framework for teachers. Blueprint was trying at the same time to assert its authority as an operator and change the status quo, which made striking this balance and forging a productive relationship with BTU leadership quite challenging. At the school site, however, Blueprint and union representatives who were part of the faculty met and worked collaboratively to navigate local issues.

By the third year (2016-17), Blueprint's efforts at Dever began to pay off as it settled into its role in the school and managed to address earlier issues. Blueprint had conducted a successful national search for a principal and designed a new principal hiring process that involved multiple

local and state-level stakeholders. Interviewees noted that Dever thrived once Blueprint identified the right school leader. They did not assert that the principal alone led the rapid improvement in student achievement, but that a strong leader who bought into the Blueprint model enabled the potential of the model. In year three, math achievement increased substantially, with improvements of approximately a third of a SD, regardless of estimation approach. While ELA trends were generally positive in the first two years, scores also increased substantially by year three.

Discussion & Conclusion

Across the three schools, we find large effects of Blueprint's turnaround efforts on student achievement. By the third year of Blueprint's involvement, impacts were equivalent to roughly a quarter of a SD in ELA, with slightly larger effects in math. These effects are comparable in size to those of turnaround schools in the state of Massachusetts as a whole (LiCalsi et al., 2015; Papay & Hannon, 2019) and in Kentucky (Bonilla & Dee, 2020), and are consistent with other settings where turnaround effects were initially small or null but accrued over time (e.g., de la Torre et al., 2013; Player & Katz, 2016). The initial difficulties that Blueprint faced with implementation were followed by improvements that reflected a flexibility to adapt to circumstances, particularly as it built relationships and established buy-in from other essential stakeholders (Schueler et al., 2017; Strunk at al., 2016).

Two core themes emerge from Blueprint's experiences in BPS, along with lessons for other turnaround interventions. First, the role that Blueprint played—i.e., as partner versus operator—was pivotal to its ability to implement its model. Each role brought distinct advantages and drawbacks, and shaped Blueprint's scope of influence. Importantly, there was widespread agreement across stakeholders that Blueprint was initially better-equipped to be an external lead

partner, as it was at EHS and EGLA, than a school operator, as it was at Dever. Blueprint had substantially more experience as a partner and excelled at several aspects of this work. However, this role also came with a lack of authority which limited its ability to help schools implement the model. Blueprint faced critical barriers in its work as an external partner. These barriers stemmed in part from Blueprint's lack of official authority over school operations, as well as from the way the partnership had been initially framed. The Math Fellows program is a primary example of such limitations. Blueprint was unable to secure sufficient time during the school schedule to implement the program as planned, and this component of Blueprint's model was ultimately abandoned at EHS. On the other hand, Blueprint initially had minimal expertise in operating a school and was less equipped to do so. It needed to learn school management tasks and build up operational capacity over the course of Dever's turnaround program. Over time, however, it was able to more fully implement its vision.

A second key lesson, also related to Blueprint's role as operator or partner, is the importance of strong and aligned school leadership for school success. Given Blueprint's strengths and limitations as a school operator, it relied heavily on a strong school leader who had a robust vision for instruction in the school. This worked effectively in EGLA, where the principal's vision tended to be more supportive of Blueprint's work, but was less effective in EHS, where the principal was less receptive to Blueprint's guidance and feedback. This lack of alignment stymied the extent to which Blueprint could fully implement its model.

This case study adds to a small but increasingly nuanced literature on school turnaround effects and factors that influence how successfully turnaround models are implemented. While our quasi-experimental estimates demonstrate large overall effects on student achievement, our qualitative analyses allow us to dig deeper into the contexts that may have differently facilitated

Blueprint's effectiveness across these three schools. These analyses demonstrate that the nature of the turnaround partnership is integral to the autonomy with which a turnaround organization can implement its model. At the same time, it is inseparable from the relationships that it has with relevant school and political stakeholders, without whom the operator cannot establish a coherent and aligned intervention. An effective turnaround partner must balance developing and maintaining these relationships with sufficient authority to faithfully roll out its model.

References

Bonilla, S., & Dee, T. (2020). The effects of school reform under NCLB waivers: Evidence from Focus Schools in Kentucky. *Education Finance and Policy, 15*(1), 75-103.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics, 90*(1), 414-427.

Carlson, D. & Lavertu, S. (2018). School Improvement Grants in Ohio: Effects on student achievement and school administration. *Educational Evaluation and Policy Analysis, 40*(3), 287-315.

De la Torre, M., Allensworth, E., Jagesic, S., Sebastian, J., Salmonowicz, S., Meyers, C., & Gerdeman, R. D. (2103). *Turning around low-performing schools in Chicago*. Chicago, IL: The University of Chicago Consortium on School Research.

Dee, T. S. (2012). *School turnarounds: Evidence from the 2009 stimulus* (NBER Working Paper No. 17990). Cambridge, MA: National Bureau of Economic Research.

Dee, T. S. & Dizon-Ross, E. (2019). School performance, accountability, and waiver reforms: Evidence from Louisiana. *Educational Evaluation and Policy Analysis, 41*(3), 316-349.

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., … & Giffin, J. (2017). *School Improvement Grants: Implementation and Effectiveness* (NCEE 2017-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Fryer, R. G. (2013). *Injecting successful charter school strategies into traditional public schools: A field experiment in Houston* (NBER Working Paper No. 17494). Cambridge, MA: National Bureau of Economic Research.

Gill, B., Zimmer, R., Christman, J., & Blanc, S. (2007). *State takeover, school restructuring, private management, and student achievement in Philadelphia*. Santa Monica, CA: RAND Corporation.

Glazer, J. L., Massell, D., & Malone, M. (2019). Charter schools in turnaround: Competing institutional logics in the Tennessee Achievement School District. *Educational Evaluation and Policy Analysis, 41*(1), 5-33.

Heissel, J. A. & Ladd, H. F. (2016). *School turnaround in North Carolina: A regression discontinuity analysis* (CALDER Working Paper No. 156). Washington, DC: National Center for the Analysis of Longitudinal Data in Education Research, American Institutes for Research.

Hemelt, S. W. & Jacob, B. (2017). *Differentiated accountability and education production: Evidence from NCLB waivers*. (NBER Working Paper No. 23461). Cambridge, MA: National Bureau of Economic Research.

Henry, G. T. & Guthrie, J. E. (2019). *The effects of Race to the Top school turnaround in North Carolina.* (EdWorkingPaper: 19-107). Retrieved from Annenberg Institute at Brown University: http://www.edworkingpapers.com/ai19-107

Henry, G. T., Guthrie, J. E., & Townsend, L. W. (2015). *Outcomes and impacts of North Carolina's initiative to turn around the lowest-achieving schools*. Chapel Hill NC: Consortium for Educational Research and Evaluation.

Henry, G.T., McNeill, S.M., & Harbatkin, E. (2019*). Effects of school turnaround on K-3 student achievement* (EdWorkingPaper No.19-66). Retrieved from Annenberg Institute at Brown University: http://edworkingpapers.com/ai19-66

Henry, G. T., Pham, L., Kho, A., & Zimmer, R. (2019). Peeking into the black box of school turnaround: A formal test of mediators and suppressors. EdWorkingPaper No. 19-44. Retrieved from Annenberg Institute at Brown University: http://edworkingpapers.com/ai19-44.

Herman, R., Dee, T. S., Greene, J. P., Maynard, R. A., Redding, S. & Darwin, M. (2008). *Turning around chronically low-performing schools: A practice guide* (NCEE #2008-4020). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Hill, H. (in press). What does it take to get a turnaround school turned around? *Education Week.*

LiCalsi, C., Citkowicz, M., Friedman, L., & Brown, M. (2015). *Evaluation of Massachusetts Office of District and School Turnaround Assistance to Commissioners' districts and schools: Impact of school redesign grants.* Washington, DC: American Institutes for Research.

Maxwell, J. A. (2005). Qualitative Research Design: An Interactive Approach (2nd Ed. ed.). Thousand Oaks, CA: Sage.

Meyers, C. & Sadler, J. (2018). District leaders engaged in school turnaround: Identified challenges and espoused responses. *NASSP Bulletin, 102*(2), 89-110.

Papay, J. P. & Hannon, M. (2019). *The effects of school turnaround strategies in Massachusetts*. Working Paper.

Player, D. & Katz, V. (2016). Assessing school turnaround: Evidence from Ohio. *The Elementary School Journal, 116*(4), 675-698.

Schueler, B. E. (2019). A third way: The politics of school district takeover and turnaround in Lawrence, MA. *Educational Administration Quarterly, 55*(1), 116-153.

Schueler, B., Goodman, J. & Deming, D. (2017). Can states take over and turn around school districts? Evidence from Lawrence, Massachusetts. *Educational Evaluation and Policy Analysis, 39(2), 311-332.*

Scott, C., McMurrer, J., McIntosh, S., & Dibner, K. (2012). *Opportunities and obstacles: Implementing stimulus-funded school improvement grants in Maryland, Michigan, and Idaho*. Washington, DC: Center on Education Policy.

Strunk, K. O., Marsh, J. A., Hashim, A. K., Bush-Mecenas, S., & Weinstein, T. (2016). The impact of turnaround reform on student outcomes: Evidence and insights from the Los Angeles Unified School District. *Education Finance and Policy, 11*(3), 251-282.

Sun, M., Liu, J., Zhu, J., & LeClair, Z. (2019). Using a text-as-data approach to understanding reform processes: A deep exploration of school improvement strategies. *Educational Evaluation and Policy Analysis, 41*(4), 510-536.

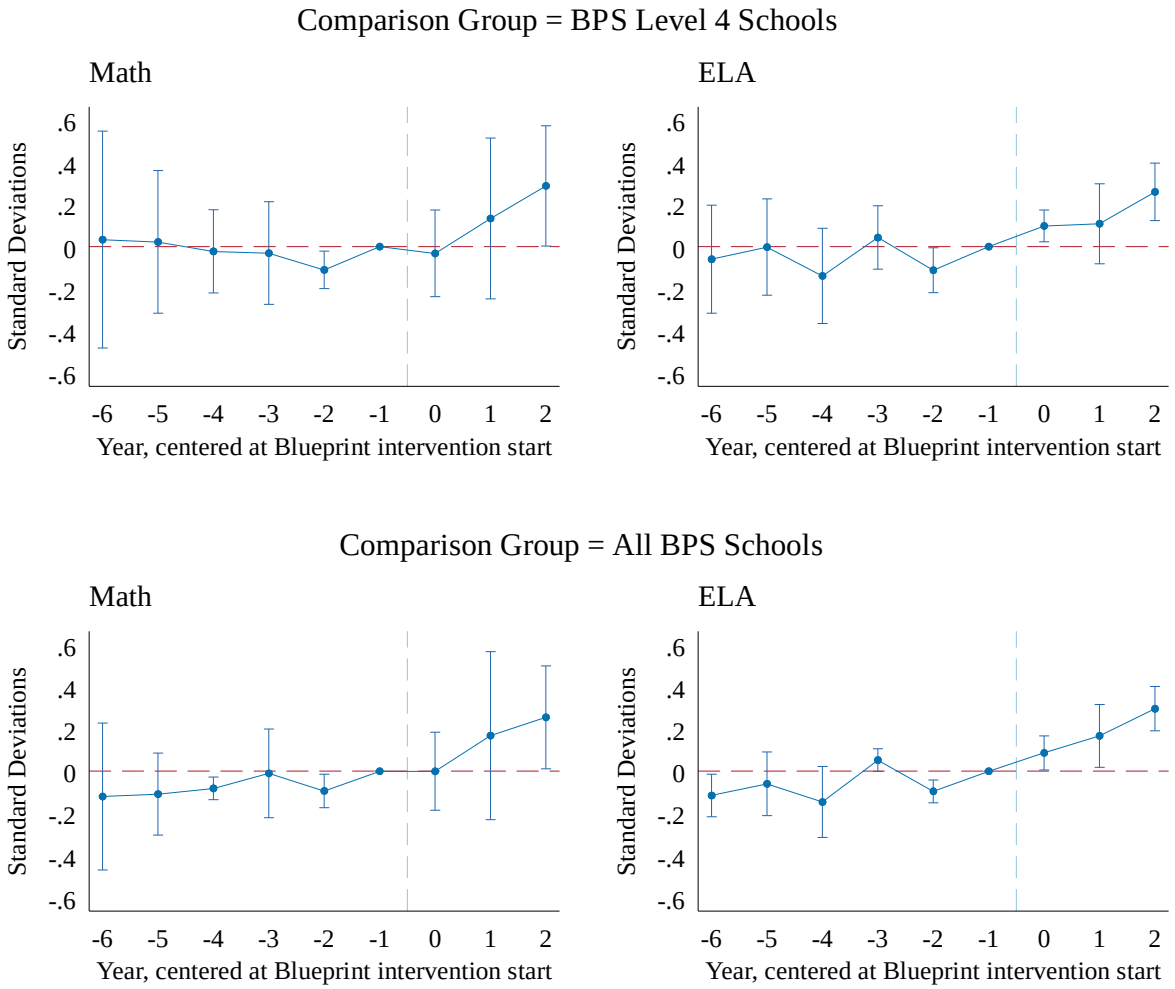
Sun, M., Penner, E., & Loeb, S. (2017). Resource- and approach-driven multidimensional change: Three-year effects of school improvement grants. *American Educational Research Journal, 54*(4), 607-543.

VanGronigen, B. A. & Meyers, C. V. (2019). How state education agencies are administering school turnaround efforts: 15 years after No Child Left Behind. *Educational Policy, 33*(3), 423-452.

Zimmer, R., Henry, G. T., & Kho, A. (2017). The effects of school turnaround in Tennessee's Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis, 39*(4), 670-696.

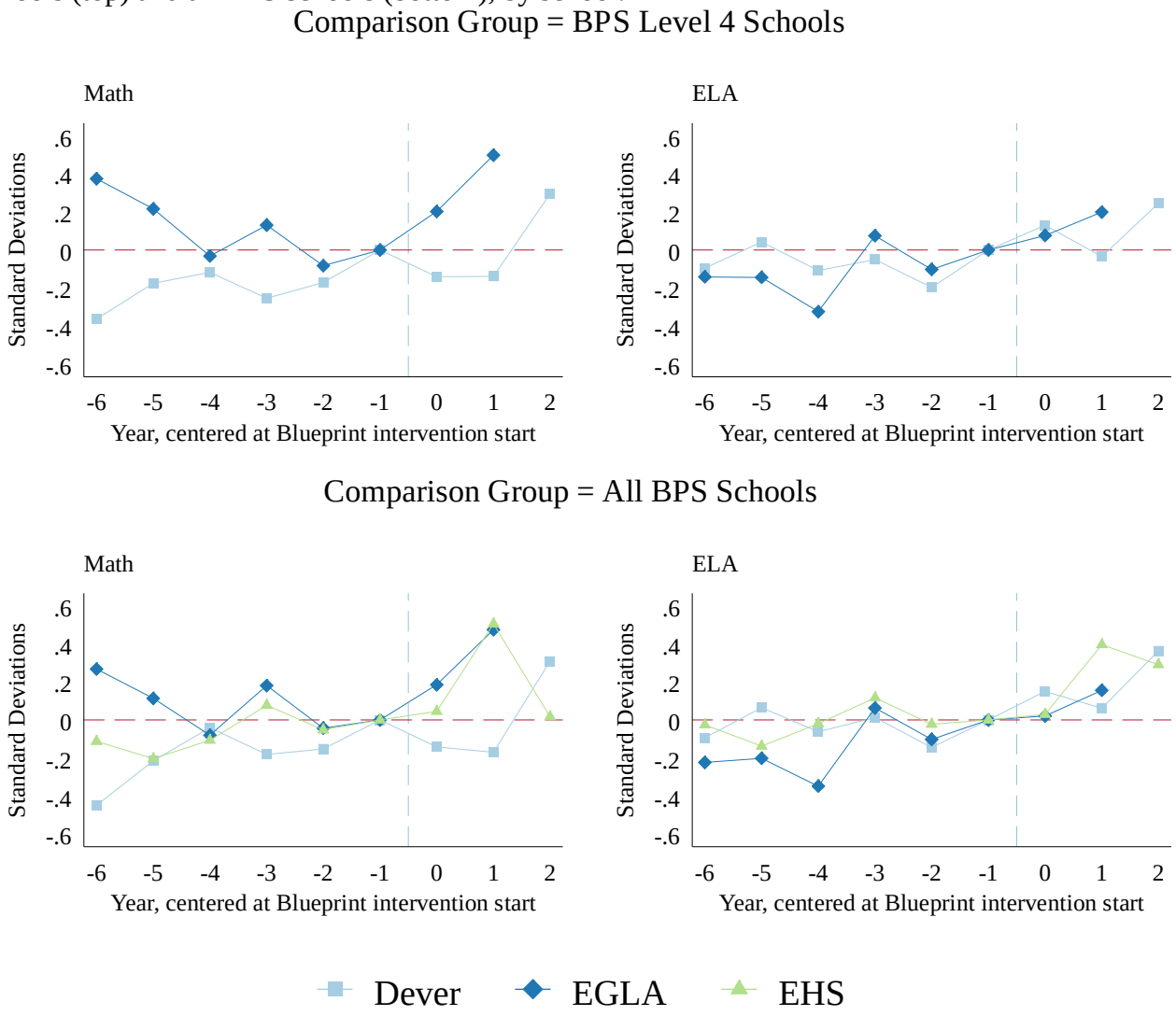
Figures and Tables

Figure 1. Event study results for Math (left) and ELA (right) students in Blueprint schools relative to students in all other Level 4 schools (top) and all BPS schools (bottom).



NOTE: Year 0 is centered to indicate the first year of Blueprint involvement. Error bars represent 95% confidence intervals.

Figure 2. Event study results for Math (left) and ELA (right) students in Blueprint schools relative to students in all other Level 4 schools (top) and all BPS schools (bottom), by school.



NOTE: Year 0 is centered to indicate the first year of Blueprint involvement. Level-4 comparisons are omitted for EHS because there is only one Level 4 high school outside of EHS.

Table 1. The Blueprint Schools Network turnaround model

PROCESS
1. **Due diligence and strategic planning:** During this phase, Blueprint spends time with school and central office administrators to understand the contexts within which they will be implementing their model. This encompasses: a. Working to identify practices, policies, and systems (e.g., human resources, data collection, scheduling, etc.) that might impede or promote implementation; b. Strategizing adaptations or alterations to district or school-level policies and systems in order to better serve partner schools' students; c. Conducting site visits—with classroom observations, student performance data analysis, and conversations students, teachers, and school leaders—to develop a baseline understanding of strengths and weaknesses; and d. Building relationships with key stakeholders in the community (e.g., network superintendents and their teams, school leaders, district foundations, community representatives, and local religious leaders). 2. **Technical assistance and implementation support:** The second phase calls for Blueprint to provide extensive technical assistance to support the implementation of customized district and school turnaround plans, systems, and structures, following their five-point framework (see the five strategies described below). 3. **Ongoing monitoring, evaluation, and reflection:** This occurs in the form of a series of formal site visits, every four to six weeks, with the partner school. These site visits include classroom observations, focus groups with teachers, tutors, and students, and debrief sessions with school leadership. This process includes: a. Delivering feedback on strengths and areas for growth; b. Tracking progress towards education goals; c. Ensuring that Blueprint's turnaround strategies are being implemented effectively throughout the network; and d. Helping schools reflect and prioritize, as well as helping judge the effectiveness of selected strategies to achieve those priorities. The results of this process are distilled into a report for each school. Action items are identified and Blueprint's field-based team and district partners work with school leadership to address challenges. Identified areas for improvement, and their corresponding action items, are re-visited in subsequent site visits.
STRATEGIES
1. Excellence in leadership and instruction: Blueprint's human capital team works with schools and districts to recruit and select highly-effective principals and teachers. Recruitment supports include: a. Human capital and recruitment staff who work to build talent for network schools; b. Partnerships with pathway organizations, including school leadership graduate programs and Teach for America; c. Access to Blueprint's website and national job posting boards for posting hiring information; and

d. Screening of candidates at both the leader and teacher levels for previous performance, as well as alignment with participating schools in terms of beliefs and experience.

Blueprint also regularly monitors instructional effectiveness and provides recommendations to improve the quality of teaching and learning.

2. **Daily tutoring in critical growth areas:** Blueprint supplements classroom instruction with individualized tutoring during the regular school day through the Blueprint Fellows Program—a comprehensive tutoring program designed to accelerate mathematics achievement. Blueprint manages the national recruitment, selection, training, and professional development of a corps of full-time Fellows for each school and district.

   Fellows meet daily with 3–4 students at a time for a 45–60-minute tutorial. These sessions are an ongoing part of each student's daily schedule. Tutorial lessons are designed to include a 5-minute warm-up activity, 15–25 minutes of practice in foundational skills (i.e., computation and problem-solving), 20–30 minutes of support in grade-level content, and an end-of-lesson assessment.

3. **Increased instructional time:** Blueprint works to increase the time students spend on-task and engaged in meaningful learning activities by working with districts to explore options for adding five to ten days to the beginning of the academic year and extending daily schedules by an hour each day. Prior to the start of the school year, Blueprint also collaborates with school principals to create master schedules that use the increased instructional time to maximize planning, intervention, re-teaching, and professional development opportunities.

4. **Culture of high expectations for all:** Prior to the start of the school year, Blueprint's model calls for it to partner with district leadership and principals to develop plans, systems, and tools to improve school safety, climate, learning environments, and expectations for students. The organization provides tools, resources, and strategies geared toward building a positive, college-focused school culture. Blueprint expects all schools to visibly reflect their high expectations for students and staff, both in the classrooms and in public spaces.

5. **Use of data from frequent assessments to improve instruction:** Blueprint works with district and school leaders to implement data-driven instructional systems so that teachers can identify struggling students and differentiate classroom instruction accordingly. Given that districts vary in the frequency and quality of interim assessments administered, as well as their capacity to collect and analyze this data, Blueprint works to understand and help build this infrastructure as needed.

NOTE: Additional information is available on Blueprint's website (https://blueprintschools.org/our-framework).

Table 2. Schools included in the Blueprint Schools Network evaluation

School	Blueprint Role	Years	Grades Served
English High School (EHS)	Partner	2013-14 to 2015-16	9-12
Elihu Greenwood Leadership Academy (EGLA)	Partner	2013-14 to 2014-15*	K-5
Dever Elementary School (Dever)	Operator	2014-15 to 2016-17	K-5

* *EGLA was closed at the end of the 2014-15 school year.*

Table 3. Demographic characteristics and prior-year test scores in Blueprint Schools, Level 4 Schools, and all Boston Public Schools

	Blueprint	**All Level 4**	**All BPS**
African-American	0.386	0.451	0.343
Asian-American	0.030	0.031	0.086
Hispanic	0.545	0.488	0.430
White	0.030	0.023	0.127
Special Educational Services	0.221	0.185	0.189
Low Income	0.886	0.881	0.770
Limited English Proficient	0.352	0.305	0.259
Math test score (prior year, std.)	-0.325	-0.266	0.025
ELA test score (prior year, std.)	-0.373	-0.325	0.022
Sample Size*	*1,383*	*8,779*	*60,512*

* NOTE: Sample sizes for prior-year test scores are substantially smaller.

Table 4. Estimated effect of Blueprint Schools implementation on student test scores in mathematics and ELA, by evaluation approach

		Math		ELA	
		Model 1	Model 2	Model 1	Model 2
Level 4	$t_c = 0$	-0.042	0.011	0.092 ***	0.120 *
		(0.109)	(0.179)	(0.034)	(0.053)
	$t_c = 1$	0.126	0.179	0.108	0.191 *
		(0.195)	(0.219)	(0.096)	(0.084)
	$t_c = 2$	0.277 +	0.347	0.251 ***	0.262 *
		(0.148)	(0.273)	(0.069)	(0.130)
		20,514	*20,514*	*20,594*	*20,594*
All BPS	$t_c = 0$	0.000	0.001	0.084 *	0.089 +
		(0.094)	(0.156)	(0.042)	(0.047)
	$t_c = 1$	0.170	0.110	0.167 *	0.177 ***
		(0.203)	(0.150)	(0.076)	(0.060)
	$t_c = 2$	0.257 *	0.219	0.296 ***	0.266 ***
		(0.125)	(0.168)	(0.054)	(0.073)
		139,730	*139,730*	*139,313*	*139,313*

NOTES: Cell entries include point estimates, robust standard errors (in parentheses), sample sizes (in italics), and approximate p-values (+ p<0.10; * p<0.05; ** p<0.01; *** p<0.001). Estimates in Model 1 are the coefficients on each year indicator; coefficients in Model 2 are the results of linear combinations of the appropriate parameter estimates.

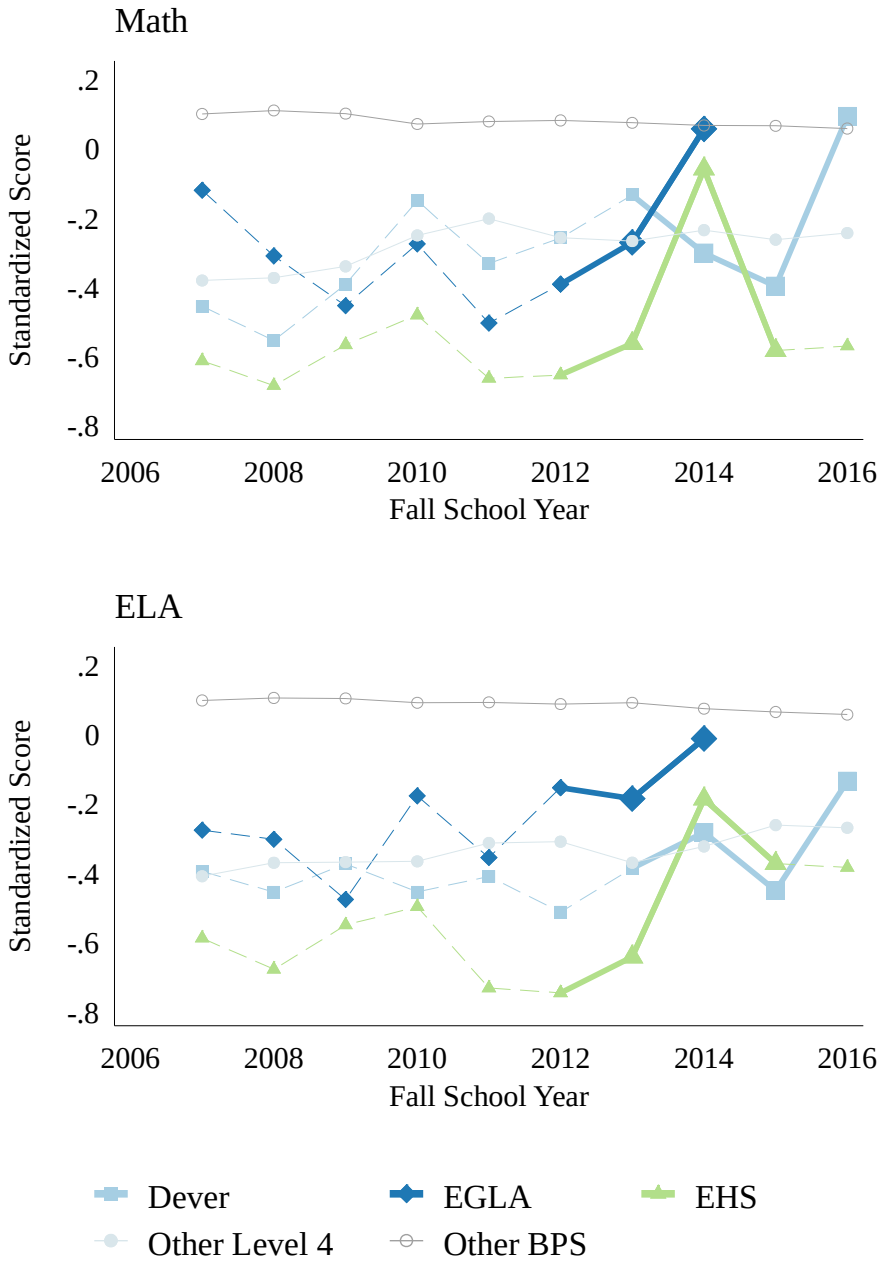
Table 5. Estimated effect of Blueprint Schools implementation on student test scores in mathematics and ELA, by school.

	EHS [a]		EGLA		Dever	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
			Math			
Comparison Group = BPS Level 4 Schools						
$t_c = 0$	--	--	0.205 ***	0.360 ***	-0.153 *	-0.178
			(0.050)	(0.082)	(0.069)	(0.112)
$t_c = 1$	--	--	0.490 ***	0.713 ***	-0.147	-0.008
			(0.082)	(0.160)	(0.090)	(0.171)
$t_c = 2$	--	--	--	--	0.283 *	0.162
			--	--	(0.129)	(0.235)
n			*12,830*	*12,830*	*16,968*	*16,968*
Comparison Group = All BPS Schools						
$t_c = 0$	0.045	0.155 ***	0.184 ***	0.258 ***	-0.141 ***	-0.247 ***
	(0.032)	(0.026)	(0.019)	(0.026)	(0.022)	(0.030)
$t_c = 1$	0.505 ***	0.097 ***	0.475 ***	0.588 ***	-0.168 ***	-0.086 *
	(0.033)	(0.029)	(0.024)	(0.042)	(0.026)	(0.040)
$t_c = 2$	0.015	0.038	--	--	0.311 ***	0.075
	(0.035)	(0.043)	--	--	(0.030)	(0.054)
n	*32,381*	*32,381*	*78,939*	*78,939*	*102,537*	*102,537*
			ELA			
Comparison Group = BPS Level 4 Schools						
$t_c = 0$	--	--	0.074 *	0.057	0.128 **	0.165 **
			(0.038)	(0.056)	(0.053)	(0.065)
$t_c = 1$	--	--	0.202 ***	0.147	-0.035	0.224 *
			(0.071)	(0.094)	(0.069)	(0.098)
$t_c = 2$	--	--	--	--	0.240 ***	0.283 +
			--	--	(0.093)	(0.151)
n			*12,942*	*12,942*	*17,037*	*17,037*
Comparison Group = All BPS Schools						
$t_c = 0$	0.029	0.066 ***	0.015	-0.032 +	0.148 ***	0.145 ***
	(0.022)	(0.025)	(0.015)	(0.019)	(0.023)	(0.026)
$t_c = 1$	0.395 ***	0.175 ***	0.156 ***	0.054	0.060 **	0.250 ***
	(0.030)	(0.024)	(0.023)	(0.033)	(0.024)	(0.027)
$t_c = 2$	0.290 ***	0.283 ***	--	--	0.363 ***	0.354 ***
	(0.030)	(0.029)	--	--	(0.024)	(0.037)
n	*32,601*	*32,601*	*78,523*	*78,523*	*101,906*	*101,906*

[a] Level-4 comparisons are omitted for EHS because there is only one Level 4 high school outside of EHS.
NOTE: Cell entries include point estimates, robust standard errors (in parentheses), and approximate *p*-values (+ $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$). Estimates in Model 1 are the coefficients on each year indicator; coefficients in Model 2 are the results of linear combinations of the appropriate parameter estimates.

Appendix A. Additional tables and figures

Appendix Figure 1. Test-score trends over time in mathematics (top panel) and English language arts (bottom panel) in the three Blueprint Schools, other Level 4 Schools, and other BPS schools, from 2007-08 to 2016-17



NOTE: Solid lines indicate years when Blueprint is engaged with a given school.

Appendix Table 1. Implementation constructs and fidelity of implementation, by school and year

Indicator		EHS	EGLA	Dever
Core Strategy 1: Excellence in Leadership and Instruction				
1. Number of Site Visits per school	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Full	Full	Partial
	$t_c = 2$	Full		Partial
2. Provision of Executive Report after Site Visit	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Full	Full	Partial
	$t_c = 2$	Full		Partial
3. Use of competency-based best practice in teacher hiring	$t_c = 0$	Full	Full	Full
	$t_c = 1$	No	Full	Full
	$t_c = 2$	No		Full
4. BPS and Blueprint agree to co-select principals	$t_c = 0$	Partial	Full	Full
	$t_c = 1$	Partial	Full	Full
	$t_c = 2$	Partial		Full
Core Strategy 2: Increased Instructional Time				
1. Initiated school schedule change to increase time in Math and ELA supports	$t_c = 0$	Partial	Full	Full
	$t_c = 1$	Partial	Full	Full
	$t_c = 2$			Full
Core Strategy 3: Using Data to Improve Instruction and Learning				
1. Blueprint provides technical assistance for school leadership teams to use data to improve instruction	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Full	Full	Full
	$t_c = 2$	Full		Full
Core Strategy 4: Culture of High Expectations				
1. Provide materials and technical assistance to support positive behavior systems, learning environments, goal-setting, and college-going culture	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Full	Full	Full
	$t_c = 2$	Full		Full
Core Strategy 5: Daily Tutoring in Critical Growth Years				
1. Proportion of Math Fellow slots selected and trained by the beginning of school	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Full	Full	Full
	$t_c = 2$	No		Full
2. Identified students received 45-60 minutes of tutorial per day per week	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Partial	Full	Full
	$t_c = 2$	No		Full
3. A site-based coordinator is identified and trained to support the Math Fellows program	$t_c = 0$	Full	Full	Full
	$t_c = 1$	Full	Full	Full
	$t_c = 2$	No		Full

Appendix Table 2. Test for student sorting

		African American	Asian-American	Hispanic	White	Special Ed. Services	Low Income	LEP
Level 4	$t_c = 0$	0.012	-0.011	0.004	-0.001	-0.011	0.010	0.063
		(0.050)	(0.009)	(0.056)	(0.004)	(0.019)	(0.015)	(0.042)
	$t_c = 1$	-0.026	-0.010	0.037	-0.001	-0.015	-0.011	0.111 +
		(0.054)	(0.016)	(0.071)	(0.005)	(0.010)	(0.025)	(0.061)
	$t_c = 2$	-0.058	-0.023	0.087	-0.002	-0.019	-0.022	0.126 ***
		(0.048)	(0.015)	(0.058)	(0.005)	(0.016)	(0.020)	(0.038)
		22,412	*22,412*	*22,412*	*22,412*	*22,412*	*22,412*	*22,412*
All BPS	$t_c = 0$	0.006	-0.012 +	0.005	0.002	-0.008	0.027 +	0.032
		(0.048)	(0.006)	(0.047)	(0.007)	(0.028)	(0.014)	(0.042)
	$t_c = 1$	-0.035	-0.010	0.031	0.009	-0.016 +	0.025	0.056
		(0.049)	(0.011)	(0.064)	(0.006)	(0.010)	(0.021)	(0.085)
	$t_c = 2$	-0.059	-0.022	0.072	0.009 +	-0.045 *	-0.022	0.079
		(0.036)	(0.013)	(0.046)	(0.005)	(0.022)	(0.030)	(0.053)
		149,002	*149,002*	*149,002*	*149,002*	*149,002*	*149,002*	*149,002*

NOTES: Estimates come from a regression taking the same form as model 1, except that instead including student demographics as control variables, each characteristic is instead on the left-hand side of the equation. Samples include any student in either math or ELA analyses. Cell entries include point estimates, robust standard errors (in parentheses), sample sizes (in italics), and approximate *p*-values (+ $p<0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$).

Appendix B. Interview Protocol

1. What is your current position in BPS (MA DESE) and what has your role been over the past few years?

2. What has been your role in relationship to Blueprint?
    a. What role did you play, if any, in the consideration or selection of Blueprint as a school operator or partner in BPS?
    b. In what ways, if any, have you interacted with Blueprint central staff or schools?
3. In your opinion, how have things gone with Blueprint over the past several years?
    a. Probe for changes over time
    b. What has Blueprint done well? What could Blueprint have done better?
4. In your opinion, what is the nature of the relationship between the state, BPS, Blueprint, and BPS schools run by Blueprint?
    a. Has Blueprint played any role in bringing these groups together productively? If so, what role?
5. I'd like to talk a bit about the state and district systems and policies that you think have helped Blueprint or provided barriers to them in implementing their model in BPS.
    a. What state/district systems and policies have supported Blueprint?
    b. Which systems and policies have constructed barriers that constrained Blueprint from implementing their model as they would like?
    c. Are there any systems or policies could Blueprint have leveraged that it did not? Which ones? How could they have used them?
    d. What restrictive policies or practices could Blueprint have used its autonomies to push against more?
6. How are state and district policymakers using the data that Blueprint collects during the site visits?
    a. Is this information useful?
7. Do you see any important differences between working with Blueprint as a school operator and Blueprint as a school partner?
    a. How much does the degree of Blueprint's influence differ across these two models?

8. What could Blueprint do to improve how it operates and works in partnership with BPS schools?

9. Is there anything else we did not ask about that we should know about?