The Role of School Climate in School Turnaround

July 2023

Education Policy Innovation Collaborative

COLLEGE OF EDUCATION | MICHIGAN STATE UNIVERSITY

236 ERICKSON HALL, 620 FARM LANE, EAST LANSING, MI 48824 | www.EPICedpolicy.org

ABSTRACT

School climate is critical to school effectiveness, but there is limited large-scale data available to examine the magnitude and nature of the relationship between school climate and school improvement. Drawing on statewide administrative data linked with unique teacher survey data in Michigan, we examine whether school climate appeared to play a role in the effects of a state-level school turnaround intervention. Using comparative interrupted time series models and descriptive mediation analysis, we find that students in schools with more positive school climate appeared to fare better than their peers in schools with less positive climate. Certain elements of climate—relational trust and school leadership—also mediated the effect of turnaround on student achievement. Our findings have implications for school improvement planning, for the design of evaluations of school turnaround initiatives, and for data collection by states aiming to improve their lowest performing schools.

DISCLAIMER

The Education Policy Innovation Collaborative (EPIC) at Michigan State University is an independent, non-partisan research center that operates as the strategic research partner to the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). EPIC conducts original research using a variety of methods that include advanced statistical modeling, representative surveys, interviews, and case study approaches. Results, information, and opinions solely represent the author(s) and are not endorsed by, nor reflect the views or positions of, grantors, MDE and CEPI, or any employee thereof. All errors are our own.

JULY 2023

The Role of School Climate in School Turnaround

AUTHORS

Erica Harbatkin, *EPIC, Florida State University*

Jason Burns, *EPIC, Seton Hall University*

Samantha Cullum, *EPIC, Michigan State University*

Contemporary school turnaround calls for improving student achievement in part by transforming the climate in which low-performing schools operate. This is because positive school climate and teacher working conditions may contribute to meaningful and sustainable school improvement (Cucchiara et al., 2015; Strunk, Marsh, Bush-Mecenas, et al., 2016; Thompson et al., 2016). However, a school turnround intervention that is not implemented well could lead to lower teacher perceptions of some school climate measures, increasing teacher turnover and undermining improvement efforts (Heissel & Ladd, 2018). However, almost no research to date has quantitively examined the extent to which school climate appears to mediate or moderate the effects of turnaround (for an exception, see Pham (2022), which found that one element of school climate—teacher collaboration—played an important role in the positive effects of turnaround in Tennessee).

However, recent federal school improvement policies have not explicitly addressed school climate. Instead, they have focused on other factors likely related to climate such as staffing (e.g., School Improvement Grants and Race to the Top) and tailoring interventions to local context identified in a needs assessment (e.g., Every Student Succeeds Act, or ESSA).

There are multiple reasons we might expect school climate to play a role in successful turnaround. First, a leading theory of action for school turnaround calls for systems-level change to disrupt the processes underlying low performance and create the educational infrastructure needed to sustain improvements (Meyers & Smylie, 2017; Peurach & Neumerski, 2015). Because school climate is an critical component of this educational infrastructure, turnaround interventions can improve student outcomes by first improving school climate. Second, meaningful change to school operations requires buy-in from staff (Datnow & Stringfield, 2000), which may be less likely to occur in the context of negative school climate (Trujillo & Renée, 2015). In other words, successful turnaround may require either a climate that is conducive to reform or for the intervention itself to successfully build buy-in (Player et al., 2014). Third, building buy-in requires a strong school leader who can rally staff around school improvement goals (Duke & Salmonowicz, 2010; Redding & Corbett, 2018).

Together, the research therefore points to two distinct but overlapping ways in which climate may explain a turnaround intervention's effectiveness. First, schools with more positive climate at the outset of reform may be better able to enact change under turnaround. Second, a turnaround model that explicitly targets a school's climate may improve school outcomes through improvements to climate, for example by setting a clear mission, clarifying staff roles, and prioritizing working conditions.

In this paper, we examine the extent to which elements of school climate may have contributed to the effects of a turnaround intervention in Michigan. Specifically, we ask:

(1) To what extent did the effect of turnaround differ across schools with stronger or weaker school climate?
(2) To what extent did school climate mediate the relationship between turnaround and student outcomes?

The first question investigates whether schools with more positive school climate fared better than schools with more negative school climate. The second is a mediation question—do school climate measures explain the observed relationship between turnaround and student achievement?

We examine these questions in the context of the Michigan Partnership Model of School and District Turnaround. Initially implemented for the first cohort of schools in fall 2017 and the second in fall 2018, the Partnership Model aims to foster improvement in district-level systems to support intermediate school outcomes such as increased educator retention and quality, and ultimately improve student achievement. Underlying these improvements is the Partnership Agreement, an improvement plan outlining improvement goals, strategies for achieving goals (including efforts to improve climate), and accountability measures for failing to achieve them.

LITERATURE REVIEW

Recent years have seen a profusion of empirical research on school turnaround, with small positive effects on average, though there has been wide variation across contexts (for reviews, see Redding & Nguyen, 2020; Schueler et al., 2020). Some empirical studies have found positive effects of school turnaround on student outcomes (Bonilla & Dee, 2017; Carlson & Lavertu, 2018; Dee, 2012; Gandhi et al., 2018; Henry et al., 2015; Henry & Guthrie, 2019; Pham et al., 2020; Schueler et al., 2017; Strunk, Marsh, Hashim, Bush-Mecenas, et al., 2016; Zimmer et al., 2017), while others have found no effects or even negative outcomes, with some identifying heterogeneity within interventions such as some cohorts with null or negative effects and some with positive effects (Burns et al., 2023; Dougherty & Weiner, 2019; Dragoset et al., 2019; Heissel & Ladd, 2018; Henry & Guthrie, 2019; Henry & Harbatkin, 2020; Pham et al., 2020; Strunk, Marsh, Hashim, Bush-Mecenas, et al., 2016, 2016; Zimmer et al., 2017).

Research on the effects of turnaround on elements of climate has highlighted both positive and negative effects of turnaround. For example, some studies found that turnaround increased meaningful teacher collaboration, thus improving instructional practices and teacher efficacy (Andreoli et al., 2020; Pham, 2022; Strunk, Marsh, Hashim, & Bush-Mecenas, 2016). Studies also found that turnaround led to a clearer, better-defined vision from school leadership and more collaboration across school leaders within a district, thus strengthening teacher expectations for students, academic rigor, and school culture (Meyers, 2020; Weixler et al., 2018). One study found the reform created distributed leadership, supporting the longer term sustainability of the reform (Patterson et al., 2021).

In contrast, other studies found that being in a school identified for reform increased teacher stress, demoralization, and workload (Collet, 2017; Cucchiara et al., 2015; Fried, 2020; Heissel & Ladd, 2018; Henry & Harbatkin, 2020; Nolan, 2018; Quartz et al., 2020; Rice & Croninger, 2005). Some found that teachers felt a loss of voice and autonomy, and in turn some engaged in satisficing behaviors to comply with reform requirements (Welsh & Williams, 2018). Teachers in some cases reported loss of trust in programs, leaders, and in colleagues after existing collaborative networks broke down due to reform-induced teacher replacements (Lenhoff & Ulmer, 2016; Malen et al., 2002; Rice & Malen, 2003). Additionally, while improving school leadership in some contexts, turnaround introduced new ambiguities and stress into the job for other school leaders—potentially reducing principal effectiveness (Daly et al., 2011; Hamilton et al., 2014; Yoon et al., 2020).

The research around readiness for reform is more theoretical in nature, but there is reason to believe that climate could affect a school's readiness. For example, a school with strong leadership and a cohesive staff would have greater capacity to immediately translate a reform model into meaningful improvement, while a school with weak relationships across staff would not be able to leverage turnaround

resources to improve student outcomes at all—or at least not without major restructuring (Slavin, 1998). District readiness for reform, including leadership, support infrastructure, talent management, and instructional infrastructure, also matters given that the district often acts as a primary catalyst for improvement (Player et al., 2014). Finally, because staffing is central to school improvement, staffing climate is critical to successful reform. Specifically, a turnaround model aimed at improving student achievement by hiring more effective teachers would require a local context in which there are sufficient effective teachers willing to work in the turnaround school, while a model aimed at improving the instruction of existing teachers would require an existing staff of willing and engaged teachers who are bought into the reform (Harbatkin, 2022; Henry & Harbatkin, 2020; Malen et al., 2002).

Together, the research on school climate and school improvement underscore that several dimensions of climate may affect or be affected by turnaround interventions. This matters because there is some limited evidence that schools with stronger climate fare better under turnaround than those with weaker climate. For example, a study examining school improvement in Chicago found that schools with strong climate components such as school leadership, community and parent involvement, safety, professional capacity, and academic climate were more likely to make academic improvements than schools that were weak in those areas (Bryk et al., 2010). School climate also appeared to be a mechanism for positive change in qualitative and descriptive studies on whole school improvement, the precursor to federal school turnaround policy (Bulach & Malone, 1994; Cohen, 2006; Dellar, 1998; Peurach & Neumerski, 2015; Thapa et al., 2013).

While there are several elements of school climate, we focus on five—school safety and discipline, academic climate, staffing climate, school leadership, and relational trust—that are largely aligned with elements of the first three of Thapa and colleagues' (2013) dimensions of school climate (safety, relationships, and teaching and learning). An unsafe school environment can introduce barriers to teaching and learning that undermine school improvement efforts (Kutsyuruba et al., 2015; Reynolds et al., 2014). Meanwhile, academic climate is critical to turnaround because successful school improvement requires strong instructional practices, effective professional development, and supports that are well-aligned with the reform (Cucchiara et al., 2015; Leithwood et al., 2010; Meyers & Hambrick Hitt, 2017; Peurach & Neumerski, 2015). A large literature underlines the importance of a stable and effective staff to effective turnaround; indeed, high teacher turnover may in fact suppress the effects of a turnaround intervention that would have otherwise improved student outcomes (Henry et al., 2020; Papay & Hannon, 2018; Player & Katz, 2016; Sun et al., 2017).

One uniquely important element of staffing climate is leadership. A large qualitative literature points to school leadership as a critical component of turnaround. Qualitative and case study research suggests the existence of so-called turnaround principals, those with a distinct set of skills and traits needed to turn around low-

performing schools (Duke, 2004; Duke & Salmonowicz, 2010; Finnigan & Stewart, 2009; Harris, 2002; Jacobson et al., 2005; Meyers & Hambrick Hitt, 2017). A handful of these studies, focused on whole-school reform efforts before NCLB, suggest that turnaround leadership is associated with improvements to school climate, highlighting that school leadership plays an important role in promoting positive school climate and ultimately in successful school turnaround (Finnigan & Stewart, 2009; Huberman et al., 2011; Leithwood & Jantzi, 1990; Meyers & Hitt, 2017). School leaders also determine working conditions and the professional environment in which teachers operate, which can either promote or stymie teacher development (Kraft & Papay, 2014)—another critical element of school turnaround.

Finally, relational trust characterizes the extent to which stakeholders across the school system (students, staff, leadership, parents, community) trust one another and in the process build shared values, perspectives, and interests (Bryk & Schneider, 2002; Forsyth et al., 2006; Torres, 2016). Relational trust is essential to a positive school climate, and in turn to improved student achievement (Bryk & Schneider, 2002; Mayger & Hochbein, 2021)

We draw on this past research to explicitly test whether elements of school climate appear to play a role in the effectiveness of a turnaround intervention in Michigan. We contribute to the literature in two ways. First, we leverage teacher survey data and statewide administrative data to explicitly examine whether dimensions of climate appear to contribute to heterogeneous effects of a turnaround intervention. In doing so, we provide evidence on an understudied element that may contribute to the heterogeneity of effects across turnaround interventions. Second, while two studies have formally tested variables such as staff turnover, chronic absenteeism, and staff collaboration as mediators of school turnaround (Henry et al., 2020; Pham, 2022), this is the first study we know of that tests whether multiple measures of school climate formally mediate the relationship between a turnaround intervention and student achievement.

MICHIGAN'S PARTNERSHIP MODEL FOR SCHOOL AND DISTRICT IMPROVEMENT

The Partnership Model is intended to boost student outcomes in Michigan's lowest performing schools. Initially implemented in spring 2017 and subsequently included as part of the state's plan under ESSA, the Partnership Model involves cooperation among local, regional, and state-level stakeholders to identify the challenges faced by struggling schools and then develop and implement a plan to turn them around. Several important characteristics distinguish the Partnership Model from previous generations of turnaround. One is that while individual schools are identified for Partnership, the Partnership Model is a district-level intervention as the districts that operate low-performing schools play a central role in crafting and implementing the

turnaround plan for identified schools. And second, the Partnership Model is built around flexibility, providing identified schools and their districts significant latitude in setting their turnaround goals and identifying the strategies to meet those goals.

Two cohorts of Partnership schools and districts were identified and received supports prior to the COVID-19 pandemic. The first cohort, identified in spring 2017, included the schools that had been identified as Priority schools for three consecutive years under Michigan's No Child Left Behind (NCLB) waiver reform. The second cohort was identified across two rounds during the 2017-18 school year. Schools in the first of these two rounds were identified in fall 2017 after they had been identified previously as a Priority school and displayed decreased student achievement over time. The second round of Cohort 2 Partnership schools was made up of Michigan's first set of Comprehensive Support and Improvement (CSI) schools under the Every Student Succeeds Act (ESSA), the schools identified in spring 2018 as the bottom 5% of schools statewide on Michigan's academic performance index. For both cohorts, implementation of their Partnership reforms began the school year following identification: 2017-18 for Cohort 1 and 2018-19 for Cohort 2. In total, 119 schools across 36 districts entered Partnership across the two cohorts. There is evidence that on average, the Partnership Model improved student outcomes, though there was heterogeneity by cohort and student characteristics (Burns et al., 2023).

The implied theory of change for the Partnership Model is illustrated in Figure 1. Beginning with the top of this figure, once a school is identified for Partnership, its district is charged with drafting a Partnership Agreement, a three-year turnaround plan that identifies the challenges faced by that school, the goals for that school, strategies aligned to those goals, and the consequences for failing to meet goals. To aid in the development of the Partnership Agreement, the district receives supports from several stakeholders. Among them are the Michigan Department of Education (MDE), which provides the district a liaison to work with district leadership as needed as well as grants to support turnaround work. Districts are encouraged to work with local community organizations to support their turnaround work, though this was de-emphasized over time. Partnership schools and districts also have access to supports from their intermediate school district (ISD), a regional organization that can provide trainings, coaching, and professional development for teachers and leaders. Importantly, a central tenet of the Partnership Model is that districts design a Partnership Agreement aligned with the context and needs of their identified school(s), and as such, the district has discretion in how its constellation of supports is mobilized to boost student outcomes.

Using the available supports, Partnership districts develop a Partnership Agreement that analyzes the strengths and weaknesses of the district, lays out goals to be met over a 36-month improvement cycle, strategies to achieve those goals, consequences for failing to meet their goals, and the roles of partners in the district's reform efforts. As shown in the lower panel of Figure 1, the Partnership Model holds that a focus on

core school- and district-level systems that support student outcomes, such as human resources, curriculum, and instructional systems, will help to make the district more effective and efficient in the intermediate-term and in the long-term yield improved student outcomes such as higher student achievement and attendance.

FIGURE 1

Although the Partnership Model allowed for variation across districts as they worked to craft an approach to turnaround that was aligned with their context, Partnership districts identified a similar set of challenges to overcome in designing their Partnership Agreements. The most prominent among these is difficulty staffing their schools with high quality, and sometimes even simply qualified, teachers and leaders, reporting high rates of turnover and difficulty in hiring new teachers (Burns et al., 2023). These challenges suggest the presence of acute issues around climate inside Partnership schools, which may mediate or moderate the impact of the Partnership Model through their effect on the educators who ultimately carry out needed reforms in Partnership districts and schools.

DATA AND SAMPLE

To examine whether and to what extent school climate plays a role in successful turnaround, we draw from statewide administrative data and teacher survey data. Specifically, we use seven years of statewide administrative data on students and teachers from 2013-14 through 2018-19 provided by the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). We pair the administrative data with two years of teacher survey data collected as part of a larger evaluation of the Partnership Model in fall 2018 and 2019. We administered surveys to all teachers in both Partnership and non-Partnership schools in Partnership districts. In total, the data include 2,719 teacher survey responses from fall 2018 (38.3% response rate) and 3,386 teacher survey responses from fall 2019 (49.2% response rate). We use these survey responses to develop constructs related to climate and then collapse teacher responses to the school level to generate school-level measures of climate. Our two research questions employ different analytic strategies and therefore draw on different samples.

The first question uses a comparative interrupted time series (CITS) design to examine the extent to which students in terciles of Partnership schools based on our climate constructs fare differently relative to a group of near-selected comparison schools over multiple years. The second question uses descriptive regressions to explore whether climate mediate the relationship between Partnership and student achievement in a single year.

CITS Sample

The analytic sample for the CITS analysis includes students in Partnership schools (treatment) and students in near-selected schools that were not identified for Partnership (comparison). This study draws on data from the first three rounds of Partnership identification rounds, which we analyze as two implementation cohorts. The first round of Partnership schools was identified in the 2016-17 school year based on being designated Priority schools (i.e., the bottom 5% in the state) for three straight years. We characterize this group of schools as Cohort 1, which began implementing Partnership in fall 2017. We construct the comparison group for this cohort as students in 2015-16 Priority schools that were not identified as Partnership in any round. These are schools that were in the bottom 5% in 2015-16 but not in each of the two prior years (which would have placed them in Cohort 1) and were not low performing in 2016-17 (which would have placed them in round 2 or 3).

The state identified round 2 and 3 schools in fall and spring of 2017-18, respectively, using test score data from spring 2016 and spring 2017. Round 2 schools were those that were not selected in Round 1 and were in the bottom 5% in 2015-16 and had continued low performance in 2016-17. Round 3 makes up the state's first round of Comprehensive Support and Improvement (CSI) schools under the Every Student Succeeds Act (ESSA), which were the bottom 5% of schools on the state index system in 2016-17. Because both round 2 and 3 schools began turnaround in fall 2018, we combine them into a single Cohort 2. The Cohort 2 comparison group includes all schools in the bottom 10% in 2016-17 that were not designated as Partnership in any of the three rounds.

We restrict the CITS analytic sample to students in treated schools meeting a minimum survey response rate threshold of 20% or at least three teacher responses for all items across all climate measures. Because we do not have survey data for comparison schools, this restriction does not apply in the comparison groups. In total, the Cohort 1 analysis includes about 100,000 student-year observations and the Cohort 2 analysis includes about 185,000 student-year observations.

Table 1 provides descriptive statistics for each cohort, the subset of each cohort meeting minimum response rate thresholds, the cohort's comparison group, and the rest of the state. Panel A shows student-level means, highlighting three takeaways. First, both cohorts of Partnership schools, shown in Columns 1 and 5, serve a very large share of economically disadvantaged and Black students relative to the rest of the state. Second, student achievement in Partnership schools is about one standard deviation below the state mean for both math and ELA. Third, comparison schools, shown in Columns 3 and 6, serve a population of students that are similar in terms of economic disadvantage but less Black than Partnership schools and slightly less low performing. Fourth, the analytic sample, shown in Columns 2 and 4, is similar on student-level observables to the full sample of Partnership schools. Panel B shows that the analytic samples of schools have higher enrollment on average than the full

sample, which is unsurprising given our minimum threshold rule because larger schools could have lower response rates and still reach at least three responses.

TABLE 1

Though there is variation across climate measures, students attending schools in the bottom tercile on many measures tend to be slightly lower performing at baseline. There are otherwise no clear patterns of observable baseline differences between terciles (Appendix Table A-1).

Mediation Sample

Because the mediation analysis requires climate constructs for all schools across all years of analysis, the mediation sample is subject to two restrictions that the moderation sample is not. First, the sample includes only schools in Partnership districts rather than all treated schools and a comparison group of near-selected schools for that year—which are located in both Partnership and non-Partnership districts. This limitation stems from survey administration; we administered surveys to both treated and untreated schools in Partnership districts, but not to teachers in any non-Partnership districts. The mediation sample uses Partnership schools as the treated group and non-Partnership schools in Partnership districts as the untreated group. The mediation comparison group therefore is fundamentally different from the CITS comparison group; because Partnership is in part a district-level treatment, non-Partnership schools in Partnership districts received some dose of treatment, though it was less targeted than the Partnership school treatment. To the extent that they did improve from Partnership-related district-level improvements, the estimated relationship between Partnership and student achievement in the mediation models would be attenuated (Burns et al., 2023). The role of school climate, however, should not be biased.

Second, the sample includes data only from the 2017-18 and 2018-19 school years. Because we administered surveys beginning in the 2018-19 school year, we do not observe climate measures for the pretreatment years. We therefore can examine mediating factors for the second year of Partnership for Cohort 1 and the first year for Cohort 2.[1]

We again limit the sample to just schools meeting our minimum response threshold, though here we aim to maximize the sample by setting the threshold at the construct level rather than across all constructs. We also restrict the sample of non-Partnership schools using the same response rate threshold. Of 34 Cohort 1 schools, for a given construct, 91-100% meet the minimum threshold. Of 83 Cohort 2 schools, 66-89% meet the minimum threshold. Of 259 non-Partnership schools in Partnership districts, 68-69% meet the minimum threshold.

EMPIRICAL STRATEGY

In this section, we begin by describing our approach to factor analysis to create measures of climate. We use these measures to create high, mid, and low climate groups for the CITS analysis, and to test as mediators in the mediation analysis. We then describe the CITS analysis to examine whether students in schools with more positive climate measures fared better under turnaround than their peers in schools with more negative climate measures. This analysis draws from the climate constructs to examine whether there are heterogeneous effects by school climate. Finally, we describe the mediation analysis, in which we empirically test whether school climate mediates the observed relationship between Partnership and student outcomes.

Factor Analysis

We use factor analysis on two years of teacher survey data to develop climate constructs, focusing on questions about school climate asked in both years of the teacher survey. We begin with an exploratory factor analysis (EFA), including a parallel analysis (Horn, 1965) to identify the appropriate number of factors within each relevant item set using the *paran* Stata package (Dinno, 2009). In cases where the parallel analysis points to multiple factors, we rotate the factors using the varimax criterion in order to identify a set of factors with meaningful interpretations. We then run confirmatory factor analyses (CFA) replicating the structure identified through the EFA, generate factor loadings, and collapse all loadings to the school-by-year level. Most questions asked teachers about climate in the current school year. One question, about school leadership, asked teachers to rate the effectiveness of their principal in the *prior* school year. For this question, we apply responses to the prior year.

Using these school-level measures, we generate terciles of Cohort 1 and Cohort 2 treated schools, respectively, based on their mean factor scores for each construct. Specifically, we create three groups (high, mid, and low climate) in each of the two cohorts for each of the five constructs that emerged from the factor analysis. In total, each treated school is assigned to a total of five different groups—the high, mid, or low group for each of the five climate measures. We create these groupings using fall 2018 measures (fall 2019 for the school leadership measure) because most of our outcomes were measured in 2018-19. Thus, data collected in fall 2018 provide the most relevant measures of climate in the outcome year. Because the school mean factor scores represent teacher average perceptions of the school's climate, these terciles provide rough groupings of schools by the strength of their climate.

CITS Analysis

To estimate separate effects of Partnership for students in high, mid, and low schools on each climate measure, we estimate a series of CITS models. This approach allows us to calculate separate effects for each of the three groups and then test the

equivalence of groups within models. There is strong evidence that CITS models have the capacity to estimate effects of social policy interventions with limited bias, especially in first years of an intervention (Coopersmith et al., 2022; Sims et al., 2022). The primary benefit of this approach is it allows us to estimate heterogeneous effects of Partnership by climate scores within a difference-in-differences framework relative to a comparison group of students in similarly low-performing schools throughout the state. The primary limitation for our purposes is that it does not directly test climate as a mediating factor of turnaround. We therefore undertake a second set of analyses in a traditional mediation framework, which we describe in the next section. Another limitation is in timing of the climate measures. While Partnership was first implemented in 2017-18 in Cohort 1 and 2018-19 in Cohort 2, our survey measures were collected in fall 2018—after Partnership began. To that end, we cannot parse the extent to which the intervention affected climate from the extent to which schools with strong climate prior to the intervention fared better. Schools that had strong climate prior to the intervention will look the same in our analysis to schools that improved climate as a result of the intervention. Thus, our estimates reflect the combined effect of Partnership on climate and student achievement.

Beginning with the terciles described above, we include three treatment groups in each model—high climate × treatment, mid climate × treatment, and low climate × treatment. We then estimate separate CITS models for each construct grouping in each of the two cohorts to estimate the effect of Partnership for students in schools in the high, mid, and low groups, respectively. For example, the Cohort 1 model takes the form

$$TestScore_{ist} = \beta_0 + \beta_1 Year + \beta_2 (Year \times Partnership_{st}) + \sum_{t=2018}^{2019} \sum_{g=low}^{high} Year_t \times Partnership_{Group} + \beta_3 TestScore_{ist-1} + \boldsymbol{X'}_{is}\theta + \psi_s + \phi_t + \varepsilon_{ist}\,, \quad (1)$$

predicting the test score for student *i* in school *s* in year *t*. *Year* is a linear year variable centered at the identification year for Cohort 1 (2016-17), and $Year \times Partnership$ is an interaction between the linear year variable and a binary indicator that takes the value of 1 for Partnership schools, with β_2 representing the deviation from the year trend for treated schools. We then interact an indicator for each treated year (i.e., 2017-18 and 2018-19 for Cohort 1, represented in the summation operator above as the spring year) with an indicator for each climate tercile, where all comparison schools are coded as zero. Because we also include year fixed effects, ϕ, the coefficients on these interactions represent the estimated effect for treated schools in the low, mid, and high climate terciles, respectively, relative to the average comparison school. $\boldsymbol{X'}$ is a vector of baseline school covariates interacted with a linear year trend and time-varying student covariates, ψ is a school fixed effect, and ε is an idiosyncratic error term clustered at the school level. Student covariates include grade level fixed effects and indicators denoting whether the student was economically disadvantaged, had English learner status, and had a special education designation, respectively. School-level covariates include mutually exclusive race/ethnicity percentages (Black; Hispanic

or Latino/a/x; Asian, multiple races, or other) with White as the reference category; economically disadvantaged, English learner, and special education percentage; and logged school enrollment.

The CITS model assumes that in the absence of Partnership, comparison schools would follow the linear trend they were following prior to the intervention (estimated as β_1), and Partnership schools would continue their own linear trend (the linear combination of β_1+ β_2), conditional on covariates. Because we allow these flexible trends, we do not need parallel trends prior to the intervention, but rather to assume that each respective group would not deviate from its own pretreatment trend in the absence of the intervention.

The model for Cohort 2 follows the same format as Equation 1 except we center the linear year trend at the Cohort 2 identification year of 2017-18, we only estimate one year of treatment effects (2018-19), and we trim the pretreatment time period to begin with 2014-15 rather than 2013-14. We trim the panel because the state changed testing instruments in 2014-15 from the Michigan Educational Assessment Program (MEAP) to the Michigan Student Test of Educational Progress (M-STEP). While we standardized scores by grade level and year, standardization would not account for differences across the treated and comparison groups. We retain the 2013-14 scores for Cohort 1 in order to retain three years of pretreatment data but caution that the 2013-14 estimates for Cohort 1 could be affected by test differences.

Mediation Analysis

We follow the Baron & Kenny (1986) approach to mediation by estimating three models shown visually in Figure 2. Model 2 estimates the effect of Partnership on the measure of school climate, Model 3 estimates the effect of Partnership on the more distal outcome (i.e., student test score) and Model 4 repeats Model 3 with the school-level culture or climate measure included as a covariate. To the extent that the culture or climate measure mediated the effect of Partnership on the outcome, the Model 4 estimate would be lower than the Model 3 estimate.

FIGURE 2

We show these models, estimated using a series of descriptive ordinary least squares regressions, below. Model 2 estimates the relationship between school climate and Partnership, with β_1 representing path *a*. Model 3 estimates the relationship between the outcome (math score in this example) and Partnership, with β_1 representing path *c*. Model 4 adds the climate measure to Model 3 as a covariate to estimate the mediated relationship between Partnership and the outcome, with β_1 representing *c'*.

$$Climate_s = \beta_0 + \beta_1 Partnership_s + \beta_2 LaggedScore_i + \boldsymbol{X'}_{ist}\theta + \psi_s + \varepsilon_{is} \quad (2)$$

$$TestScore_s = \beta_0 + \beta_1 Partnership_s + \beta_2 LaggedScore_i + \boldsymbol{X'}_{ist}\theta + \psi_s + \varepsilon_{is} \quad (3)$$

$$TestScore_s = \beta_0 + \beta_1 Partnership_s + \beta_2 LaggedScore_i + \beta_3 Climate_s + \boldsymbol{X'}_{ist}\theta + \psi_s + \varepsilon_{is} \quad (4)$$

We include in the mediation models the same set of covariates as in the CITS models, though here we use covariates measured at time *t*. We do not include school or year fixed effects because we only observe a single year of data (2018-19).

The benefit of this approach is it allows us to examine the relationship between Partnership and school climate and to test whether and to what extent climate mediates the relationship between Partnership and student outcomes. There are two main limitations. First, we only have survey data from schools in Partnership districts. Because both Partnership schools and non-Partnership schools within Partnership districts received some dose of treatment, the estimated relationship will be attenuated by the reduced contrast between the treatment and comparison schools. Second, our survey data cover just two school years rather than the full panel of years required for an event study. Thus, we can only estimate the mediation models using one treated year for each cohort. The results therefore should not be interpreted as causal. Instead, they provide an estimate of the extent to which students in Partnership schools fared better than students in non-Partnership schools in the observed year, and whether those differences appear to have been mediated by school climate factors.

RESULTS

In this section, we provide a summary of the factor analysis and mean comparisons of the factor scores across teachers in treated and comparison schools. We then describe the CITS findings followed by results from the mediation analysis.

Factor analysis

The factor analysis yields five factors, which we classify as academic climate, school safety and discipline, relational trust, staffing climate, and school leadership. The first four factors emerged from items from two questions—one asking teachers to grade their school from A-F in a variety of areas and one asking teachers the extent to which they agree with a series of statements. Table 2 provides the factor loadings and uniqueness (ψ) for each variable in each factor. Cronbach's α range from 0.79 for relational trust to 0.96 for school leadership.

TABLE 2

Table 3 provides means and standard deviations for each of the five constructs by treatment condition and study year. Because we create these measures to have a mean of zero and a standard deviation of one across the full sample of Partnership district educators across the study period, negative values can be interpreted as below the mean for Partnership districts and positive values can be interpreted as above the mean for Partnership districts. The first five columns show values for 2018-19, the first year the survey was administered. The 2018-19 school year represents the second year of Partnership implementation for Cohort 1 and the first for Cohort 2. Row 1

shows substantial variation in academic climate across cohorts, with Cohort 1 right around the sample mean and Cohort 2 about a quarter of a standard deviation below the mean. In sum, academic climate was higher in Cohort 1 than Cohort 2 schools and similar to comparison schools in 2018-19. On school safety and discipline, relational trust, and staffing climate, both cohorts of Partnership schools are lower than comparison schools in 2018-19 in particular.

TABLE 3

On school leadership, comparison schools are lower than Cohort 1 schools in 2018-19 and both cohorts in 2019-20. The school leadership measure needs to be interpreted differently from the others because the question asks about teacher perceptions of their school leader in the *prior* school year.[2] To that end, we highlight two caveats. First, the 2018-19 measure applies to the 2017-18 school year (i.e., year 1 for Cohort 1 and the identification year for Cohort 2) and the 2019-20 measure applies to the 2018-19 school year (i.e., year 2 for Cohort 1 and year 1 for Cohort 2). Second, the question is posed only to returning teachers; as such, the measures may be subject to survivorship bias. The school leadership score for Cohort 1 is higher than both Cohort 2 and the comparison schools in 2017-18. Cohort 2 is the lowest group in 2017-18, the year the schools were identified for Partnership. Cohort 1 remains similar in 2018-19, its second year of services, while Cohort 2 rebounds in its first year.

For the other four measures, the 2019-20 values show the descriptive change in these climate measures during the intervention. We find that Cohort 2 in particular improves from 2018-19 to 2019-20 on all five measures. Cohort 1, by comparison, remains stable over each of the two years, though we note that we do not observe these measures over the same relative time period. In particular, fall 2018 to fall 2019 roughly coincides with the first year of the intervention for Cohort 2 and the second for Cohort 1. We cannot say from our data whether Cohort 1 made similar improvements over the first year of their intervention. However, this year-to-year difference in Cohort 2 does provide some evidence that some Partnership schools improved on these measures of climate during the intervention period.

CITS findings

The CITS results are provided in Table 4, with math in Panel A and ELA in Panel B. Each panel shows Cohort 1 in Columns 1-5 and Cohort 2 in Columns 6-10. The estimates of interest are in the rows labeled Partnership 2017-18 × Low, Partnership 2017-18 × Mid, and Partnership × 2017-18 x High (Cohort 1 only), and Partnership 2018-19 × Low, Partnership 2018-19 × Mid, and Partnership 2018-19 × High (both cohorts). These coefficient estimates represent the estimated effect of Partnership for treated schools in the low, mid, and high terciles, respectively, relative to all comparison schools in the observed year.

TABLE 4

In Cohort 1 in 2018-19, we find suggestive evidence of positive effects in math for the high tercile schools in school safety and discipline, relational trust, and staffing climate. Further, *F*-tests show that the coefficients on the high and low terciles in these three models are significantly different from one another ($p<0.05$), providing evidence that these dimensions of climate are important for school turnaround. We do not observe the same patterns in 2017-18, though this may be unsurprising because the climate constructs were measured in fall 2018 after the 2017-18 school year concluded. While tercile groupings are moderately correlated from year to year (0.35 to 0.50 for the first four measures and 0.28 to 0.40 for school leadership, which asks about the prior year and is therefore less precise), noisier terciles will mask differences by true climate levels. The implications of this imprecision will be greater in Cohort 1 because these terciles include only nine schools each (Cohort 2 terciles, in contrast, contain 15-17 schools each, depending on the construct). As a result, any movement around the margins of the tercile groupings will have a large impact on the treated group and the resulting effect estimate for that group.

Cohort 2 shows clearer patterns, especially in math (Panel A). The most pronounced pattern is that there are positive effects in math for the high tercile on each of the five dimensions of climate. These positive effects range from 0.12 standard deviations (SDU) for students in the high school leadership tercile schools to 0.17 SDUs for students in the high school safety and discipline tercile schools. Estimated effects are smaller in ELA, shown in Panel B, but are positive and significant for the high terciles in relational trust, staffing climate, and school leadership.

By contrast, we do not detect effects for students in the low-tercile Cohort 2 schools in any of the five climate dimensions. *F*-tests find that the low and high tercile coefficients are significantly different from one another for academic climate, school safety and discipline, relational trust, and staffing climate ($p<0.01$), and marginally significant for school leadership ($p<0.10$).

Estimated effects for students in mid-tercile Cohort 2 schools are generally between the low and high tercile estimates, with a positive and significant estimate in math for students in mid-tercile school leadership schools. Figure 3 illustrates these patterns. In Cohort 2 math in particular, the coefficient estimates follow an upward trend with the lowest estimates for students in the low-tercile schools, higher and in some cases positive and significant estimates for students in the mid-tercile schools, and the strongest positive estimates for students in the high-tercile schools. In Cohort 1, we see similar patterns for school safety and discipline, relational trust, and staffing climate in 2018-19, the year these variables were measured.

FIGURE 3

Together, these results suggest that schools with stronger climate fared better than schools with weaker climate. This relationship is strongest where we have more precise climate measures (in the 2018-19 when we measure climate; in Cohort 2,

which has larger terciles; and on the first four constructs, which we measure in the survey year rather than asking retrospectively to returning teachers only as in the case of school leadership).

Mediation Analysis

We turn next to results from the descriptive mediation analysis, beginning with Cohort 1 and then moving to Cohort 2. Table 5 provides the results for Cohort 1, with Panel A showing path *a* and Panels B and C providing the estimates for paths *b*, *c*, and *c'* for math and ELA, respectively. The first-order requirement in mediation is that Partnership has a relationship to the factor score itself. Therefore, the potential mediators are those with significant effects on path *a* in Panel A. In Cohort 1, we find that only two factors are potential mediators by conventional standards—relational trust and school leadership. Specifically, being in a Partnership school is associated with a 0.23 standard deviation (SDU) increase in relational trust a 0.27 SDU increase in school leadership. We note that there is a substantively strong estimate, 0.21 SDU, on academic climate as well.

To examine whether these mediate the relationship between Partnership and student achievement, we turn to Panels B and C. The odd-numbered columns provide the estimates for path *c*, or the direct (unmediated) relationship between Partnership and student achievement within the sample of schools meeting minimum survey threshold requirements for each construct. The even-numbered columns provide the estimate for paths *c'* (the coefficient on Partnership) and *b* (the coefficient on the climate factor score). As expected, higher scores on each climate factor are associated with higher test scores, shown by the positive and significant coefficients on each of the factor scores.

The extent to which the climate construct mediates the relationship between Partnership and student achievement is represented in the difference between the *c* and *c'* estimates on Partnership. For example, the estimate on the *c* path in the relational trust model (Column 5) shows that that achievement is -0.062 SDU lower in Partnership than non-Partnership schools. The estimate on the *c'* path (Column 6) shows that after controlling for relational trust, this relationship decreases by 35% to 0.085 SDU. This suggests that about 35% of the relationship between Partnership and math achievement can be explained by the relational trust construct. A Chow test shows that the difference between the *c* and *c'* coefficients is marginally significant (p=0.080). Moving to the school leadership models, the coefficient estimate decreases from -0.089 on *c* to -0.106 on *c'*, suggesting that 19% of the relationship between Partnership and math achievement is explained by the school leadership construct (Chow test p=0.077). While path *a* was not statistically significant by conventional standards in the academic climate model, we also highlight that Panel B suggests academic climate may also be a mediator, with the relationship moving from -0.058 (*c*) to -0.078 (*c'*). Estimates on ELA, shown in Panel C, follow similar patterns to math but are more muted.

TABLE 5

Table 6 provides the same set of estimates for Cohort 2. Here, none of the path *a* relationships are significant. While the insignificant estimates on path *a* rule out these measures as formal mediators within the Baron & Kenny (1986) framework, the positive and significant estimates on the factor score again show that these factors are associated with higher test scores in the sample—they just do not mediate the relationship between Partnership and test scores. Additionally, while the path a relationships are not statistically significant, the math estimate in the relational trust models does decrease from -0.029 in path *c* to -0.038 and marginally significant in path *c*' and the ELA estimates follow similar though tempered patterns, providing some additional suggestive evidence that relational trust may have mediated the relationship between Partnership and student achievement in Cohort 2.

TABLE 6

DISCUSSION

We find that school climate appears to play a role in successful turnaround, though the extent to which climate is a mediator versus a type of reform readiness is unclear. In particular, our CITS analysis finds consistent evidence that students in Cohort 2 schools with stronger academic climate, school safety, relational trust, staffing climate, and school leadership fared better under Partnership than both their peers in near-selected comparison schools and in schools with weaker climate measures. We find similar but less precise patterns for school safety and discipline, relational trust, and staffing climate in Cohort 1. We also find that relational trust and school leadership mediate the relationship between Partnership and math achievement in Cohort 1.

There are some important limitations to these analyses. Ideally, we would like to observe climate measures over time for all treated and comparison schools, which would allow us to (a) estimate the effects of Partnership on climate, (b) examine the extent to which schools with stronger climate at baseline fared better under turnaround, and (c) estimate the extent to which climate formally mediated the effects of Partnership on student achievement rather than just the relationship at one point in time. However, due to the timing of the teacher survey, we only observe climate for one year for which we have student outcomes—Year 2 for Cohort 1 and Year 1 for Cohort 2. In our CITS analysis, we can therefore only estimate heterogeneous effects based on a snapshot of each climate measure. We cannot isolate the effect of Partnership on climate from the differential effects of Partnership for schools with stronger or weaker climate at the time of identification. Similarly, the mediation analysis draws on only one year of data. While we control for lagged test scores to account for growth of individual students, we cannot estimate mediation effects that control for pretreatment trends. The mediation analyses therefore need to be interpreted as associational and not causal.

Still, taken together, these two sets of estimates provide multifaceted context on the role of school-level factors in successful turnaround. Future research can draw on these results to design evaluations of turnaround reforms that account for measures such as school leadership and school climate drawing, for example, on data from statewide teacher working conditions surveys administered annually or biannually in some states (e.g., Tennessee, North Carolina, Kentucky).

There are several policy implications. State turnaround policy under ESSA could account for climate both in identification of and needs assessments in low-performing schools. In particular, states are required to include some measure of school quality and student success (SQSS) in their ESSA school performance index, which determines the bottom 5% of schools that will be slated for turnaround. Many states use measures related to chronic absenteeism, college readiness, and advanced course taking, but they could consider adding available climate measures to their indices moving forward. ESSA also calls for needs assessments in all low-performing schools. While these needs assessments have historically focused on needs directly related to student test scores, including explicit guidance for measuring school climate would prompt state and local education agencies to address climate weaknesses as part of the turnaround process. Finally, school turnaround is intended as a schoolwide intervention aimed at transforming the processes and systems underlying low performance. Turnaround policies, especially at the federal and state levels, often focus on systems related to human resources and curriculum and instruction. Explicit attention to interventions promoting positive school climate may help to engender the system-level change necessary for successful turnaround.

ENDNOTES

[1] The survey questions about climate asked about the current year while the survey questions about school leadership asked about the prior year. We can therefore also examine school leadership as a mediator for the first year of Partnership for Cohort 1.
[2] This decision was made because the survey was administered in the fall of each school year, at which point teachers would have little information on which to assess a principal who was new to the school.

REFERENCES

Andreoli, P. M., Klar, H. W., Huggins, K. S., & Buskey, F. C. (2020). Learning to lead school improvement: An analysis of rural school leadership development. *Journal of Educational Change*, *21*(4), 517–542. https://doi.org/10.1007/s10833-019-09357-z

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037//0022-3514.51.6.1173

Bonilla, S., & Dee, T. (2017). *The Effects of School Reform Under NCLB Waivers: Evidence from Focus Schools in Kentucky* (Working Paper No. 23462). National Bureau of Economic Research. https://doi.org/10.3386/w23462

Bryk, A., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation. https://www.jstor.org/stable/10.7758/9781610440967

Bryk, A., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing Schools for Improvement: Lessons from Chicago*. The University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/O/bo8212979.html

Bulach, C., & Malone, B. (1994). The Relationship of School Climate to the Implementation of School Reform. *ERS Spectrum*, *12*(4), 3–8.

Burns, J., Harbatkin, E., Strunk, K. O., Torres, C., Mcilwain, A., & Frost Waldron, S. (2023). The efficacy and implementation of Michigan's Partnership Model of school and district turnaround: Mixed-methods evidence from the first 2 years of reform implementation. *Educational Evaluation and Policy Analysis*, 1–33. https://doi.org/10.3102/01623737221141415

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. https://doi.org/10.3102/0162373718760218

Cohen, J. (2006). Social, Emotional, Ethical, and Academic Education: Creating a Climate for Learning, Participation in Democracy, and Well-Being. *Harvard Educational Review*, *76*(2), 201-237,285.

Collet, V. S. (2017). Lesson Study in a Turnaround School: Local Knowledge as a Pressure-Balanced Valve for Improved Instruction. *Teachers College Record*, *119*(6), 1–58. https://doi.org/10.1177/016146811711900605

Coopersmith, J., Cook, T. D., Zurovac, J., Chaplin, D., & Forrow, L. V. (2022). Internal and External Validity of the Comparative Interrupted Time-Series Design: A Meta-Analysis. *Journal of Policy Analysis and Management*, *41*(1), 252–277. https://doi.org/10.1002/pam.22361

Cucchiara, M. B., Rooney, E., & Robertson-Kraft, C. (2015). "I've never seen people work so hard!" Teachers' working conditions in the early stages of school turnaround. *Urban Education*, *50*(3), 259–287. https://doi.org/10.1177/0042085913501896

Daly, A. J., Der-Martirosian, C., Ong-Dean, C., Park, V., & Wishard-Guerra, A. (2011). Leading Under Sanction: Principals' Perceptions of Threat Rigidity, Efficacy, and Leadership in Underperforming Schools. *Leadership and Policy in Schools*, *10*(2), 171–206. https://doi.org/10.1080/15700763.2011.557517

Datnow, A., & Stringfield, S. (2000). Working Together for Reliable School Reform. *Journal of Education for Students Placed at Risk (JESPAR)*, *5*(1–2), 183–204. https://doi.org/10.1080/10824669.2000.9671386

Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus* (Working Paper No. 17990). National Bureau of Economic Research. https://doi.org/10.3386/w17990

Dellar, G. B. (1998). School Climate, School Improvement and Site-based Management. *Learning Environments Research*, *1*(3), 353–367. http://dx.doi.org.proxy1.cl.msu.edu/10.1023/A:1009970210393

Dinno, A. (2009). Implementing Horn's Parallel Analysis for Principal Component Analysis and Factor Analysis. *The Stata Journal: Promoting Communications on Statistics and Stata*, *9*(2), 291–298. https://doi.org/10.1177/1536867X0900900207

Dougherty, S. M., & Weiner, J. M. (2019). The Rhode to Turnaround: The Impact of Waivers to No Child Left Behind on School Performance. *Educational Policy*, *33*(4), 555–586. https://doi.org/10.1177/0895904817719520

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., & Luca, D. L. (2019). The impact of School Improvement Grants on student outcomes: Findings from a national evaluation using a regression discontinuity design. *Journal of Research on Educational Effectiveness*, *12*(2), 215–250. https://doi.org/10.1080/19345747.2019.1571654

Duke, D. L. (2004). *The turnaround principal: High-stakes leadership* (No. EJ693848). ERIC. https://eric.ed.gov/?id=EJ693848

Duke, D. L., & Salmonowicz, M. (2010). Key Decisions of a First-year 'Turnaround' Principal. *Educational Management Administration & Leadership*, *38*(1), 33–58. https://doi.org/10.1177/1741143209345450

Finnigan, K. S., & Stewart, T. J. (2009). Leading Change under Pressure: An Examination of Principal Leadership in Low-Performing Schools. *Journal of School Leadership*, *19*(5), 586–621. https://doi.org/10.1177/105268460901900504

Forsyth, P. B., Barnes, L. L. B., & Adams, C. M. (2006). Trust-effectiveness patterns in schools. *Journal of Educational Administration*, *44*(2), 122–141. https://doi.org/10.1108/09578230610652024

Fried, S. A. (2020). State Takeover: Managing Emotions, Policy Implementation, and the Support/Sanction Duality in the Holyoke Public Schools Receivership. *Harvard Educational Review*, *90*(1), 75–101. https://doi.org/10.17763/1943-5045-90.1.75

Gandhi, A. G., Slama, R., Park, S. J., Russo, P., Winner, K., Bzura, R., Jones, W., & Williamson, S. (2018). Focusing on the Whole Student: An Evaluation of Massachusetts's Wraparound Zone Initiative. *Journal of Research on Educational Effectiveness*, *11*(2), 240–266. https://doi.org/10.1080/19345747.2017.1413691

Hamilton, M. P., Heilig, J. V., & Pazey, B. L. (2014). A Nostrum of School Reform? Turning Around Reconstituted Urban Texas High Schools. *Urban Education*, *49*(2), 182–215. https://doi.org/10.1177/0042085913475636

Harbatkin, E. (2022). Staffing for School Turnaround in Rural Settings. *Leadership and Policy in Schools*, 1–23. https://doi.org/10.1080/15700763.2022.2058963

Harris, A. (2002). Effective Leadership in Schools Facing Challenging Contexts. *School Leadership & Management*, *22*(1), 15–26. https://doi.org/10.1080/13632430220143024a

Heissel, J. A., & Ladd, H. F. (2018). School turnaround in North Carolina: A regression discontinuity analysis. *Economics of Education Review*, *62*, 302–320. https://doi.org/10.1016/j.econedurev.2017.08.001

Henry, G. T., & Guthrie, J. E. (2019). *The effects of race to the top school turnaround in North Carolina* (EdWorkingPaper No. 19–107). Annenberg Institute at Brown University. https://doi.org/10.26300/w488-pf83

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*. Consortium for Educational Research and Evaluation-North Carolina. http://cerenc.org/wp-content/uploads/2015/09/ES-FINAL-Final-DST-Report-9-3-15.pdf

Henry, G. T., & Harbatkin, E. (2020). The Next Generation of State Reforms to Improve their Lowest Performing Schools: An Evaluation of North Carolina's School Transformation Intervention. *Journal of Research on Educational Effectiveness*, *13*(4), 702–730. https://doi.org/10.1080/19345747.2020.1814464

Henry, G. T., Pham, L. D., Kho, A., & Zimmer, R. (2020). Peeking into the black box of school turnaround: A formal test of mediators and suppressors. *Educational Evaluation and Policy Analysis*, *42*(2), 232–256. https://doi.org/10.3102/0162373720908600

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Huberman, M., Parrish, T., Hannan, S., Arellanes, M., & Shambaugh, L. (2011). *Turnaround schools in California: Who are they and what strategies do they use?* (No. ED592859). ERIC. https://files.eric.ed.gov/fulltext/ED592859.pdf

Jacobson, S. L., Giles, C., Ylimaki, R., & Johnson, L. (2005). Successful leadership in challenging US schools: Enabling principles, enabling schools. *Journal of Educational Administration*, *43*(6), 607–618. https://doi.org/10.1108/09578230510625700

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. https://doi.org/10.1007/BF02289233

Kraft, M. A., & Papay, J. P. (2014). Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Effectiveness and Policy Analysis*, *36*(4), 476–500.

Kutsyuruba, B., Klinger, D. A., & Hussain, A. (2015). Relationships among school climate, school safety, and student achievement and well-being: A review of the literature. *Review of Education*, *3*(2), 103–135. https://doi.org/10.1002/rev3.3043

Leithwood, K., Harris, A., & Strauss, T. (2010). *Leading School Turnaround: How Successful Leaders Transform Low-Performing Schools*. Jossey-Bass.

Leithwood, K., & Jantzi, D. (1990). Transformational leadership: How principals can help reform school cultures. *School Effectiveness and School Improvement*, *1*(4), 249–280. https://doi.org/10.1080/0924345900010402

Lenhoff, S. W., & Ulmer, J. B. (2016). Reforming for "all" or for "some": Misalignment in the discourses of education reformers and implementers. *Education Policy Analysis Archives*, *24*, 108–108. https://doi.org/10.14507/epaa.24.2273

Malen, B., Croninger, R., Muncey, D., & Redmond-Jones, D. (2002). Reconstituting schools: "Testing" the "theory of action." *Educational Evaluation and Policy Analysis*, *24*(2), 113–132. https://doi.org/10.3102/01623737024002113

Mayger, L. K., & Hochbein, C. D. (2021). Growing Connected: Relational Trust and Social Capital in Community Schools. *Journal of Education for Students Placed at Risk (JESPAR)*, *26*(3), 210–235. https://doi.org/10.1080/10824669.2020.1824676

Meyers, C. V. (2020). District-Led School Turnaround: A Case Study Of One U.S. District's Turnaround Launch For Multiple Schools. *Leadership and Policy in Schools*, *19*(4), 710–729. https://doi.org/10.1080/15700763.2019.1637902

Meyers, C. V., & Hambrick Hitt, D. (2017). School Turnaround Principals: What Does Initial Research Literature Suggest They Are Doing to Be Successful? *Journal of Education for Students Placed at Risk (JESPAR)*, *22*(1), 38–56. https://doi.org/10.1080/10824669.2016.1242070

Meyers, C. V., & Smylie, M. A. (2017). Five Myths of School Turnaround Policy and Practice. *Leadership and Policy in Schools*, *16*(3), 502–523. https://doi.org/10.1080/15700763.2016.1270333

Nolan, K. (2018). The Lived Experience of Market-Based School Reform: An Ethnographic Portrait of Teachers' Policy Enactments in an Urban School. *Educational Policy*, *32*(6), 797–822. https://doi.org/10.1177/0895904816673742

Papay, J., & Hannon, M. (2018, November 8). *The Effects of School Turnaround Strategies in Massachusetts*. 2018 APPAM Fall Research Conference: *Evidence for Action: Encouraging Innovation and Improvement*, Washington, D.C. https://appam.confex.com/appam/2018/webprogram/Paper26237.html

Patterson, J. A., AlSabatin, H., Anderson, A., Klepacka, M., Lawrence, J., & Miner, B. (2021). A Distributed Leadership Perspective on Implementing Instructional Reform: A Case Study of an Urban Middle School. *Journal of School Leadership*, *31*(3), 248–267. https://doi.org/10.1177/1052684620904942

Peurach, D. J., & Neumerski, C. M. (2015). Mixing metaphors: Building infrastructure for large scale school turnaround. *Journal of Educational Change*, *16*(4), 379–420. https://doi.org/10.1007/s10833-015-9259-z

Pham, L. D. (2022). Why do we find these effects? An examination of mediating pathways explaining the effects of school turnaround. *Journal of Research on Educational Effectiveness*, 82–105. https://doi.org/10.1080/19345747.2022.2081276

Pham, L. D., Henry, G. T., Kho, A., & Zimmer, R. (2020). Sustainability and Maturation of School Turnaround: A Multiyear Evaluation of Tennessee's Achievement School District and Local Innovation Zones. *AERA Open*, *6*(2). https://doi.org/10.1177/2332858420922841

Player, D., Hambrick Hitt, D., & Robinson, W. (2014). District Readiness to Support School Turnaround: A Users' Guide to Inform the Work of State Education Agencies and Districts. In *Center on School Turnaround at WestEd*. Center on School Turnaround at WestEd. https://eric.ed.gov/?id=ED559742

Player, D., & Katz, V. (2016). Assessing School Turnaround: Evidence from Ohio. *The Elementary School Journal*, *116*(4), 675–698. https://doi.org/10.1086/686467

Quartz, K. H., Geller, R. C., & Mcqueen, S. S. (2020). A Beautiful Struggle: Reimagining Neighborhood Schools in Urban Communities. *Teachers College Record*, *122*(2), 1–46. https://doi.org/10.1177/016146812012200204

Redding, C., & Nguyen, T. D. (2020). The Relationship Between School Turnaround and Student Outcomes: A Meta-Analysis. *Educational Evaluation and Policy Analysis*, *41*(4), 0162373720949513. https://doi.org/10.3102/0162373720949513

Redding, S., & Corbett, J. (2018). *Shifting school culture to spark rapid improvement: A quick start guide for principals and their teams. The center on school turnaround four domains series* (No. ED592993). ERIC. https://eric.ed.gov/?id=ED592993

Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): A state-of-the-art review. *School Effectiveness and School Improvement*, *25*(2), 197–230. https://doi.org/10.1080/09243453.2014.885450

Rice, J. K., & Croninger, R. G. (2005). Resource Generation, Reallocation, or Depletion: An Analysis of the Impact of Reconstitution on School Capacity. *Leadership and Policy in Schools*, *4*(2), 73–103. https://doi.org/10.1080/15700760590965569

Rice, J. K., & Malen, B. (2003). The Human Costs of Education Reform: The Case of School Reconstitution. *Educational Administration Quarterly*, *39*(5), 635–666. https://doi.org/10.1177/0013161X03257298

Schueler, B. E., Asher, C. A., Larned, K. E., Mehrotra, S., & Pollard, C. (2021). Improving low-performing schools: A meta-analysis of impact evaluation studies. *American Educational Research Journal*, *59*(5), 975–1010. https://doi.org/10.3102/00028312211060855

Schueler, B. E., Goodman, J. S., & Deming, D. J. (2017). Can States Take Over and Turn Around School Districts? Evidence From Lawrence, Massachusetts. *Educational Evaluation and Policy Analysis*, *39*(2), 311–332. https://doi.org/10.3102/0162373716685824

Sims, S., Anders, J., & Zieger, L. (2022). The Internal Validity of the School-Level Comparative Interrupted Time Series Design: Evidence From Four New Within-Study Comparisons. *Journal of Research on Educational Effectiveness*, *15*(4), 876–897. https://doi.org/10.1080/19345747.2022.2051652

Slavin, R. E. (1998). Sand, Bricks, and Seeds: School Change Strategies and Readiness for Reform. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *International Handbook of Educational Change: Part One* (pp. 1299–1313). Springer Netherlands. https://doi.org/10.1007/978-94-011-4944-0_62

Strunk, K. O., Marsh, J. A., Bush-Mecenas, S. C., & Duque, M. R. (2016). The Best Laid Plans: An Examination of School Plan Quality and Implementation in a School Improvement Initiative. *Educational Administration Quarterly*, *52*(2), 259–309. https://doi.org/10.1177/0013161X15616864

Strunk, K. O., Marsh, J. A., Hashim, A. K., & Bush-Mecenas, S. (2016). Innovation and a return to the status quo: A mixed-methods study of school reconstitution. *Educational Evaluation and Policy Analysis*, *38*(3), 549–577. https://doi.org/10.3102/0162373716642517

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. https://doi.org/10.1162/EDFP_a_00188

Sun, M., Penner, E. K., & Loeb, S. (2017). Resource- and Approach-Driven Multidimensional Change: Three-Year Effects of School Improvement Grants. *American Educational Research Journal*, *54*(4), 607–643. https://doi.org/10.3102/0002831217695790

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research*, *83*(3), 357–385. https://doi.org/10.3102/0034654313483907

Thompson, C., Henry, G., & Preston, C. (2016). School Turnaround Through Scaffolded Craftsmanship. *Teachers College Record*, *118*(13), 1–26.

Torres, A. C. (2016). The uncertainty of high expectations: How principals influence relational trust and teacher turnover in no excuses charter schools. *Journal of School Leadership*, *26*(1), 61–91. https://doi.org/10.1177/105268461602600103

Trujillo, T., & Renée, M. (2015). Irrational Exuberance for Market-based Reform: How Federal Turnaround Policies Thwart Democratic Schooling. *Teachers College Record*, *117*(6), 34.

Weixler, L. B., Harris, D. N., & Barrett, N. (2018). Teachers' Perspectives on the Learning and Work Environments Under the New Orleans School Reforms. *Educational Researcher*, *47*(8), 502–515. https://doi.org/10.3102/0013189X18787806

Welsh, R. O., & Williams, S. M. (2018). Incentivizing improvement or imposition? An examination of the response to gubernatorial school takeover and statewide turnaround districts. *Education Policy Analysis Archives*, *26*, 124–124. https://doi.org/10.14507/epaa.26.3679

Yoon, Y., Kim, K.-R., Park, H., Kim, S., & Kim, Y.-J. (2020). *Stepwise school opening online and off-line and an impact on the epidemiology of COVID-19 in the pediatric population*. medRXiv. https://doi.org/10.1101/2020.08.03.20165589

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. https://doi.org/10.3102/0162373717705729

TABLES

Table 1. Baseline school-level characteristics by cohort and treatment condition

Panel A. Student level

	(1) Cohort 1	(2) Cohort 1 analytic	(3) Cohort 1 comparison	(4) Rest of state 2016-17	(5) Cohort 2	(6) Cohort 2 analytic	(7) Cohort 2 comparison	(8) Rest of state 2017-18
Economically disadvantaged	0.865 (0.341)	0.867 (0.340)	0.857 (0.350)	0.477 (0.499)	0.924 (0.265)	0.925 (0.263)	0.896 (0.306)	0.504 (0.500)
English learner	0.025 (0.157)	0.027 (0.161)	0.131 (0.337)	0.064 (0.245)	0.087 (0.281)	0.083 (0.275)	0.154 (0.361)	0.065 (0.246)
Special education	0.155 (0.362)	0.157 (0.364)	0.130 (0.336)	0.119 (0.323)	0.148 (0.355)	0.151 (0.358)	0.128 (0.334)	0.120 (0.325)
White	0.044 (0.205)	0.047 (0.211)	0.185 (0.388)	0.688 (0.463)	0.076 (0.265)	0.064 (0.245)	0.160 (0.366)	0.701 (0.458)
Black	0.890 (0.313)	0.883 (0.322)	0.615 (0.487)	0.153 (0.360)	0.786 (0.410)	0.806 (0.396)	0.650 (0.477)	0.136 (0.343)
Hispanic or Latinx	0.036 (0.186)	0.038 (0.192)	0.148 (0.355)	0.078 (0.268)	0.097 (0.297)	0.098 (0.297)	0.147 (0.354)	0.078 (0.268)
Asian, multiple races, or other	0.031 (0.172)	0.033 (0.178)	0.053 (0.223)	0.081 (0.272)	0.040 (0.196)	0.033 (0.178)	0.043 (0.202)	0.085 (0.278)
Math score	-1.160 (0.775)	-1.151 (0.779)	-0.817 (0.849)	0.031 (0.990)	-1.057 (0.761)	-1.095 (0.753)	-0.834 (0.814)	0.063 (0.981)
ELA score	-1.103 (0.797)	-1.100 (0.808)	-0.719 (0.902)	0.028 (0.991)	-0.964 (0.755)	-0.999 (0.739)	-0.740 (0.832)	0.056 (0.987)
Observations	5,565	4,955	12,646	526,833	13,417	9,699	21,706	507,166

Panel B. School level

	(1) Cohort 1	(2) Cohort 1 analytic	(3) Cohort 1 comparison	(4) Rest of state 2016-17	(5) Cohort 2	(6) Cohort 2 analytic	(7) Cohort 2 comparison	(8) Rest of state 2017-18
Economically disadvantaged	0.874 (0.068)	0.875 (0.067)	0.839 (0.155)	0.536 (0.256)	0.925 (0.061)	0.927 (0.037)	0.884 (0.113)	0.564 (0.246)
English learner	0.026 (0.068)	0.028 (0.074)	0.112 (0.218)	0.069 (0.136)	0.070 (0.147)	0.073 (0.149)	0.104 (0.217)	0.069 (0.133)
Special education	0.175 (0.055)	0.169 (0.054)	0.153 (0.058)	0.163 (0.147)	0.165 (0.047)	0.178 (0.050)	0.145 (0.084)	0.165 (0.149)
White	0.041 (0.062)	0.046 (0.067)	0.218 (0.272)	0.676 (0.292)	0.091 (0.143)	0.079 (0.106)	0.256 (0.305)	0.695 (0.272)
Black	0.882 (0.155)	0.869 (0.168)	0.592 (0.343)	0.166 (0.269)	0.788 (0.250)	0.810 (0.228)	0.595 (0.365)	0.141 (0.238)
Hispanic or Latinx	0.043 (0.080)	0.048 (0.087)	0.131 (0.221)	0.077 (0.113)	0.075 (0.154)	0.073 (0.168)	0.098 (0.185)	0.078 (0.110)
Asian, multiple races, or other	0.034 (0.041)	0.037 (0.044)	0.059 (0.052)	0.081 (0.088)	0.045 (0.057)	0.038 (0.046)	0.050 (0.056)	0.086 (0.092)
Math score	-1.267 (0.518)	-1.325 (0.518)	-0.928 (0.954)	-0.059 (1.001)	-0.947 (0.768)	-0.972 (0.804)	-0.709 (0.850)	-0.027 (0.971)
ELA score	-1.153 (0.665)	-1.192 (0.664)	-0.726 (0.879)	-0.053 (1.013)	-1.003 (0.718)	-0.976 (0.745)	-0.675 (0.862)	-0.011 (0.997)
Enrollment	493.7 (144.8)	522.8 (138.4)	471.7 (440.6)	417.5 (235.0)	411.5 (206.6)	467.5 (212.7)	386.8 (229.4)	422.9 (241.4)
Observations	24	20	63	2,340	73	43	115	2,200

Note: Panel A provides student-level means and Panel B provides school-level means. Standard deviations in parentheses. Values based on identification year, which is 2016-17 for Cohort 1 and 2017-18 for Cohort 2. Sample restricted to students in grades 4-8.

Table 2. Factor loadings

Panel A. Academic climate

	Factor loading	ψ (uniqueness)
Grade school: Professional development and support for teachers	0.708	0.499
Grade school: Curriculum	0.763	0.417
Grade school: Academic achievement	0.840	0.294
Grade school: Literacy practice and instruction	0.836	0.302
Grade school: Ability to support student subgroups	0.743	0.448
Grade school: Ability to meet academic needs	0.755	0.431
N	4,176	
α	0.864	

Panel B. School safety and discipline

	Factor loading	ψ (uniqueness)
Agree: Safe and orderly environment	0.863	0.256
Agree: Student fights are frequent (r)	0.610	0.627
Agree: Administrators consistently enforce behavioral standards	0.746	0.444
Agree: Students listen to staff	0.739	0.454
Grade school: Student discipline	0.799	0.362
Agree: Teachers effectively manage student behavior	0.750	0.438
Agree: School meets students' social-emotional needs	0.761	0.421
N	4,461	
α	0.869	

Panel C. Relational trust

	Factor loading	ψ (uniqueness)
Agree: Teachers have high expectations for students	0.803	0.356
Agree: Staff works to build relationships with parents	0.722	0.479
Agree: Most of my colleagues share my beliefs about the central mission of school	0.595	0.647
Agree: Teachers have strong rapport with students	0.751	0.436
Agree: Students enthusiastic to come to school	0.625	0.609
Agree: Teachers consistently enforce behavioral standards	0.735	0.460
N	4,557	
α	0.794	

Panel D. Staffing climate

	Factor loading	ψ (uniqueness)
Agree: There is a high degree of staff turnover	0.774	0.401
Grade school: Teacher retention	0.924	0.145
Grade school: Staff retention	0.916	0.161
Grade school: Teacher attendance	0.617	0.619
N	4,127	
α	0.833	

Panel E. School leadership

	Factor loading	ψ (uniqueness)
Work with staff to meet curriculum standards	0.897	0.195
Communicate central mission of the school	0.891	0.207
Use evidence to make data-driven decisions	0.895	0.199
Establish clear discipline policies	0.828	0.314
Work with community partners	0.866	0.251
Facilitate and encourage teacher PD	0.881	0.224
Encourage parental engagement	0.861	0.259
Communicate strategies and goals with teachers	0.909	0.174
N	3,908	
α	0.957	

NOTE: Factor loadings from principal factors analysis. In Panels A through D, items prefixed with "Grade school" come from the question, "We are interested in how well you believe your school is implementing activities in the following areas. Please give your school a grade, from A (high) to F (low) in each of the following areas." Items prefixed with "Agree" are from the question, "Please indicate the extent to which you agree or disagree with the following statements about your school." Response options were strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree. All items in Panel D come from the question, "Indicate how effectively your principal or school leader performed each of the following." Response options were not at all effectively, slightly effectively, somewhat effectively, very effectively, and extremely effectively. Response options coded as 1-5, with 1 representing the lowest response (i.e., F, strongly disagree, not at all effectively) and 5 the highest. Items denoted with "(r)" were reverse-coded.

Table 3. Factor descriptive statistics by treatment condition

	2018-19				2019-20			
	Cohort 1	Cohort 2	Comparison	Total	Cohort 1	Cohort 2	Comparison	Total
Academic climate	-0.043	-0.257	-0.035	-0.095	-0.047	0.004	0.049	0.025
	(0.548)	(0.642)	(0.635)	(0.629)	(0.430)	(0.575)	(0.655)	(0.610)
School safety and discipline	-0.166	-0.218	0.079	-0.036	-0.187	0.055	0.077	0.039
	(0.520)	(0.625)	(0.763)	(0.707)	(0.464)	(0.670)	(0.740)	(0.697)
Relational trust	-0.095	-0.149	0.037	-0.032	-0.111	0.046	0.100	0.060
	(0.565)	(0.472)	(0.623)	(0.581)	(0.397)	(0.556)	(0.563)	(0.546)
Staffing climate	-0.206	-0.218	0.160	0.005	-0.172	-0.045	0.152	0.059
	(0.539)	(0.611)	(0.768)	(0.719)	(0.586)	(0.550)	(0.751)	(0.691)
School leadership [1]	0.205	-0.068	-0.055	-0.019	0.214	0.159	-0.085	0.011
	(0.608)	(0.659)	(0.600)	(0.621)	(0.406)	(0.526)	(0.650)	(0.607)

Note: School-level means with standard deviations in parentheses. Sample limited to those schools that meet minimum participation thresholds in each year.

[1] Question items ask about the previous school year. 2018-19 statistics therefore apply to 2017-18 while 2019-20 statistics apply to 2018-19.

Table 4. Estimated effects by tercile, grades 4-8 student achievement

Panel A. Math

	(1) Cohort 1 Academic climate	(2) School safety	(3) Relational trust	(4) Staffing climate	(5) Leadership	(6) Cohort 2 Academic climate	(7) School safety	(8) Relational trust	(9) Staffing climate	(10) Leadership
Partnership × Year	-0.018	-0.018	-0.018	-0.017	-0.018	-0.011	-0.011	-0.011	-0.011	-0.011
	(0.016)	(0.015)	(0.015)	(0.015)	(0.016)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Partnership	0.087+	0.106+	0.084+	0.047	0.101+					
2017-18 × Low	(0.052)	(0.054)	(0.046)	(0.071)	(0.055)	*a*	*a*	*a*	*a*	*a*
Partnership	0.065	0.087+	0.095	0.116**	0.168*					
2017-18 × Mid	(0.067)	(0.050)	(0.064)	(0.044)	(0.084)	*a*	*a*	*a*	*a*	*a*
Partnership	0.084	0.040	0.065	0.069	0.020					
2017-18 × High	(0.068)	(0.089)	(0.071)	(0.075)	(0.060)	*a*	*a*	*a*	*a*	*a*
Partnership	0.063	0.003	-0.019	-0.006	0.079	0.015	0.021	0.012	0.035	0.043
2018-19 × Low	(0.091)	(0.053)	(0.060)	(0.080)	(0.085)	(0.032)	(0.029)	(0.039)	(0.031)	(0.034)
Partnership	0.033	0.094	0.054	0.057	0.014	0.044	0.048	0.060+	0.064+	0.080*
2018-19 × Mid	(0.068)	(0.070)	(0.067)	(0.060)	(0.056)	(0.032)	(0.033)	(0.032)	(0.037)	(0.035)
Partnership	0.098	0.108	0.130+	0.119+	0.087	0.138***	0.168***	0.137***	0.138***	0.117**
2018-19 × High	(0.064)	(0.074)	(0.066)	(0.070)	(0.068)	(0.037)	(0.032)	(0.033)	(0.032)	(0.036)
Constant	-0.281	-0.253	-0.208	-0.149	-0.269	-0.300	-0.312	-0.324	-0.308	-0.288
	(0.284)	(0.287)	(0.298)	(0.286)	(0.266)	(0.293)	(0.292)	(0.292)	(0.295)	(0.295)
N	100,480	100,480	100,480	100,480	100,480	185,594	185,594	185,594	185,594	185,594
R^2	0.519	0.519	0.519	0.519	0.519	0.485	0.485	0.485	0.484	0.484
Within R^2	0.447	0.447	0.447	0.447	0.447	0.452	0.452	0.452	0.452	0.452

Panel B. ELA

	(1) Cohort 1 Academic climate	(2) School safety	(3) Relational trust	(4) Staffing climate	(5) Leadership	(6) Cohort 2 Academic climate	(7) School safety	(8) Relational trust	(9) Staffing climate	(10) Leadership
Partnership × Year	-0.004	-0.004	-0.004	-0.004	-0.004	-0.000	-0.000	0.000	-0.000	-0.000
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Partnership 2017-18 × Low	0.077+	0.107*	0.106*	0.094	0.048					
	(0.043)	(0.052)	(0.053)	(0.059)	(0.036)	*a*	*a*	*a*	*a*	*a*
Partnership 2017-18 × Mid	0.145**	0.062	0.077	0.095+	0.190***					
	(0.047)	(0.051)	(0.057)	(0.051)	(0.045)	*a*	*a*	*a*	*a*	*a*
Partnership 2017-18 × High	0.037	0.068	0.062	0.057	0.057					
	(0.050)	(0.059)	(0.051)	(0.050)	(0.052)	*a*	*a*	*a*	*a*	*a*
Partnership 2018-19 × Low	0.072	0.053	0.033	0.048	0.070	0.048	0.046	0.013	0.018	-0.021
	(0.080)	(0.060)	(0.060)	(0.067)	(0.072)	(0.039)	(0.033)	(0.040)	(0.033)	(0.030)
Partnership 2018-19 × Mid	0.060	0.049	0.069	0.047	0.060	-0.020	0.014	-0.016	-0.002	0.048
	(0.058)	(0.061)	(0.058)	(0.063)	(0.058)	(0.032)	(0.041)	(0.029)	(0.044)	(0.041)
Partnership 2018-19 × High	0.049	0.076	0.068	0.073	0.049	0.044	0.017	0.079*	0.066*	0.059*
	(0.061)	(0.071)	(0.066)	(0.063)	(0.061)	(0.035)	(0.029)	(0.033)	(0.026)	(0.029)
Constant	-0.536+	-0.546+	-0.553+	-0.548+	-0.530+	-0.558+	-0.557+	-0.552+	-0.555+	-0.525
	(0.280)	(0.277)	(0.288)	(0.289)	(0.279)	(0.319)	(0.325)	(0.316)	(0.319)	(0.318)
N	100,667	100,667	100,667	100,667	100,667	185,975	185,975	185,975	185,975	185,975
R^2	0.538	0.538	0.538	0.538	0.538	0.508	0.508	0.508	0.508	0.508
Within R^2	0.474	0.474	0.474	0.474	0.474	0.477	0.476	0.477	0.477	0.477

Note: Estimates from comparative interrupted time series models with heterogeneous effects estimated in treated years only. All models include lagged test score, year fixed effects, grade fixed effects, baseline school covariates interacted with a linear time trend, and student covariates. School covariates include share of students who were economically disadvantaged, special education, English learners, race/ethnicity (Black, Hispanic or Latinx, Asian/other/multiple races, with White as the reference category), and logged school enrollment. Student covariates include indicators for race/ethnicity (Black, Asian, Hispanic, and other race, with White as the reference category) female, economically disadvantaged, and special education. Robust standard errors clustered at the school level in parentheses.

[a] 2017-18 is a pre-intervention year for Cohort 2 and therefore included as part of the linear pretreatment trend.

*+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

Table 5. Mediated results, grades 4-8 student achievement, Cohort 1

Panel A. Path A

	(1) Academic climate	(2) School safety	(3) Relational trust	(4) Staffing climate	(5) Leadership
Partnership	0.209	0.034	0.225+	-0.030	0.273*
	(0.139)	(0.142)	(0.134)	(0.142)	(0.136)
N	26,487	26,870	27,382	26,926	27,115
R^2	0.235	0.317	0.251	0.363	0.199

Panel B. Math

	Academic climate		School safety		Relational trust		Staffing climate		Leadership	
	(1) C	(2) C'	(3) C	(4) C'	(5) C	(6) C'	(7) C	(8) C'	(9) C	(10) C'
Partnership	-0.058+	-0.078*	-0.063+	-0.067*	-0.062+	-0.085**	-0.064+	-0.062*	-0.089**	-0.106***
	(0.035)	(0.033)	(0.034)	(0.030)	(0.033)	(0.030)	(0.034)	(0.031)	(0.033)	(0.031)
Factor score		0.098***		0.097***		0.105***		0.084***		0.062**
(path *b*)		(0.018)		(0.018)		(0.020)		(0.019)		(0.020)
N	26,487	26,487	26,870	26,870	27,382	27,382	26,926	26,926	27,115	27,115
R^2	0.621	0.625	0.620	0.623	0.619	0.622	0.621	0.623	0.665	0.666

Panel C. ELA

	Academic climate		School safety		Relational trust		Staffing climate		Leadership	
	(1) C	(2) C'	(3) C	(4) C'	(5) C	(6) C'	(7) C	(8) C'	(9) C	(10) C'
Partnership	-0.033	-0.050	-0.044	-0.046	-0.043	-0.065*	-0.044	-0.042	-0.053+	-0.066*
	(0.033)	(0.031)	(0.032)	(0.030)	(0.032)	(0.030)	(0.033)	(0.031)	(0.032)	(0.031)
Factor score		0.078***		0.080***		0.101***		0.074***		0.048**
(path *b*)		(0.017)		(0.019)		(0.020)		(0.019)		(0.017)
N	26,487	26,487	26,870	26,870	27,382	27,382	26,926	26,926	27,115	27,115
R^2	0.609	0.612	0.610	0.612	0.609	0.612	0.610	0.612	0.644	0.645

*Note: Ns vary slightly by construct because we retain students in all schools meeting minimum threshold requirements for a given construct. Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

Table 6. Mediated results, grades 4-8 student achievement, Cohort 2

Panel A. Path A

	(1) Academic climate	(2) School safety	(3) Relational trust	(4) Staffing climate	(5) Leadership
Partnership	0.069	-0.031	0.086	-0.198	0.062
	(0.116)	(0.120)	(0.098)	(0.123)	(0.124)
N	33,399	33,379	34,169	33,250	32,085
R^2	0.185	0.283	0.256	0.322	0.157

Panel B. Math

	Academic climate		School safety		Relational trust		Staffing climate		Leadership	
	(1) C	(2) C'	(3) C	(4) C'	(5) C	(6) C'	(7) C	(8) C'	(9) C	(10) C'
Partnership	-0.028	-0.035	-0.036	-0.033	-0.029	-0.038+	-0.035	-0.017	-0.027	-0.031
	(0.024)	(0.022)	(0.025)	(0.022)	(0.024)	(0.023)	(0.025)	(0.023)	(0.025)	(0.023)
Factor score		0.104***		0.100***		0.100***		0.092***		0.057**
(path *b*)		(0.016)		(0.015)		(0.019)		(0.016)		(0.019)
N	33,399	33,399	33,379	33,379	34,169	34,169	33,250	33,250	32,085	32,085
R^2	0.595	0.599	0.595	0.599	0.595	0.598	0.597	0.601	0.646	0.647

Panel C. ELA

	Academic climate		School safety		Relational trust		Staffing climate		Leadership	
	(1) C	(2) C'	(3) C	(4) C'	(5) C	(6) C'	(7) C	(8) C'	(9) C	(10) C'
Partnership	-0.030	-0.034	-0.034	-0.031	-0.030	-0.038+	-0.031	-0.017	-0.033	-0.036
	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.022)	(0.023)	(0.023)	(0.023)	(0.022)
Factor score		0.063***		0.070***		0.098***		0.071***		0.048**
(path *b*)		(0.016)		(0.016)		(0.020)		(0.016)		(0.016)
N	33,399	33,399	33,379	33,379	34,169	34,169	33,250	33,250	32,085	32,085
R^2	0.590	0.591	0.590	0.592	0.590	0.593	0.591	0.594	0.629	0.630

*Note: Ns vary slightly by construct because we retain students in all schools meeting minimum threshold requirements for a given construct. Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001*

FIGURES

Figure 1. Partnership theory of change

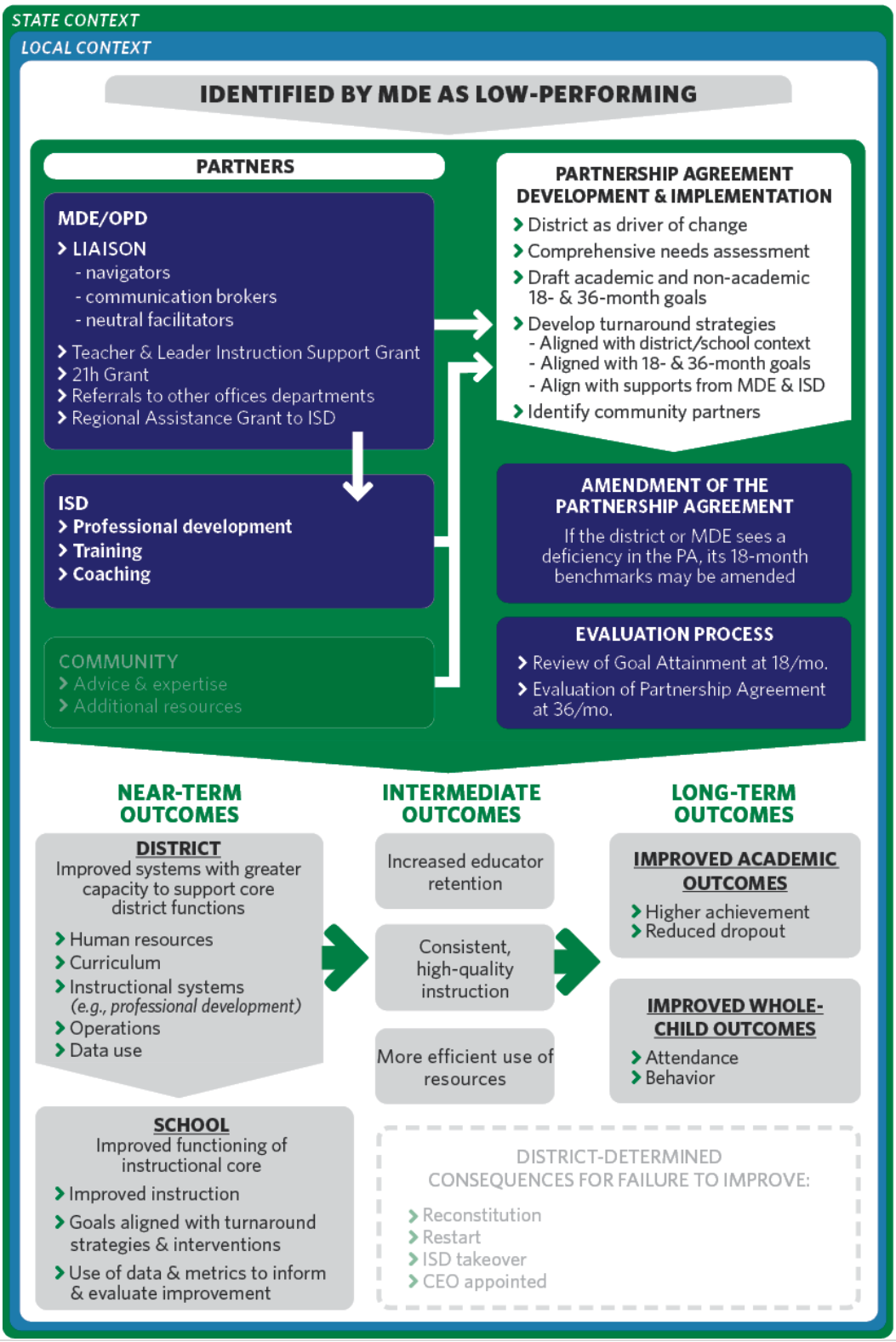


Figure 2. Mediation model

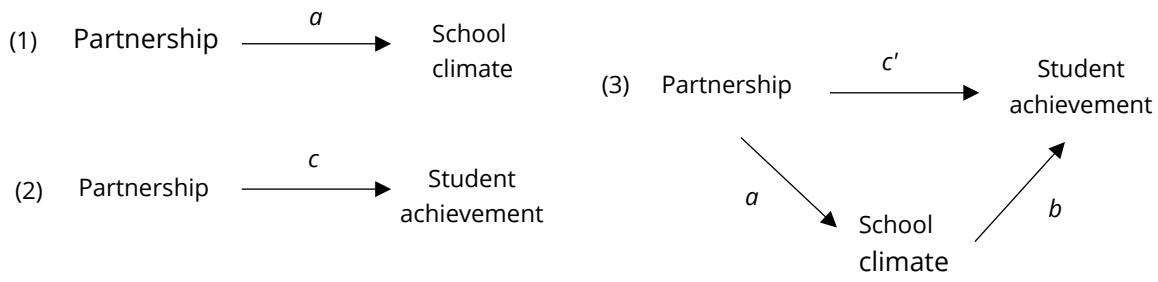
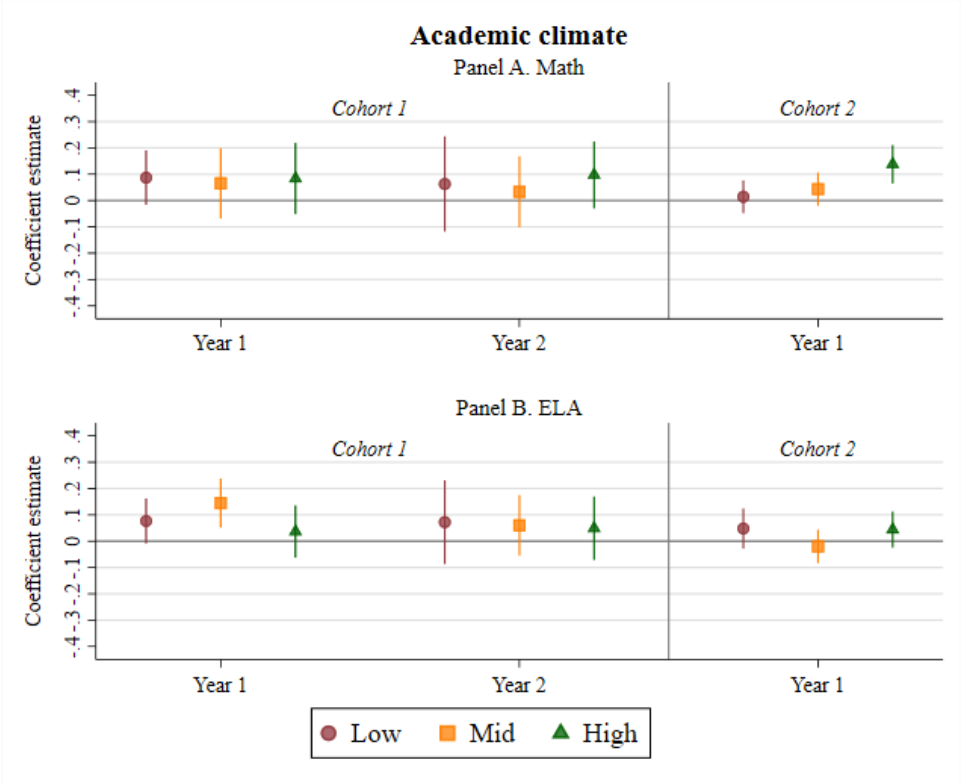
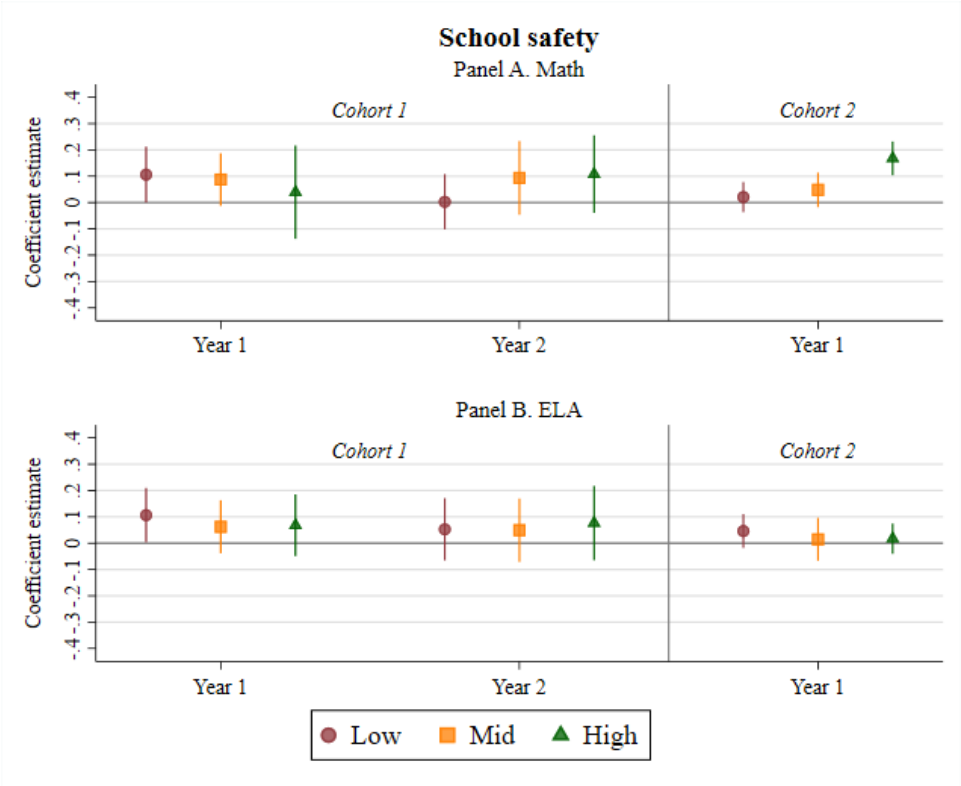


Figure 3. CITS estimates by climate tercile

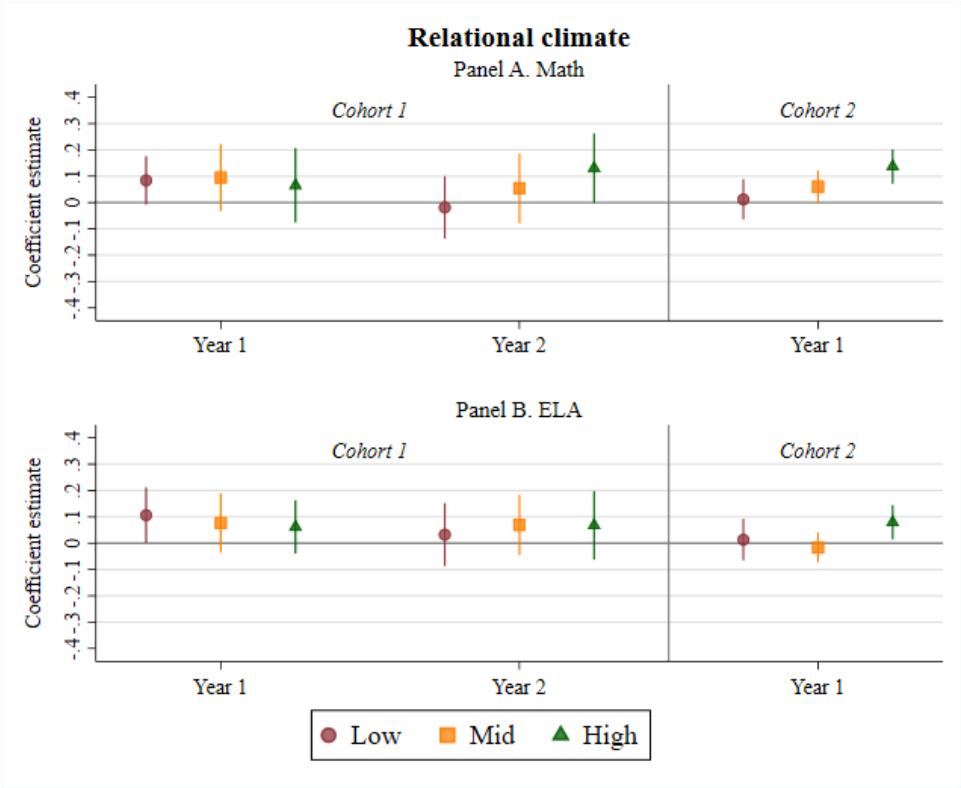
First Panel



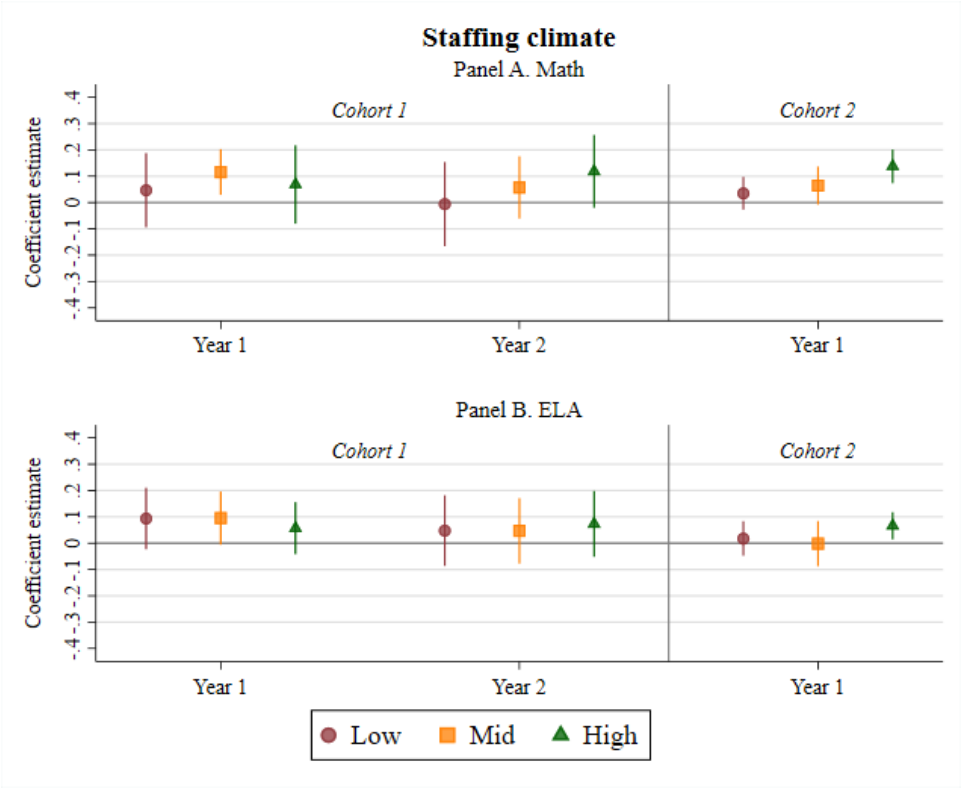
Second Panel



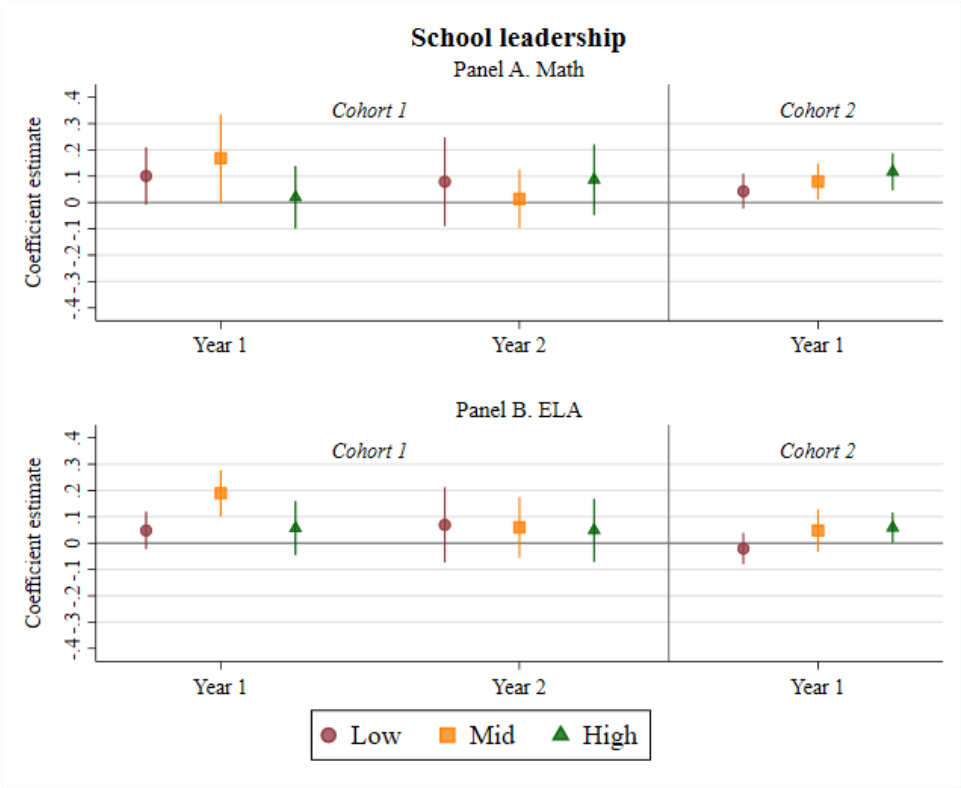
Third Panel



Fourth Panel



Fifth Panel


School leadership
Panel A. Math
Cohort 1
Cohort 2
Coefficient estimate
-.4 -.3 -.2 -.1 0 .1 .2 .3 .4
Year 1
Year 2
Year 1
Panel B. ELA
Cohort 1
Cohort 2
Coefficient estimate
-.4 -.3 -.2 -.1 0 .1 .2 .3 .4
Year 1
Year 2
Year 1
Low
Mid
High

SUPPLEMENTAL TABLES

Table A-1. Baseline descriptive statistics by factor terciles

Panel A. Academic climate

	Cohort 1				Cohort 2			
	Low	Mid	High	Missing	Low	Mid	High	Missing
Black	0.900	0.928	0.832	0.951	0.820	0.786	0.866	0.781
	(0.162)	(0.141)	(0.174)	(0.030)	(0.199)	(0.226)	(0.241)	(0.277)
Hispanic	0.047	0.028	0.074	0.018	0.065	0.069	0.075	0.079
	(0.117)	(0.053)	(0.098)	(0.009)	(0.104)	(0.168)	(0.196)	(0.133)
Economically disadvantaged	0.877	0.926	0.928	0.895	0.917	0.939	0.915	0.923
	(0.043)	(0.022)	(0.026)	(0.040)	(0.053)	(0.039)	(0.054)	(0.079)
English learner	0.031	0.005	0.048	0.012	0.046	0.066	0.077	0.042
	(0.074)	(0.004)	(0.084)	(0.011)	(0.082)	(0.136)	(0.170)	(0.095)
Special education	0.179	0.216	0.171	0.225	0.200	0.201	0.163	0.160
	(0.036)	(0.107)	(0.047)	(0.150)	(0.073)	(0.042)	(0.055)	(0.046)
Math 3-8	-1.197	-1.119	-1.042	-1.184	-1.174	-1.031	-0.906	-1.001
	(0.210)	(0.307)	(0.254)	(0.230)	(0.179)	(0.123)	(0.183)	(0.240)
ELA 3-8	-1.104	-1.005	-0.971	-1.083	-1.051	-0.932	-0.906	-0.886
	(0.247)	(0.151)	(0.218)	(0.272)	(0.137)	(0.137)	(0.141)	(0.229)
Observations	9	9	9	7	15	16	17	31

Panel B. School safety

	Cohort 1				Cohort 2			
	Low	Mid	High	Missing	Low	Mid	High	Missing
Black	0.855	0.897	0.908	0.951	0.788	0.803	0.878	0.781
	(0.195)	(0.133)	(0.157)	(0.030)	(0.236)	(0.206)	(0.226)	(0.277)
Hispanic	0.065	0.037	0.047	0.018	0.075	0.076	0.060	0.079
	(0.122)	(0.050)	(0.097)	(0.009)	(0.173)	(0.129)	(0.180)	(0.133)
Economically disadvantaged	0.884	0.929	0.918	0.895	0.924	0.925	0.922	0.923
	(0.047)	(0.031)	(0.023)	(0.040)	(0.053)	(0.053)	(0.044)	(0.079)
English learner	0.030	0.018	0.035	0.012	0.057	0.055	0.078	0.042
	(0.075)	(0.025)	(0.086)	(0.011)	(0.133)	(0.099)	(0.167)	(0.095)
Special education	0.192	0.173	0.201	0.225	0.209	0.196	0.159	0.160
	(0.043)	(0.055)	(0.104)	(0.150)	(0.063)	(0.045)	(0.061)	(0.046)
Math 3-8	-1.049	-1.098	-1.136	-1.184	-1.121	-1.054	-0.902	-1.001
	(0.318)	(0.236)	(0.268)	(0.230)	(0.153)	(0.211)	(0.143)	(0.240)
ELA 3-8	-0.998	-0.976	-1.053	-1.083	-1.025	-0.928	-0.915	-0.886
	(0.250)	(0.224)	(0.161)	(0.272)	(0.125)	(0.182)	(0.114)	(0.229)
Observations	9	9	9	7	15	16	17	31

Panel C. Relational trust

	Cohort 1				Cohort 2			
	Low	Mid	High	Missing	Low	Mid	High	Missing
Black	0.890	0.867	0.903	0.951	0.827	0.814	0.833	0.781
	(0.158)	(0.179)	(0.154)	(0.030)	(0.169)	(0.258)	(0.238)	(0.277)
Hispanic	0.051	0.049	0.049	0.018	0.037	0.107	0.063	0.079
	(0.116)	(0.067)	(0.096)	(0.009)	(0.051)	(0.202)	(0.177)	(0.133)
Economically disadvantaged	0.883	0.921	0.927	0.895	0.938	0.900	0.933	0.923
	(0.050)	(0.026)	(0.019)	(0.040)	(0.053)	(0.046)	(0.042)	(0.079)
English learner	0.032	0.016	0.035	0.012	0.028	0.080	0.079	0.042
	(0.074)	(0.026)	(0.086)	(0.011)	(0.031)	(0.155)	(0.166)	(0.095)
Special education	0.188	0.215	0.164	0.225	0.201	0.191	0.172	0.160
	(0.046)	(0.097)	(0.056)	(0.150)	(0.057)	(0.072)	(0.047)	(0.046)
Math 3-8	-1.113	-1.066	-1.115	-1.184	-1.174	-1.005	-0.935	-1.001
	(0.167)	(0.361)	(0.236)	(0.230)	(0.170)	(0.156)	(0.179)	(0.240)
ELA 3-8	-1.012	-0.990	-1.022	-1.083	-1.057	-0.980	-0.864	-0.886
	(0.201)	(0.283)	(0.159)	(0.272)	(0.126)	(0.139)	(0.119)	(0.229)
Observations	9	9	9	7	15	16	17	31

Panel D. Staffing climate

	Cohort 1				Cohort 2			
	Low	Mid	High	Missing	Low	Mid	High	Missing
Black	0.808	0.902	0.950	0.951	0.820	0.771	0.880	0.781
	(0.201)	(0.151)	(0.083)	(0.030)	(0.241)	(0.245)	(0.175)	(0.277)
Hispanic	0.092	0.041	0.017	0.018	0.078	0.113	0.022	0.079
	(0.118)	(0.097)	(0.023)	(0.009)	(0.174)	(0.212)	(0.040)	(0.133)
Economically disadvantaged	0.909	0.898	0.925	0.895	0.911	0.925	0.933	0.923
	(0.047)	(0.041)	(0.024)	(0.040)	(0.049)	(0.051)	(0.047)	(0.079)
English learner	0.039	0.037	0.007	0.012	0.058	0.104	0.031	0.042
	(0.076)	(0.085)	(0.009)	(0.011)	(0.133)	(0.172)	(0.082)	(0.095)
Special education	0.169	0.233	0.165	0.225	0.212	0.210	0.144	0.160
	(0.040)	(0.090)	(0.055)	(0.150)	(0.060)	(0.039)	(0.053)	(0.046)
Math 3-8	-1.009	-1.044	-1.201	-1.184	-1.147	-0.991	-0.949	-1.001
	(0.315)	(0.199)	(0.235)	(0.230)	(0.162)	(0.183)	(0.183)	(0.240)
ELA 3-8	-0.929	-0.966	-1.098	-1.083	-1.033	-0.923	-0.921	-0.886
	(0.211)	(0.199)	(0.184)	(0.272)	(0.152)	(0.148)	(0.128)	(0.229)
Observations	9	9	9	7	15	16	17	31

Panel E. School leadership

	Cohort 1				Cohort 2			
	Low	Mid	High	Missing	Low	Mid	High	Missing
Black	0.889	0.935	0.836	0.951	0.784	0.831	0.859	0.781
	(0.158)	(0.141)	(0.177)	(0.030)	(0.235)	(0.231)	(0.204)	(0.277)
Hispanic	0.052	0.024	0.073	0.018	0.106	0.063	0.041	0.079
	(0.115)	(0.053)	(0.099)	(0.009)	(0.199)	(0.168)	(0.097)	(0.133)
Economically disadvantaged	0.893	0.905	0.933	0.895	0.918	0.928	0.924	0.923
	(0.052)	(0.024)	(0.025)	(0.040)	(0.046)	(0.050)	(0.054)	(0.079)
English learner	0.031	0.004	0.048	0.012	0.094	0.046	0.051	0.042
	(0.074)	(0.005)	(0.085)	(0.011)	(0.161)	(0.129)	(0.111)	(0.095)
Special education	0.168	0.235	0.164	0.225	0.183	0.197	0.181	0.160
	(0.026)	(0.089)	(0.062)	(0.150)	(0.053)	(0.060)	(0.067)	(0.046)
Math 3-8	-1.186	-1.050	-1.063	-1.184	-1.021	-1.051	-0.991	-1.001
	(0.208)	(0.351)	(0.259)	(0.230)	(0.138)	(0.221)	(0.209)	(0.240)
ELA 3-8	-1.103	-1.004	-0.950	-1.083	-0.952	-0.965	-0.944	-0.886
	(0.232)	(0.202)	(0.178)	(0.272)	(0.141)	(0.168)	(0.145)	(0.229)
Observations	9	9	9	7	16	16	16	31

School-level means with standard deviations in parentheses. Missing values indicate no schools with the listed measure in a particular quantile. Values based on identification year, which is 2016-17 for Cohort 1 and 2017-18 for Cohort 2.