ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 19-44

Peeking into the Black Box of School Turnaround: A Formal Test of Mediators and Suppressors

Gary T. Henry
Vanderbilt University

Lam Pham
Vanderbilt University

Adam Kho
University of Southern California

Ron Zimmer
University of Kentucky

A growing body of research evaluates the effects of turnaround on chronically low-performing schools. We extend this research to formally test factors that either mediate or suppress the effects of two turnaround initiatives in Tennessee: the Achievement School District (ASD) and local Innovation Zones (iZones). Using difference-in-differences models within a mediational framework, we find that hiring highly effective teachers and employing effective principals partially explain positive effects of iZone interventions. In the ASD, high levels of teacher turnover suppress potential positive effects after the first year. In iZone schools, several factors suppress even larger positive effects: hiring more novice teachers; hiring more principals with less experience; and high levels of student chronic absenteeism and student in-migration.

VERSION: May 2019

Suggested citation: Henry, G.T., Pham, L., Kho, A., & Zimmer, R. (2019). Peeking into the Black Box of School Turnaround: A Formal Test of Mediators and Suppressors (EdWorkingPaper No.19-44). Retrieved from Annenberg Institute at Brown University: http://edworkingpapers.com/ai19-44

Peeking into the Black Box of School Turnaround:

A Formal Test of Mediators and Suppressors

Gary T. Henry, Lam Pham

Vanderbilt University

Adam Kho

University of Southern California

Ron Zimmer

University of Kentucky

April 25, 2019

Abstract: (116 words)

A growing body of research evaluates the effects of turnaround on chronically low-performing schools. We extend this research to formally test factors that either mediate or suppress the effects of two turnaround initiatives in Tennessee: the Achievement School District (ASD) and local Innovation Zones (iZones). Using difference-in-differences models within a mediational framework, we find that hiring highly effective teachers and employing effective principals partially explain positive effects of iZone interventions. In the ASD, high levels of teacher turnover suppress potential positive effects after the first year. In iZone schools, several factors suppress even larger positive effects: hiring more novice teachers; hiring more principals with less experience; and high levels of student chronic absenteeism and student in-migration.

Keywords: school turnaround, mediation, staff turnover, student mobility, chronic absenteeism

ACKNOWLEDGMENTS: We wish to thank the Achievement School District and Tennessee Department of Education for their feedback and assistance. Funding for this research was provided by John and Laura Arnold Foundation and the Walton Family Foundation. Any opinions or errors are solely attributable to us, however.

Introduction

Fueled by sustained policy interest in school reform, a growing body of research estimates the effects of school turnaround on persistently low-performing schools (Dickey-Griffith, 2013; Dougherty & Weiner, 2017; Gill, Zimmer, Christman, & Blanc, 2007; Harris & Larsen, 2016; Henry & Harbatkin, 2018; Schueler, Goodman, & Deming, 2017; Zimmer, Henry, & Kho, 2017), but these studies report a range of effects from positive to null to, alarmingly, negative results (Dickey-Griffith, 2013; Dougherty & Weiner, 2017; Henry & Harbatkin, 2018). As new school reform plans are being written and implemented across the nation under the Every Student Succeeds Act (ESSA), there is a sense of urgency to better understand why some prior reforms succeeded in turning around low-performing schools while others have failed.

As a requirement for receiving federal funds, recent efforts aimed at turning around the lowest performing schools include a mandate to disrupt the status quo in these schools, often by replacing the principal and most of the teachers (Dee, 2012; Zimmer, Henry, & Kho, 2017). By replacing school staff, these reforms intentionally increased staff turnover during the first year of reforms, but were mainly silent about the effects of ongoing instability from elevated rates of staff turnover in later years (Herman et al., 2008). At the student level, these turnaround reforms assume students would regularly attend school, but many not explicitly address student mobility or attendance that could hinder access to improved instructional practices.

In this study, we formally test the extent to which the effects of school turnaround can be explained by mediating factors that can be directly influenced by the reforms, such as staff turnover and student attendance. We also examine the quality of staff hired to fill vacancies. In doing so, this study builds upon a prior impact evaluation of two turnaround interventions in Tennessee: the Achievement School District (ASD) and local Innovation Zones (iZones) (Zimmer, Henry, & Kho,

2017). Data from the original study, which we draw upon to test mediation and suppression, showed that schools placed into local iZones increased student achievement, while ASD schools had no effect on student achievement. These data are ideal in that we have two distinct reform interventions, allowing us to examine the extent to which potential mediators explain positive effects in iZones and to formally test suppression for both interventions.

To test potential mediators, we use difference-in-differences models within the mediational framework originally proposed by Baron and Kenny (1986). This approach allows us to essentially test whether ASD and iZone interventions have significant indirect effects on student achievement through each potential mediator. Additionally, we report similar average mediation effects using simulation-based algorithms proposed by Imai, Keele, and Tingley (2010) and test the sensitivity of our results to potential omitted confounders that may be associated with both the mediators and the student achievement outcomes.

Our results across three years of Tennessee turnaround reforms show that iZone schools hired more effective teachers in the first year than in previous years. In the second and third years, principals in iZone schools were more effective than in similarly low-performing comparison schools receiving no turnaround interventions. These mediators explain between 10 and 38 percent of the positive effects in iZone schools. In ASD schools, teacher turnover rates were substantially higher than in comparison schools across all three years of reform, and more replacement teachers had fewer than three years of experience. These high rates of teacher turnover suppressed potentially positive effects from ASD interventions. Moreover, increased levels of student chronic absenteeism and student in-migration suppressed some of potentially larger positive effects of iZone schools. Overall, our results provide formal evidence that the ability to

stabilize schools by retaining experienced and effective staff substantially mediates the effects of school reforms.

The rest of this paper proceeds as follows. First, we review the relevant literature on school turnaround and present an overview of Tennessee's turnaround efforts. Then, we describe our methodological approach and results. We conclude with a discussion of how our findings inform a larger understanding of school reform, especially in the advent of new approaches to school reform under ESSA.

Review of the Literature on School Turnaround

For over half a century, whole-school reform efforts have been the subject of substantial policy interest and investment, with varying strategies for improving chronically low-performing schools (Aladjem et al., 2010; Berends, Bodilly, & Kirby, 2002; Egbert, 1981; Gross, Booker, & Goldhaber, 2009). Under No Child Left Behind, school reform experienced an era of remarkable federal activism with unprecedented federal resource investments including over $7 billion allocated to school improvement grants (SIGs; Carlson & Lavertu, 2018). In order to receive SIG funding, schools were required to implement one of four models: (1) transformation, which requires schools to replace the principal, change the curriculum, make student achievement a component of teacher evaluations, and use teacher evaluations to make personnel decisions; (2) turnaround, which requires replacing at least 50% of the staff in addition to all the requirements of the transformation model; (3) restart, which requires conversion into a charter or privately managed school; or (4) closure. Collectively and hereafter referred to as school turnaround, these four models have guided school reform efforts over the last decade.

In tandem with increased resource investments, studies evaluating the effects of school turnaround have proliferated in recent years, but results vary from positive (Bonilla & Dee, 2017;

Carlson & Lavertu, 2018; Dee, 2012; Harris & Larsen, 2016; Henry, Guthrie, & Townsend, 2015; Papay & Hannon, 2018; Schueler et al., 2017; Sun, Penner, & Loeb, 2017) to negative (Dickey-Griffith, 2013; Dougherty & Weiner, 2017; Henry & Harbatkin, 2018), to mixed (Heissel & Ladd, 2017; Strunk, Marsh, Hashim, Bush-Mecenas, et al., 2016; Zimmer, Henry & Kho, 2017). Overall, these studies suggest that turnaround interventions can potentially bring about school improvement, but there is not enough large-scale, quantitative evidence to identify the mediators that lead to success in some turnaround models and or the suppressors that lead to failure in others. Toward this end, qualitative and descriptive studies of school turnaround have provided valuable insight into a number of mechanisms that are likely mediators of turnaround (e.g., Glazer & Egan, 2018; Henry, Zimmer, Attridge, Kho, & Viano, 2014; Henry et al., 2017; Le Floch et al., 2016; Leithwood & Steinbach, 2003; Leithwood & Strauss, 2008; Malen, Croninger, Muncey, & Redmond-Jones, 2002; Rice & Malen, 2010; Scott, 2009; Strunk, Marsh, Hashim, & Bush-Mecenas, 2016). We contribute to this literature by formally testing potential mediators in order to inform a theory of action for school turnaround that makes efficient use of resource investments to support effective practices and address barriers to improvement.

The theory of action behind school turnaround is rooted in the perspective that persistently low-performing schools suffer from chronic instabilities that cannot be addressed through incremental changes, and instead require bold, school-wide interventions (Herman et al., 2008). As shown in Figure 1 below, the underlying theory of action for school turnaround begins with identifying persistently low-performing schools in order to signal a need for and commitment to change (Chiang, 2009). After identifying low-performing schools, school turnaround efforts are designed to bring about swift and dramatic changes, using policy tools such as school takeover and restructuring.

[Insert Figure 1 Here]

One common turnaround reform involves replacing school leadership and a significant portion of the instructional staff (Malen et al., 2002; Rice & Malen, 2010; Strunk, Marsh, Hashim, & Bush-Mecenas, 2016). After implementing dramatic changes to school staffing, the theory of action for school turnaround relies on the sense of urgency, motivation, and skill of new leaders and teachers to implement innovative practices that will lead to organizational improvement. Together, the sequence of identification, disruption, and organizational improvement are intended to produce improved student achievement. However, over a decade of implementation and research suggests that a number of barriers tend to threaten the effectiveness of turnaround under this theory of action.

In order to implement dramatic staff replacements, a pool of leaders and teachers must be available and willing to transfer into low-performing schools, but previous research has found that low-performing schools have difficulty attracting effective staff (Boyd, Lankford, Loeb, & Wyckoff, 2005; Guarino, Santibanez, & Daley, 2006; Le Floch et al., 2016; Redding & Henry, 2018). Some case studies of reconstituted schools find that new teachers in these schools tend to have less experience (Hamilton, Heilig, & Pazey, 2014; Hess, 2003; Malen et al., 2002; Rice & Malen, 2010). These findings highlight the need to investigate how the effects of turnaround are either driven or suppressed by the quality of staff recruited by these schools.

After completing initial efforts to hire replacements at the school and alter its management, ongoing staff turnover represents a second unanticipated obstacle that can impede the school improvement process. Multiple studies have reported a connection between ongoing staff turnover and diminished positive effects in turnaround schools, including in Philadelphia (Gill et al., 2007), North Carolina (Henry et al., 2015), and Los Angeles (Strunk, Marsh, Hashim, Bush-Mecenas, et

al., 2016). These findings suggest a fundamental tension within the theory of action for school turnaround where chronic instability in low-performing schools is addressed by initially creating more instability through personnel disruptions. The theoretical tension between intended and unintended disruptions shows that an examination of staff instability must attend to its timing within the turnaround process – initial personnel disruptions may be desired whereas ongoing personnel disruptions may act as barriers to improvement.

In addition to staff turnover, student mobility and attendance comprise another set of obstacles schools should address to create stability. Existing evaluations of school turnaround have begun examining how interventions affect student mobility with some descriptive findings to suggest that turnaround schools tend to experience elevated rates of student mobility (Dougherty & Weiner, 2017). Additionally, although student attendance has received less attention in the turnaround literature, previous work strongly suggests that student attendance generally has a positive relationship with achievement (Gottfried, 2010; Parke & Kanyongo, 2012). These findings support the theory that schools will have a difficult time making and sustaining improvement if students leave or are not in school to receive instruction.

Given our current understanding of school turnaround, this study makes a number of contributions to the turnaround literature. First, we build on the existing literature to identify and analyze potential mechanisms as either positive mediators or negative suppressors of improvement from turnaround interventions. This analysis helps shed light on the mixed findings from prior evaluations of school turnaround. Second, we illuminate an existing tension within the theory of action for school turnaround by examining personnel disruptions both as an intended intervention when turnaround is first initiated and as a possible barrier in later years when ongoing staff turnover is likely to impede school improvements. Finally, to the best of our knowledge, this study

presents the first formal examination of the extent to which various measures of instability at the principal, teacher, and student levels mediate the effects of school turnaround. We accomplish these goals in the context of school turnaround in Tennessee's ASD and iZones.

Background on Tennessee Turnaround

The Tennessee General Assembly created the ASD in 2010 as part of legislation called First to the Top (Public Chapter No. 2, 2010). After adopting this legislation and applying for the funding, Tennessee was awarded over $500 million in Race to the Top (RttT) funds. The proposal called for the State Commissioner of Education to identify the state's lowest-achieving five percent of schools, known as *priority schools*, and to turn around these schools using the four turnaround models (transformation, turnaround, restart, and closure) as defined by the U.S. Department of Education. With funding from RttT, Tennessee named 83 schools as priority schools eligible to receive turnaround interventions in 2012. Approximately 80 percent of these priority schools were located in Memphis, and all but two of the remaining schools were located in either Nashville or Chattanooga. Between 2012-13 and 2014-15, Tennessee's priority schools were either placed into the ASD, became part of a local district iZone, closed, or continued operating with neither ASD nor iZone interventions. Schools were chosen to become part of the ASD through a combination of criteria including the schools' feeder patterns and matches between the school and CMOs eligible to operate the school (ESEA Flexibility Request, 2012). Districts could then choose to place priority schools that were not selected for the ASD (or closed) into a local iZone.

As the state's boldest and most controversial reform strategy, the ASD applied the federal restart model by requiring that school governance be transferred from the local district to the state and that school management be shifted to another entity such as a CMO or the ASD itself, which

were granted autonomy in operating these schools (ESEA Flexibility Request, 2012). It is important to note that the ASD schools remained neighborhood schools, not schools of choice with which many of the CMOs had much more experience. The ASD model has received significant policy attention, and at least four other states have adopted or attempted to adopt this model to restart schools under the auspices of the state and managed by a CMO or other private organization (Delaney, 2017; Downey, 2016; Granados, 2017; Riley, 2015).

Local iZones, on the other hand, placed schools into an intra-district network that received increased funding, support, and autonomy. As part of the iZone, schools were managed by a separate organizational structure composed of local district staff dedicated to supporting iZone schools. Schools joining local iZones were also required to replace the principal and at least a portion of the instructional staff. The iZones also implemented the state's requirement to offer financial incentives for effective teachers and principals to transfer into and remain in these schools. During the study period, local iZones were created in three Tennessee districts: Shelby County Schools District (Memphis), Metropolitan Nashville Public Schools, and Hamilton County Schools (Chattanooga). Throughout our analysis, we aggregate across all three iZones, but restricting analyses to only the largest iZone (Memphis) produced substantively similar results.[1]

Methods

Data. Data for the analysis were provided by the Tennessee Department of Education and managed by the Tennessee Education Research Alliance. These data link students, teachers, and principals with a school in each year from 2009-10 to 2014-15, allowing us to identify when a student, teacher, or principal moves to a different Tennessee school. The teacher and principal-level data include demographic and professional characteristics, including value-added measures of effectiveness on Tennessee's value-added assessment system (TVAAS scores), observation

ratings, years of experience, and highest degree earned. The student-level data include demographic variables, attendance information, and test scores.

Our analytical sample is restricted to students in tested subjects and grades. In Tennessee, students in grades 3-8 take the Tennessee Comprehensive Assessment Program (TCAP) exams in reading, math, and science, and in later grades, students take end-of-course (EOC) exams after completing courses in seven subjects: English I, English II, English III, Algebra I, Algebra II, Biology, and Chemistry. Given these two testing systems, we standardize the statewide data for TCAP scores by subject, year, and grade, and EOC scores by subject, year, and semester.

Our study spans the time period between 2010-11 and 2014-15, which encompasses two years of pre-turnaround data and three years of post-turnaround outcomes: 2012-13 through 2014-15. Since both the ASD and iZones targeted additional schools in each post-turnaround year using a cohort model, we code the year indicators such that the first year after turnaround is 2012-13 for the first cohort, 2013-14 for the second cohort, and 2014-15 for the third cohort. Years two and three of turnaround interventions are coded following the same logic. This coding is preferred because of the disruption intended for the first year of turnaround reforms.

Comparison schools are identified using Tennessee's 2012 list of 83 priority schools (i.e., the lowest performing five percent of schools in the state). Listed in Table 1 below, the comparison sample is composed of priority schools that were not included in either the ASD or an iZone between 2012-13 and 2014-15.

[Insert Table 1 Here]

As shown in Table 1, both the ASD and iZones took over an increasing number of schools in every year across three cohorts. We note that in years two and three, the ASD began operating new-start schools that did not exist previously and were not part of the Tennessee's

priority list. We do not include these schools in our analysis. Moreover, we exclude any schools where we did not have scores for students in the years prior to turnaround. We also exclude schools that did not have tested grades in the relevant years, such as schools with only K-2 grades. By the end of the study period (2014-15), 28 priority schools that received no ASD or iZone interventions remained, which will serve as our comparison group.

Measures. Our potential mediators are proximal outcomes that are likely to be affected when a school implements turnaround reforms that lie in the pathway between these turnaround interventions and student achievement. At the student-level, we test measures of student mobility and chronic absenteeism. Students are characterized as mobile in years when they transfer to a different school, except in cases when the student makes an expected, structural move due to changing school levels. For example, students are not considered mobile if they move from an elementary to a middle school after completing the final grade offered at the elementary school. Students are considered chronically absent if they miss more than 10 percent of the instructional days during the time they are enrolled in the school. We use this definition of chronic absenteeism that aligns with Tennessee's policy for using student attendance as part of district and school accountability under ESSA. Other student characteristics used as covariates in our models include prior year achievement scores, gender, race, eligibility for free or reduced-price lunch (FRPL), English language learner status (ELL), and special education status (SpED).

Our primary teacher-level mediator is a measure of teacher turnover, where turnover is a dichotomous indicator equal to one if teachers' current school is not the same as their school in the previous year. Following Ronfeldt (2015), we aggregate this teacher-level indicator to the school-level to obtain a measure of teacher turnover operationalized as the proportion of teachers currently in the school who were not there in the previous year. As part of our effort to

understand the quality of teachers that are recruited to ASD and iZone schools, we also examined teachers' TVAAS scores and years of total teaching experience. Tested subject teachers in Tennessee receive TVAAS scores ranging from 1 to 5, where 1 indicates not effective and 5 indicates highly effective. We examine the recruitment of effective teachers with a variable that captures the proportion of tested teachers who were new to the school and received a TVAAS score of four or five in their previous school. In order to examine teacher experience, we use the proportion of teachers who are both new to the school and have three or fewer years of total teaching experience (i.e., a novice teacher).[2] We examine novice teachers because previous research has found that teachers are at their lowest level of effectiveness in their first three years (Henry, Bastian, & Fortner, 2011). Other characteristics of teachers include degree attainment (an indicator for attaining a master's degree or above), total years of experience, gender, and race.

Principal-level mediators mirror those for teachers including a principal turnover indicator equal to one if the current principal is different from the principal in the previous year. We also examined measures of whether the current principal has fewer than three years of experience as a principal and a measure of the principal's overall observation score, standardized statewide by year. Other characteristics of principals include years of experience as an educator, gender, race, and an indicator for whether the principal has fewer than three years of experience as a principal.

Mediational Framework. In order to test for mediation, we use a series of three regression equations to separately test each mediator, as shown in Figure 2 below. These models follow a framework outlined in Baron and Kenny (1986). In all three models, $Turnaround$ represents either ASD or iZone interventions.

[Insert Figure 2 Here]

Path coefficient c in the first model estimates the overall effect of ASD or iZone turnaround interventions on student test scores. Path coefficient a in the second model estimates the effect of ASD or iZone interventions on the mediator. Path coefficient b in the third model estimates the association between the mediator and student test scores, and path coefficient c' in model three reveals the direct effect of ASD and iZone interventions on student test scores after controlling for the mediator (Baron & Kenny, 1986; Preacher, 2015).

Comparing the relationships between paths $a, b, c,$ and c' allows for assessing mediation. There is evidence that the mediator explains some of the effects of the reforms if (1) the turnaround interventions have an effect on the mediator (i.e., path a is statistically significant); (2) the mediator has a relationship with student test scores (i.e., path b is statistically significant); and (3) the direct effect of turnaround interventions on student test scores changes after controlling for the mediator (e.g., c' is meaningfully different from c). If all three conditions are met, the mediator can be described as having either a positive, mediating role or a negative, suppressing role. The mediator has a positive influence if controlling for it decreases the magnitude of path c (i.e., $c - c' > 0$). Full mediation occurs if c is statistically significant and c' is not, because the mediator completely explains the effect of turnaround interventions on student test scores. Partial mediation occurs if $c - c' > 0$ but c' continues to be statistically significant, because the mediator will have explained some, but not all, of the effect of turnaround interventions. On the other hand, the mediator has a suppression effect if controlling for it increases the value of path c (i.e., $c - c' < 0$). For example, suppression occurs when turnaround interventions have a positive effect on the mediator, but the mediator has a negative relationship with student test scores. Under this framework, we can obtain a descriptive measure of the

indirect effect of turnaround interventions on student test scores through the mediator by directly multiplying the a and b paths.

We use the classic mediational framework prescribed by Baron and Kenny (1986), because of our large sample sizes, the descriptive nature of our research questions, and preferred estimation strategy. However, below we show that our substantive conclusions remain unaltered when using more modern, simulation-based algorithms for obtaining average mediational effects (Imai, Keele, & Tingley, 2010). We also provide sensitivity analyses to help quantify the robustness of our estimates to various assumptions embedded in this mediational framework.

Estimation Strategy. We estimate the three mediational models for each mediator, separately comparing ASD schools and iZone schools with similarly low-achieving, comparison schools that did not receive any systematic intervention. First, we estimate Model 1 below to establish the effect of ASD and iZone interventions on student test scores. In an ideal situation, we would randomly assign students and schools to receive ASD interventions, iZone interventions, or to an untreated comparison condition. This sort of random assignment is not feasible in this context, so we implement what we consider to be the next best approach – the difference-in-differences (DID) model. First, the DID model takes the average difference between student test score gains in the post-turnaround years with average student test score gains the pre-turnaround (i.e., baseline) years. Then, the model compares this pre-post difference in ASD and iZone schools with the same difference in comparison schools. This model relies on the key identifying assumption that, conditional on covariate adjustment, student test score gains in either ASD or iZone schools from pre-intervention to post-intervention would have differed by the same amount as average test score gains in comparison schools over the same period had they not received any turnaround interventions. Model 1 follows a similar approach to the

existing study evaluating the effect of ASD and iZone interventions (Zimmer, Henry, & Kho, 2017). That is, we model student test scores y, for student i in grade g, school s, subject c, and time t, such that:

$$y_{igsct} = \beta_0 + \sum_{t=1}^{3} \beta_{1t} After_{gst} + \sum_{t=1}^{3} \beta_{2t} Turnaround_s * After_{gst} + \mathbf{X}_{igsct} + \mathbf{\Gamma}_{st} + \boldsymbol{d}_s + \boldsymbol{p}_g + \boldsymbol{k}_c + \varepsilon_{igsct} \quad \text{(Model 1)}$$

where $After$ includes three dichotomous indicator variables, one for each of the three years after schools joined either the ASD or an iZone. The $After$ variable varies across grades within a school because ASD schools took over schools using a phase-in process where lower grades were taken over by the ASD in the first year and upper grades were added in later years. **X** is a vector of student characteristics including a prior year measure of student achievement, gender, race, free-reduced price lunch (FRPL) eligibility, English language learner (ELL) status, and special education (SpED) status. **Γ** is a vector of school-level time-varying covariates including the proportion minority, proportion ELL status, proportion SpED status, and proportion of FRPL eligible students. Finally, $\boldsymbol{d}_s$ is a vector of school indicators; $\boldsymbol{p}_g$ is a vector of grade indicators; and $\boldsymbol{k}_c$ is a vector of subject indicators. The model estimates effects for each of the three post-turnaround years, with β_{2t} as the coefficients of interest for the overall effect of turnaround reforms in each of the three years. In the parlance of our mediational framework, the β_{2t} coefficients in Model 1 estimate path c for each post-turnaround year. These coefficients can be interpreted as pre and post achievement gains in ASD or iZone schools relative to the pre and post achievement gains in comparison schools.

We apply Model 1 separately on two subsamples. The first subsample contains only ASD schools and comparison schools, and the second sample contains only iZone schools and

comparison schools.[3] We estimate the same model on both subsamples such that $Turnaround$ is an indicator for whether the school was in ASD during the study period for the first subsample. In the second subsample, the $Turnaround$ variable is an indicator for whether the school was in an iZone during the study period. We run models on two separate samples to allow for the mediators to have different effects in ASD schools and iZone schools. For example, teacher turnover in some schools may be a positive mediator if newly recruited teachers are effective and replace low-performing teachers, but teacher turnover may be a suppressor if newly recruited teachers are lower performing than the teachers they replace. We discuss these issues further below. All standard errors are clustered at the school level.

Second, we estimate the effect of turnaround interventions on each mediator using a similar DID model that replaces the outcome with the mediator.

$$M_{igst} = \mu_0 + \sum_{t=1}^{3} \mu_{1t} After_{gst} + \sum_{t=1}^{3} \mu_{2t} Turnaround_s * After_{gst} + \mathbf{X}_{igst} + \mathbf{\Gamma}_{st} + \boldsymbol{d}_s + \boldsymbol{p}_g + \varepsilon_{igst} \quad \text{(Model 2)}$$

Model 2 establishes the effect of ASD or iZone turnaround interventions on the mediator, such that the μ_{2t} coefficients estimate path a in each year of turnaround implementation. Model 2 includes the same set of covariates as Model 1 with the exception of subject indicators, because the mediators are not subject-specific measures. Our preferred results report path a from model 2 estimated at the student level in order to facilitate comparisons with model 1, and we cluster standard errors at the school level. However, since the teacher and principal-level mediators are measured at the school level, we also estimate model 2 at the school level for these mediators in

Appendix Table 1 and obtain similar estimates of path a, leading to the same substantive conclusions.

Third, we adjust for the mediator's relationship with student test score gains.

$$y_{igsct} = \beta_0' + \sum_{t=1}^{3} \beta_{1t}' After_{gst} + \sum_{t=1}^{3} \beta_{2t}' Turnaround_s * After_{gst} + \beta_3 Mediator_{igst} + \mathbf{X}_{igsct} + \mathbf{\Gamma}_{st} + \boldsymbol{d}_s + \boldsymbol{p}_g + \boldsymbol{k}_c + \varepsilon_{igst} \quad \text{(Model 3)}$$

In model 3, the β_{2t}' coefficients now estimate path c' for each post-turnaround year and β_3 estimates path b. Comparing the relationship between c and c' using the β_{2t} coefficient from Model 1 and the β_{2t}' coefficient from Model 3 shows whether the mediator has a positive or suppressing influence on the effect of turnaround interventions.

We note that these models explicitly recognize that variation in mediators is not exogenous. The DID model leverages exogeneous pre-post differences in turnaround versus comparison schools in order to estimate the effect of turnaround interventions on student achievement gains. However, variation in each of the mediators is theorized to be dependent on the turnaround interventions. For example, in the first year of turnaround, teacher turnover rates are expected to be higher in ASD and iZone schools due to explicit staff replacement efforts. Therefore, the effect of the teacher turnover on student test scores is not intended to be a causal impact estimate. Rather, we are more interested in describing how the causal effect of ASD or iZone turnaround interventions *change* after controlling for each mediator separately.

Results

Before turning to results from the mediation models, Table 2 presents descriptive statistics comparing ASD, iZone, and the comparison (non-ASD, non-iZone) priority schools in the baseline year (year 0) before turnaround and each year after reforms began. Table 2 shows

that in the baseline year, school compositional characteristics are similar between ASD, iZone, and comparison schools. Appendix Table 2 shows results from *t*-tests comparing the compositional characteristics of ASD and iZone schools with comparison schools in the baseline year, and shows no statistically significant differences. These compositional characteristics include the proportion of students who are female, minority race, FRPL eligible, ELL status, SpED status, chronically absent, and new to the school. The descriptive results show that all schools in the sample are low-performing, with average test scores that range between -0.84 and -1.13 standard deviation units in the baseline year. Moreover, the priority schools in our sample primarily serve low-income, minority race students throughout the entire study period. Student enrollment also remains relatively stable across the three years of implementation for ASD, iZone, and comparison schools. Stable enrollment numbers provide evidence that staff turnover in these schools are not an artifact of systematic increases or decreases in school size. Similar descriptive characteristics between ASD, iZone, and comparison schools in the baseline year also support our identifying assumption that turnaround schools would have been similar to comparison schools had they not received any turnaround interventions.

[Insert Table 2 Here]

Teacher turnover, defined as the proportion of teachers new to the school, differs descriptively between ASD and iZone schools. To illustrate, Figure 3 below shows that trends in teacher turnover rates in both ASD and iZone schools were similar to those in comparison schools prior to turnaround but increased dramatically in the first year of reforms. About 64 percent of teachers in ASD schools and 54 percent of teachers in iZone schools were new to the school during the first year of turnaround, compared to about 26 percent in the baseline year and about 24 percent in comparison schools. These results are unsurprising because replacing

teachers is an intentional part of turnaround efforts in the first year. After the first year, however, teacher turnover continues to be high in ASD schools (55-57 percent), whereas turnover in iZone schools drops to levels that are more similar to turnover in comparison schools (between 33 and 35 percent). Together with the results from Table 2, these descriptive averages suggest that ASD schools were faced with elevated teacher turnover rates well after intentional efforts to replace staff in the first year. Also, ASD and iZone schools both appear to have been successful in recruiting teachers who have higher TVAAS scores, but newly recruited teachers tended to also have fewer years of experience.

[Insert Figure 3 Here]

Table 2 also shows that principal characteristics mirror many of the same trends as the teacher-level variables. Principal turnover increased in the first year of turnaround in both ASD and iZone schools, but principal turnover rates in iZone schools decrease after the first year, whereas turnover rates remain high in ASD schools. Principals recruited to ASD and iZone schools tended to have fewer years of experience and somewhat higher observation scores.

Mediation results. Across all mediators, estimating the first mediational model (without including the mediator) replicates findings from the previous evaluation of ASD and iZone interventions, which concluded that iZone schools produced positive and significant gains in student achievement, whereas ASD schools did not gain more or less that non-ASD, non-iZone comparison schools (Zimmer, Henry, & Kho, 2017).

Table 3 below shows results from models testing characteristics of teachers as mediators. Model 2 confirms the descriptive trends in teacher turnover, showing that both ASD and iZone schools experienced increases in the proportion of teachers new to the school during the first year of turnaround. However, in the second and third years, the effect of iZone interventions on

teacher turnover is much smaller in magnitude and no longer statistically significant, while turnover rates in ASD schools are nearly 30 percentage points higher than in comparison schools. Model 3 shows that teacher turnover is associated with a reduction in student test score gains in ASD schools of 26 percent of a standard deviation unit (SDU), but not in iZone schools. Moreover, controlling for teacher turnover increases the ASD effect, resulting in positive and statistically significant effects of 14 and 18 percent of an SDU in years 2 and 3, respectively. Together, these models suggest that ASD interventions would likely have led to detectable gains in student achievement if these schools had not continued to experience such high rates of teacher turnover in years 2 and 3.

Although teacher turnover appears to have been an obstacle to improvements in ASD schools, changing the composition of teachers in a school is not necessarily detrimental if the newly recruited teachers are more effective than the teachers they replace. To test this idea, we examined the mediational influence of teachers who are new to the school but have a prior-year TVAAS score of four or five, out of five. While both ASD and iZone schools experienced increases in the proportion of new teachers who are high performing in their first year, the effect of the iZone intervention is reduced from 15 percent of an SDU to about 9 percent of an SDU after controlling for hiring effective teachers in these schools. These results show that about 40 percent of the positive gains experienced in the first year of iZone implementation can be explained by hiring high performing teachers.

In addition to examining new teachers with high TVAAS scores, we also tested the proportion of new teachers with fewer than three years of teaching experience prior to entering these schools. The results suggest that, like overall turnover, ASD and iZone schools both experienced increases in the proportion of new teachers who are novices in the first year of

turnaround implementation, but only ASD schools continued to experience higher rates of incoming novice teachers after the first year. The results also show that recruiting more novice teachers is negatively associated with student test scores and significant only in ASD schools. Moreover, these findings suggest the effects of ASD interventions would have resulted in larger but only marginally significant gains in student test score gains in years 2 and 3 had the ASD reduced the influx of novice teachers.

[Insert Table 3 Here]

Turning to hypothesized principal mediators, Table 4 shows the mediating influence of principal turnover, experience, and effectiveness. The results for principal turnover are similar to the trends for teacher turnover. That is, ASD schools experienced elevated rates of principal turnover in all three years, which suppressed marginally significant positive effects of 12 to 14 percent of an SDU in years 2 and 3. The iZone schools experienced somewhat elevated rates of principal turnover in the first year, but the turnover did not have a suppressing influence on the effect of iZone schools. The mediating influence of principal experience is only marginally significant in ASD schools and insignificant in iZone schools, but the point coefficients suggest that principals in ASD schools had a higher probability of being inexperienced, especially in years 2 and 3, and suggesting that principals with less experience had a negative but only marginally significant association with test score gains in ASD schools. Finally, principal observation scores were used to measure principal effectiveness in ASD and iZone schools, and including the observation scores in the models show that average principal effectiveness increased in iZone schools in years 2 and 3, but did not change in ASD schools (with some marginal evidence that principal effectiveness was lower in ASD schools compared to the other Priority schools in the first year). Increased principal effectiveness in iZone schools in years 2

and 3 provide suggestive evidence that principal effectiveness was a mediator in iZone schools, because 13 and 24 percent of the positive effect of iZone schools in years 2 and 3, respectively, was explained by principal effectiveness.

[Insert Table 4 Here]

Table 5 below shows results for student-level mediators. Model 2 finds that students in iZone schools have a positive and significantly increased probability of being chronically absent. The point coefficients show a similar increase in ASD schools but are only marginally significant in years one and three. Results from model 3 show that chronically absent students in both ASD and iZone schools tend to receive lower test scores, and the effect of turnaround interventions increase after controlling for chronic absenteeism. These results suggest that chronic absenteeism is a slight suppressor of the turnaround effects in iZone schools, because average student achievement gains would have been larger by about 1 percent of an SDU had chronic absenteeism not increased in iZone schools. In ASD schools, the mediating influence of chronic absenteeism follows similar trends, but are only marginally significant. Similar to chronic absenteeism, student mobility also suppresses a modest amount, approximately one half of a percent of a standard deviation unit, of the effects in iZone schools in the second and third years, but not in ASD schools. Together, these results provide suggestive evidence that chronic absenteeism and student mobility very slightly suppressed positive effects in iZone schools.

[Insert Table 5 Here]

To summarize our results, Table 6 below presents a measure of effect size for the partially standardized indirect effect of turnaround interventions through each mediator, obtained from multiplying the estimated a and b paths for each mediator. The partially standardized indirect effect is more suitable is this context than a fully standardized indirect effect because our

turnaround "treatment" indicators are binary, making the fully standardized indirect effect size more difficult to interpret. The indirect effect is partially standardized, because the outcome variable (student test scores) is standardized. This measure of effect size is interpreted as the number of SDUs by which test scores are expected to change indirectly through the mediator when the turnaround indicator is changed from 0 to 1. Another way to interpret this effect size is the change in the effect of the turnaround interventions on student test scores (in SDUs) when the mediator is added to the model (Preacher & Kelley, 2011).

In order to make inferences regarding these effect sizes, we use bias-corrected bootstrap 95% confidence intervals (CI). To obtain bootstrap samples, we draw from the original data with replacement and estimate $\hat{a}$ and $\hat{b}$ in each bootstrap sample. Note that we use cluster sampling of schools with the number of clusters equal to that of the original sample in order to preserve the nature of our data with students clustered within schools. We repeat this process 5,000 times and construct confidence intervals by finding the estimates of ab corresponding to the 2.5th and 97.5th percentiles of the bootstrap sampling distribution. Using this same distribution, the bias-corrected CI was calculated using the bias adjustment described by MacKinnon (2012), Hayes (2017), and Preacher & Hayes (2004). This nonparametric bootstrapping approach has the advantage of not requiring assumptions about the shape of the distribution of variables and the sampling distribution of the ab statistic but can produce CIs that are slightly asymmetric. Some researchers (Hayes & Scharkow, 2013) point out that bias-corrected bootstrap CIs may suffer from Type I error inflation, so we also calculated percentile bootstrap CIs, which are nearly identical for these data (see Appendix Table 3).

Table 6 quantifies the indirect effect sizes corresponding to the patterns we observe from our difference-in-differences mediation models. First, high levels of teacher turnover in all three

years suppressed the effects of ASD interventions. For example, the indirect effect of ASD interventions on student test scores through teacher turnover in the first year of implementation is -0.11 SDUs. Moreover, the suppression effects of high teacher turnover in the ASD appears to be largely driven by the hiring of novice teachers in all three years. Additionally, the indirect effect of ASD interventions on student test scores via high rates of principal turnover is about -0.032 SDUs and -0.053 SDUs in years 1 and 2, respectively. The indirect effect of iZone interventions on student test scores via hiring effective teachers is 0.056 SDUs in the first year of implementation, whereas the indirect effect of iZone interventions via principal effectiveness is 0.023 SDUs (year 2) and 0.033 SDUs (year 3). Finally, higher rates of chronic absenteeism and student in-migration appears to slightly suppress some effects of iZone interventions. For example, the indirect effect of iZone reforms in the third year on student achievement through chronic absenteeism is -0.009 SDUs and through student mobility is -0.007 SDUs.

Validity Checks and Alternative Specifications

The validity of the DID model has been thoroughly examined in previous work evaluating the effects of ASD and iZone interventions (Zimmer et al., 2017). These validity checks include testing for anticipatory or announcement effects, a Granger test for pre-treatment differences, checking that turnaround interventions did not change school compositional characteristics, testing alternative comparison groups, and checking that the results are robust to a number of different model specifications (e.g., including vs. excluding various covariates). Our DID model passes all of these validity tests and we do not reproduce them here to conserve space since results are available in a prior publication (Zimmer et al., 2017).

Since our main results are aggregated across all three subjects, we include an examination of the parallel trends assumption when test scores are aggregated across all three

tested subjects. The DID model relies on the assumption that trends in achievement in turnaround schools would have been similar to those in the comparison schools in the absence of treatment. While this counterfactual is unobservable, we graphically examine the parallel trends assumption in Figure 4 below.

[Insert Figure 4 Here]

Figure 4 graphs average standardized student test scores for all subjects and all grades at the school-level for six years prior to turnaround interventions. Consistent with our main analysis, we center the year variable such that year 0, the baseline year, is the year prior to the beginning of turnaround interventions (i.e., the baseline year will differ depending on when ASD or iZone interventions began in each school). To the extent that pre-treatment trends are good approximations of what the counterfactual post-treatment trends would have been in the absence of treatment, Figure 4 shows that test score trends in ASD and iZone schools are similar to test score trends in comparison schools throughout the pre-turnaround period. Test score trends in schools that will eventually join the ASD and iZone do not cross with the trend in comparison schools in any pre-turnaround year. To further support the parallel trends assumption, our descriptive results above show that average student characteristics in ASD and iZone schools in the baseline year do not differ from those of comparison schools.

We note that another statistically sound approach to mediation analysis uses structural equation models (SEM) to simultaneously estimate the effect of turnaround interventions on the mediator and student test scores. We find that SEM estimates are nearly identical to our ordinary least squares (OLS) estimates. This is likely due to our relatively large sample size and use of cluster-robust standard errors. For parsimony, we only report estimates from our OLS models, but our SEM estimates are available upon request.

Additionally, our preferred estimates allow the effect of turnaround interventions on each mediator (path a) to vary by year but estimates one value for the average relationship between the mediator and student test scores across all three years (path b). However, just as the effects of path a can differ across years, path b may also differ across the three years. To allow for this possibility, we estimate an alternative version of the third mediation model where the mediator is interacted with indicators for each year after turnaround reforms are implemented. We find that indirect effect estimates are very similar when we include these interactions and allow path b to vary across each year (see Appendix Table 4).

Another possible threat to the validity of our estimates occurs if high teacher turnover is driven by school enrollment changes after the school begins turnaround reforms (e.g., if parents move their students out the turnaround school). Our balance checks suggest that student enrollment does not change significantly when schools enter the ASD or an iZone, nor does enrollment change significantly across the three years of implementation. However, we also test models where we include log student enrollment as a covariate and reach similar conclusions (see Appendix Table 5).

Sensitivity to the Sequential Ignorability Assumption

While our mediation findings rely on endogenous variation of the mediator, they do help to explain the relationships between certain proximal outcomes (e.g., recruiting highly effective teachers) and student test score gains reported in prior causal work (Zimmer et al., 2017). To increase the rigor of our findings, we follow existing methodological recommendations to test the sensitivity of our results to potential omitted confounders (Imai, Keele, Tingley, & Yamamoto, 2011). That is, mediational analyses rely on the two successive assumptions of sequential ignorability. First, it assumes that assignment to turnaround interventions are

independent of potential outcomes and potential mediators, given observed covariates. This first assumption is the conventional assumption of exogeneity in the turnaround "treatment." Previous work has shown that this assumption is plausible given our difference-in-difference model and extensive tests for potential confounding factors (Zimmer et al., 2017). The second assumption under sequential ignorability implies that the observed mediators are also exogeneous given the observed turnaround treatment status and observed covariates.

The second assumption under sequential ignorability is untestable and cannot be fully ruled out even under experimental conditions where both the turnaround treatment status and mediators are randomly assigned (see Imai, Keele, Tingley, and Yamamoto, 2011 for a deeper discussion of this issue). Nevertheless, we use a sensitivity analysis to quantify the extent to which our results rely on the assumption of ignorable mediators. To do so, we follow methods developed by Imai, Keele, and Tingley (2010) that have been applied in prior research investigating mediation (e.g., Desmond & Travis, 2018; Pedulla, 2016; VanderWeele, 2015). First, we estimate average mediation effects using the generalized approach proposed by Imai and colleagues (2010).[4] The simulation-based strategy proposed by Imai et al. (2010) is computationally intensive (we run 1000 simulations for each mediator to obtain approximations of parameter uncertainty).[5] However, this method uses the potential outcomes framework to estimate the change in the outcome corresponding to a change in the mediator as a result of treatment, all while holding constant the treatment status. By holding the treatment status constant and allowing the mediator to vary, this approach isolates the process where treatment affects the outcome through the mediator. In Table 7, we show average mediation effects estimated using methods developed by Imai et al. (2010). These estimates of the average mediation effect lead to the same substantive conclusions as our preferred results reported in

Table 6. These average mediated effects differ from our estimates by less than one-tenth of a standard deviation unit.

[Insert Table 7 Here]

Next, we quantify the extent to which sequential ignorability must be violated before our conclusions are reversed. This sensitivity analysis is based on the correlation, ρ, between the error term for the mediator model and the error in the outcome model. If sequential ignorability holds, then all confounders will have been controlled and ρ equals zero. Thus, our findings are sensitive if the average mediation effects vary widely as a function of ρ. Although the true value of ρ is unknown, it is possible to calculate the value of ρ where the average mediation effect equals zero. We report these values of ρ in Table 8 below.

Although ρ quantifies the degree of sensitivity, it is difficult to interpret substantively (e.g., "large" versus "small" values of ρ). To aid with this interpretation, Table 8 also expresses the sensitivity of our results in terms of a change in R^2 for both the mediator and outcome models. This alternative formulation of the sensitivity parameter can be interpreted as the proportion of total variance in the mediator and outcome explained by a hypothetical unobserved confounder. Using this parameter, our findings would be considered sensitive if a hypothetical unobserved confounder needs to explain only a small portion of the variance in the mediator and outcome for the average mediation effect to become indistinguishable from zero. Table 8 below shows these sensitivity parameters for all mediators for both the ASD and iZones. The results suggest that turnover among effective teachers and novice principals are the most sensitive to potential confounding from an omitted variable, whereas student chronical absenteeism and mobility are relatively less sensitive.

[Insert Table 8 Here]

For example, in order for the average mediation effect of teacher turnover in ASD schools to become indistinguishable from zero, there must be an unobserved confounder that both increases teacher turnover and decreases student test scores such that the magnitude of the correlation between the two error terms is greater than 0.035. Interpreted as a change in R^2, an unobserved confounder would render the average mediation effect statistically insignificant if the product of the proportion of variance explained in student test scores and the proportion of variance explained in teacher turnover is greater than 0.0005.

Although a proportion of 0.0005 for an omitted confounder appears quite modest, we note that student and school characteristics commonly available in education research commonly explain very little of the variance in student achievement outcomes. Moreover, the proportions shown in Table 8 are further diminished because they are products of two R^2 values that are less that one. To put these results in context, we examine the proportion of variance explained in the mediator and outcome for each of the covariates that we can observe and currently include in our models because previous research has shown them to be important confounding variables. For example, FRPL eligibility would be a relevant confounder if it were not included as a covariate, because prior research suggests that teachers are more likely to transfer out of schools with higher proportions of FRPL eligible students and being eligible for FRPL is negatively associated with student test scores.

When we regress student test scores on FRPL eligibility, we find that $R^2 = 0.0017$, suggesting that FRPL eligibility explains 0.17 percent (i.e., less than 1 percent) of the variation in test scores. Likewise, $R^2 = 0.000011$ when we regress teacher turnover on FRPL eligibility. Therefore, the proportion of total variance in student test scores and teacher turnover explained by FRPL eligibility is 0.0000000181 (0.0017 * 0.000011). Compared with the sensitivity

parameters list in Table 8, this calculation suggests that a potential omitted confounder must explain many times more of the variation in test scores and teacher turnover than FRPL eligibility for the average mediating effect of teacher turnover to become statistically insignificant. In fact, none of the observed covariates included in our model, which prior research has shown to be predictive of both student achievement and teacher turnover, can explain enough of the variation in both to render the average indirect effect of teacher turnover statistically insignificant. Appendix Table 7 shows R^2 values from (1) regressing student test scores on each observed covariate; (2) regressing of each mediator on each covariate; and (3) the product of (1) and (2). Comparing these results with Table 2, we conclude that our results are robust to potential unobserved confounders, because these confounders must explain a relatively large proportion of the variance in both test scores and each mediator compared to our observed covariates.

Discussion and Conclusion

This mediational analysis suggests that hiring effective teachers in Year 1 and staffing the schools with effective principals in Year 2 and 3 partially explain the positive effects of the iZones on gains in student test scores. However, hiring comparatively more new novice teachers and new novice principals in Year 1 appear to have suppressed immediately larger positive effects. Also, throughout the first three years of implementation of the iZones, chronic absenteeism among students and in-migration slightly suppressed even larger effects of the iZones on student test scores gains. The mediation and suppression analysis of the iZone schools does provide evidence that, in the lowest performing schools, hiring effective teachers and staffing these schools with effective principals are important ingredients for improving the performance of these schools. This suggests that recruitment and retention of effective staff

should be a main focus of plans to improve the lowest performing schools. In addition, our findings support a need to address student transfers and missed class time in these reform plans.

For the ASD, the proximal outcomes and, therefore, the factors suppressing effects were different than for iZones. In the second and third year of operation, when turnover could have been expected to revert to levels before the schools were placed in the ASD, teacher turnover was roughly 30 percentage points higher in the ASD schools and principal turnover was 84 percentage points higher in the second year and 56 percentage points higher in the third year than in the comparison schools. Our results suggest that ASD schools may have been able to produce positive gains in student achievement in the second and third years of implementation had they not experienced ongoing disruptions from continually replacing teachers and principals. The current theory of action for school turnaround often features replacing teachers and principals in the first year of turnaround in an attempt to disrupt the status quo; however, it appears that the continuation of staff instability after the initial replacement policies were implemented undermined positive outcomes. Our results suggest that turnaround efforts should attend to retaining effective teachers and principals after they are recruited to the turnaround school in the first year. The finding that ongoing, high levels of turnover suppresses positive effects from turnaround is unsurprising given the existing evidence that finds teacher turnover has a negative effect on student achievement (Hanushek, Rivkin, and Schiman, 2016; Henry & Redding, 2018; Ronfeldt, Loeb, & Wyckoff, 2013).

Both ASD and iZone schools hired new teachers and principals at very high rates in the first year of the reform, which was a required component of the federal turnaround models. In fact, both ASD and the iZones hired more effective teachers as measured by their value-added scores, 24 and 19 percentage points more than the comparison schools, respectively. In the

iZone schools, hiring effective teachers appears to have overcome the disruption of hiring new teachers, but even higher rates of hiring effective teachers did not offset the initial disruption in the ASD. Moreover, neither the iZone nor ASD hired more effective teachers than comparison schools after the first year of turnaround. The difference appears to be that iZone schools retained their newly hired teachers but the ASD hired teachers with less than three years of experience at higher rates than the comparison schools in all three years of the reform. It may be that the pool of effective teachers willing to move into low-performing schools after the initial year of turnaround was very limited. In years two and three, we find that the ASD no longer recruited effective teachers. At the principal level, we find a parallel situation in years two and three where ASD schools were led by mostly inexperienced principals while iZones were led by principals who had higher observation scores. This result suggests that turnaround administrators should be attentive to planning, implementing, and monitoring reforms that are clearly focused on recruiting and retaining effective teachers and principals and replacing lower performing teachers and principals with individuals that are more effective and experienced.

An additional takeaway from these findings is that the positive effects in the iZone schools was only partially explained by the measures of instability of teachers, principals, and students. This leads us to conclude that either the iZones intervention directly influenced the test score gains or the effects unexplained in this analysis were mediated by other variables. In other words, the efforts to establish an educational infrastructure in the iZone schools, including increasing the capacity for effective leadership and instruction likely play important roles in improving the performance of the lowest performing schools. Measures of these factors should be developed and data on those measures collected in research to formally test additional mechanisms that explain effective school reform. In addition, states and districts engaged in the

comprehensive support and improvement should carefully measure, monitor and report labor force variables and malleable student factors that interfere with instructional time for their lowest performing schools undergoing reform. In addition, these reform leaders will need to act to stabilize these schools, hire effective personnel, and work with students and their families to reduce absenteeism and school transfers.

Finally, we do find evidence that disrupting the status quo by setting a quota, even though the quota appeared to arbitrarily set, for the replacement of the staff of the lowest performing schools brings about performance improvement. This strategy appears to be a rather blunt policy instrument that may have negative side effects, including lowering initial school performance and as other research suggests, alienating the communities the schools serve (Glazer & Egan, 2018). The sheer volume of turnover from the quotas on staff replacement may divert attention from hiring effective staff and retaining them. Also, to the extent that the demand for effective teachers and principals to fill the positions vacated to meet the quota exceeds the supply of such personnel, the quotas can be counterproductive.

Endnotes

[1] Results for individual districts are available upon request.
[2] We also tested the sensitivity of our results to different definitions for a novice teacher ranging from 2 to 5 years of experience and found similar results, available upon request.
[3] We use two different subsamples instead of including indicators for both ASD and iZone schools in one model for clarity of presentation and to ease the interpretation of coefficients, because including both turnaround interventions requires three-way interactions with mediators that obscure the mediational pathways of interest.
[4] Imai and colleagues (2010) refer to this effect as the average causal mediation effect (ACME) but we prefer to use the term average mediation effect to avoid implying that our estimates are causal.
[5] We use our entire sample to estimate the average effect as proposed by Imai and colleagues (2010). However, our sensitivity analysis could not be conducted on the full sample because the large number of observations made the simulations intractable. Therefore, we conduct the sensitivity analysis using a random subsample of 10 percent of our data, holding the relative school sizes constant. Our sensitivity results are robust to using a random subsample of 15 and 20 percent of our full sample.

References

Aladjem, D. K., Birman, B. F., Orland, M., Harr-Robins, J., Heredia, A., Parrish, T. B., & Ruffini, S. J. (2010). Achieving Dramatic School Improvement: An Exploratory Study. A Cross-Site Analysis from the Evaluation of Comprehensive School Reform Program Implementation and Outcomes Study. *Office of Planning, Evaluation and Policy Development, US Department of Education*. Retrieved from http://eric.ed.gov/?id=ED526783

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

Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a decade*. Retrieved from Rand Corporation website: https://books.google.com/books?hl=en&lr=&id=IAFT_Ln9JwcC&oi=fnd&pg=PR3&dq=berends+kirby+facing+the+challenges&ots=cy3UELrPXl&sig=-coXvQr4H_wXxWGbvKVG92sCxfU

Bonilla, S., & Dee, T. (2017). *The Effects of School Reform Under NCLB Waivers: Evidence from Focus Schools in Kentucky*. National Bureau of Economic Research.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *The American Economic Review*, *95*(2), 166–171.

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

Chiang, H. (2009). How accountability pressure on failing schools affects student achievement. *Journal of Public Economics*, *93*(9), 1045–1057. https://doi.org/10.1016/j.jpubeco.2009.06.002

Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus*. National Bureau of Economic Research.

Delaney, M. (2017). Charter operators OK'd for Achievement School District program. Retrieved November 3, 2018, from Las Vegas Review-Journal website: https://www.reviewjournal.com/news/education/charter-operators-okd-for-achievement-school-district-program/

Desmond, M., & Travis, A. (2018). Political Consequences of Survival Strategies among the Urban Poor. *American Sociological Review*, *83*(5), 869–896.

Dickey-Griffith, D. (2013). *A Preliminary Analysis of the Effects of the School Improvement Grant Program on Student Achievement in Texas* (PhD Thesis). Georgetown University.

Dougherty, S. M., & Weiner, J. M. (2017). The Rhode to Turnaround: The Impact of Waivers to No Child Left Behind on School Performance. *Educational Policy*, 0895904817719520.

Downey, M. (2016). Will state takeover improve Georgia schools? Retrieved November 3, 2018, from The Atlanta Journal-Constitution website: https://www.ajc.com/news/local-education/will-state-takeover-improve-georgia-schools/AV9JC3qxyymcNNH0OGe3WI/

Egbert, R. L. (1981). *Some Thoughts about Follow Through Thirteen Years Later.* (No. (ERIC Document Reproduction Service No. ED244733)). Lincoln: Nebraska University.

ESEA Flexibility Request. (2012). Submitted to the U.S. Department of Education.

Gill, B., Zimmer, R., Christman, J., & Blanc, S. (2007). State Takeover, School Restructuring, Private Management, and Student Achievement in Philadelphia. *RAND Corporation*.

Glazer, J. L., & Egan, C. (2018). The Ties That Bind: Building Civic Capacity for the Tennessee Achievement School District. *American Educational Research Journal*, *55*(5), 928–964. https://doi.org/10.3102/0002831218763088

Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools: An instrumental variables approach. *American Educational Research Journal*, *47*(2), 434–465.

Granados, A. (2017). Superintendent shares plans for NC's new Achievement School District :: WRAL.com. Retrieved November 3, 2018, from WRAL-TV website: https://www.wral.com/superintendent-shares-plans-for-nc-s-new-achievement-school-district/16807037/

Gross, B., Booker, T. K., & Goldhaber, D. (2009). Boosting student achievement: The effect of comprehensive school reform on student achievement. *Educational Evaluation and Policy Analysis*, *31*(2), 111–126.

Guarino, C. M., Santibanez, L., & Daley, G. A. (2006). Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research*, *76*(2), 173–208.

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

Harris, D. N., & Larsen, M. F. (2016). *The Effects of the New Orleans Post-Katrina School Reforms on Student Academic Outcomes*. Technical Report, Education Research Alliance for New Orleans.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Publications.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: does method really matter? *Psychological Science*, *24*(10), 1918–1927.

Heissel, J. A., & Ladd, H. F. (2017). School turnaround in North Carolina: A regression discontinuity analysis. *Economics of Education Review*, *62*, 302–320.

Henry, G. T., Bastian, K. C., & Fortner, C. K. (2011). Stayers and leavers early-career teacher effectiveness and attrition. *Educational Researcher*, *40*(6), 271–280.

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. *The Friday Institute for Educational Innovation, North Carolina State University. Google Scholar*.

Henry, G. T., & Harbatkin, E. (2018). The Next Generation of School Reform: Improving the Lowest Performing Schools without Disrupting the Status Quo. *Paper Presented at the Annual Conference of the Association for Public Policy Analysis and Management*.

Henry, G. T., Zimmer, R., Attridge, J., Kho, A., & Viano, S. (2014). *Teacher and student migration in and out of Tennessee's Achievement School District*. Retrieved from Tennessee Consortium on Research, Evaluation & Development website: https://peabody.vanderbilt.edu/TERA/files/ASD_Teacher_Student_Migration.pdf

Henry, G. T., Zimmer, R., Kho, A., & Pham, L. (2017). *Recruitment and Retention of Teachers in Tennessee's Achievement School District and iZone Schools*. Retrieved from Tennessee Education Research Alliance website: https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/Teacher_Retention_in_ASD_and_iZone_vF.pdf

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning Around Chronically Low-Performing Schools. IES Practice Guide*. Retrieved from National Center for Education Evaluation and Regional Assistance website: http://eric.ed.gov/?id=ED501241

Hess, G. A. (2003). Reconstitution—Three Years Later: Monitoring the Effect of Sanctions on Chicago High Schools. *Education and Urban Society*, *35*(3), 300–327. https://doi.org/10.1177/0013124503035003004

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309.

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, *105*(4), 765–789.

Le Floch, K., O'Day, J., Birman, B., Hurlburt, S., Nayfack, M., Halloran, C., … Wei, T. (2016). *Case Studies of Schools Receiving School Improvement Grants: Final Report* [Indexes; Offices]. Retrieved from NCEE website: http://ies.ed.gov/ncee/pubs/20164002/index.asp

Leithwood, K., & Steinbach, R. (2003). Successful leadership for especially challenging schools. *B. Davies & J. West-Burnham (Eds.), Handbook of Educational Leadership and Management*, 25–43.

Leithwood, K., & Strauss, T. (2008). *Turnaround Schools and the Leadership They Require.* Canadian Education Association.

MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Routledge.

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.

Papay, J., & Hannon, M. (2018). *The Effects of School Turnaround Strategies in Massachusetts*. Working Paper, Brown University.

Parke, C., & Kanyongo, G. (2012). Student Attendance, Mobility, and Mathematics Achievement in an Urban School District. *The Journal of Educational Research*, *105*(3), 161–175. https://doi.org/10.1080/00220671.2010.547231

Pedulla, D. S. (2016). Penalized or protected? Gender and the consequences of nonstandard and mismatched employment histories. *American Sociological Review*, *81*(2), 262–289.

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731.

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*(2), 93.

Public Chapter No. 2. (2010). Retrieved from http://www.tennessee.gov/sos/acts/106/pub/pc0002EOS.pdf

Redding, C., & Henry, G. T. (2018). New evidence on the frequency of teacher turnover: Accounting for within-year turnover. *Educational Researcher*, *47*(9), 577–593.

Rice, J. K., & Malen, B. (2010). *School Reconstitution as an Education Reform Strategy*. Retrieved from National Education Association website: http://199.223.128.58/assets/docs/School_Reconstruction_and_an_Education_Reform_Strategy.pdf

Riley, B. (2015). Upgrade S.C. education with Achievement School District. Retrieved November 3, 2018, from The Post and Courier website: https://www.postandcourier.com/opinion/upgrade-s-c-education-with-achievement-school-districts/article_b4f563b6-f258-55fa-be72-4f8a3b9cf155.html

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, *50*(1), 4–36.

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.

Scott, C. (2009). Improving Low-Performing Schools: Lessons from Five Years of Studying School Restructuring Under No Child Left Behind. *Center on Education Policy*.

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.

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

Sun, M., 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

VanderWeele, T. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.

Zimmer, R., Henry, G. T., & Kho, A. (2016). The Effects of School Turnaround in Tennessee's Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis*, 0162373717705729.

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.

Table 1. Number of Schools by Reform Approach

Year	Total Priority	Non-iZone, Non-ASD Priority	iZone Schools	ASD Schools in Operation	ASD Schools Included in Analysis [4]
2012-13	82[1]	65	11	6	6
2013-14	84[2]	45	22	17	11
2014-15	77[3]	28	26	23	16

Note: [1] The original 2012-13 list of priority schools in Tennessee included 83 schools, but one was closed in that year. [2] The increase in the total number of Priority schools from 2012-13 to 2013-14 comes from the addition of four new ASD schools, the splitting of one school into two separate schools by the ASD, and the closure of three Priority schools. [3] The decrease in the total number of Priority schools from 2013-14 to 2014-15 comes from the addition of two new ASD schools, the creation of a second school at a former school the ASD took over in 2012-13, the merging of two ASD schools into other ASD schools, and the closure of eight other Priority schools. [4] During this time period, the ASD opened new-start schools that did not exist previously and were not named a priority school by Tennessee on its 2012 Priority list. Also, in these years, the ASD began operating in untested grades in some priority schools. New start schools and schools where the ASD had not yet began operating tested grades were not included in this analysis, and one priority school was not included because the ASD had only begun operating in untested grades in this school during the study period.

Table 2. Descriptive Statistics for ASD Schools, iZone Schools, and Non-ASD, Non-iZone Priority Schools

	Non-ASD, Non-iZone Priority				ASD				iZone			
	Year 0	Year 1	Year 2	Year 3	Year 0	Year 1	Year 2	Year 3	Year 0	Year 1	Year 2	Year 3
Test Scores												
Reading Test Scores	-0.93	-0.90	-0.83	-0.74	-1.00	-0.99	-1.01	-0.93	-1.04	-0.91	-0.77	-0.79
Math Test Scores	-0.91	-0.79	-0.79	-0.68	-0.84	-0.89	-0.68	-0.54	-0.88	-0.63	-0.54	-0.63
Science Test Scores	-1.04	-0.92	-0.80	-0.81	-1.10	-1.14	-1.04	-0.67	-1.13	-0.86	-0.67	-0.74
Teacher Characteristics												
Female	0.73	0.73	0.74	0.75	0.76	0.77	0.82	0.79	0.72	0.73	0.74	0.77
Minority Race	0.71	0.69	0.69	0.67	0.75	0.57	0.49	0.57	0.65	0.68	0.63	0.72
Average Years of Experience	10.95	10.36	10.16	9.82	11.66	6.27	3.76	3.85	11.67	9.56	8.75	9.66
New to School	0.22	0.24	0.31	0.26	0.26	0.64	0.55	0.57	0.26	0.54	0.35	0.33
New Teacher with TVAAS >= 4	0.01	0.03	0.06	0.04	0.03	0.28	0.11	0.14	0.05	0.23	0.07	0.09
New Teacher with < 3 Years of Experience	0.11	0.15	0.18	0.14	0.14	0.45	0.39	0.39	0.12	0.27	0.19	0.15
Principal Characteristics												
Female	0.51	0.47	0.43	0.56	0.71	0.52	0.53	0.51	0.59	0.43	0.51	0.61
Minority Race	0.91	0.88	0.83	0.84	0.97	0.56	0.77	0.79	0.74	0.88	0.84	0.84
Average Years of Experience as Educator	15.12	15.47	15.93	16.53	17.35	13.15	9.37	11.23	19.03	12.78	15.99	15.09
New to School	0.17	0.25	0.17	0.37	0.20	0.73	0.77	0.70	0.48	0.63	0.22	0.26
Fewer than 3 Years of Principal Experience	0.57	0.39	0.27	0.36	0.25	0.74	0.96	0.91	0.66	0.71	0.39	0.44
Observation Score (1-5)	3.61	3.88	3.45	3.55	3.39	3.50	3.55	3.53	3.51	3.83	4.03	4.20
Student Characteristics												
Female	0.50	0.50	0.50	0.50	0.51	0.50	0.48	0.48	0.50	0.49	0.49	0.47
Minority Race	0.99	0.99	0.99	0.99	0.98	0.99	0.98	0.98	0.97	0.97	0.96	0.97
FRPL Eligible	0.86	0.87	0.90	0.88	0.92	0.88	0.96	0.92	0.90	0.94	0.92	0.98
ELL Status	0.03	0.03	0.03	0.05	0.01	0.01	0.01	0.01	0.02	0.03	0.04	0.01
SpED Status	0.11	0.11	0.12	0.16	0.11	0.14	0.11	0.17	0.10	0.11	0.13	0.19
Chronic Absent Indicator	0.19	0.19	0.21	0.19	0.18	0.27	0.21	0.34	0.22	0.26	0.23	0.20
New to School (Nonstructural Move)	0.31	0.33	0.33	0.29	0.35	0.35	0.34	0.27	0.29	0.30	0.34	0.37
Average Student Enrollment	527.0	539.8	507.9	556.5	497.1	513.2	479.2	395.8	519.5	524.4	487.5	434.6

Note. FRPL is free or reduced-price lunch. ELL is English Language Learner. SpED is special education. TVAAS and observation scores range from 1-5. Chronically absent students miss more that 10 percent of the instructional days they are enrolled in the school.

Table 3. Difference-in-Differences Mediation Models: Teacher Characteristics

	Mediator:	Proportion of Teachers who are New to the School			Proportion of Teachers who are New to the School and have a TVAAS Score >= 4 (out of 5)			Proportion of Teachers who are New to the School and have Fewer than Three Years of Experience		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Outcome:	Test Scores	Mediator	Test Scores	Test Scores	Mediator	Test Scores	Test Scores	Mediator	Test Scores
ASD										
ASD*Year 1		-0.063	0.417***	0.044	-0.055	0.243***	-0.068	-0.063	0.322***	0.024
		(0.074)	(0.050)	(0.076)	(0.074)	(0.059)	(0.087)	(0.074)	(0.039)	(0.077)
ASD*Year 2		0.066	0.299***	0.142*	0.073	0.051	0.071	0.066	0.226***	0.127+
		(0.069)	(0.081)	(0.069)	(0.069)	(0.049)	(0.068)	(0.069)	(0.045)	(0.072)
ASD*Year 3		0.103	0.286***	0.176*	0.115	0.046	0.112	0.103	0.213***	0.160+
		(0.088)	(0.067)	(0.086)	(0.088)	(0.082)	(0.091)	(0.088)	(0.041)	(0.090)
Mediator				-0.256**			0.051			-0.270**
				(0.079)			(0.132)			(0.101)
Adjusted R Squared		0.419	0.336	0.420	0.421	0.304	0.421	0.419	0.420	0.419
Observations		161642	161642	161642	159455	159455	159455	161642	161642	161642
iZone										
iZone*Year 1		0.146***	0.313***	0.169***	0.150***	0.186***	0.093**	0.146***	0.132***	0.168***
		(0.041)	(0.044)	(0.048)	(0.041)	(0.052)	(0.034)	(0.041)	(0.025)	(0.042)
iZone*Year 2		0.179***	0.033	0.181***	0.188***	-0.024	0.195***	0.179***	0.014	0.181***
		(0.039)	(0.041)	(0.039)	(0.038)	(0.023)	(0.037)	(0.039)	(0.024)	(0.039)
iZone*Year 3		0.165***	0.019	0.167***	0.175***	0.013	0.171***	0.165***	-0.001	0.165**
		(0.048)	(0.057)	(0.049)	(0.048)	(0.028)	(0.047)	(0.048)	(0.032)	(0.050)
Mediator				-0.072			0.303***			-0.161
				(0.088)			(0.077)			(0.099)
Adjusted R Squared		0.421	0.299	0.421	0.423	0.323	0.424	0.421	0.194	0.421
Observations		213683	213683	213683	209776	209776	209776	213683	213683	213683
Student and School Level Covars.		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School and Grade Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note. Standard errors in parentheses clustered at the school level. + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 4. Difference-in-Differences Mediation Models: Principal Characteristics

	Mediator:	Principal is New to the School			Principal Has Fewer than 3 Years of Experience as a Head Principal			Principal's Composite Observation Score		
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Outcome:	Test Scores	Mediator	Test Scores	Test Scores	Mediator	Test Scores	Test Scores	Mediator	Test Scores
ASD										
ASD*Year 1		-0.063	0.501**	-0.032	-0.063	0.410	-0.048	-0.046	-0.272+	-0.033
		(0.074)	(0.177)	(0.067)	(0.074)	(0.256)	(0.071)	(0.071)	(0.165)	(0.068)
ASD*Year 2		0.066	0.843***	0.119+	0.066	0.859***	0.097	0.075	-0.140	0.082
		(0.069)	(0.165)	(0.062)	(0.069)	(0.165)	(0.067)	(0.083)	(0.197)	(0.075)
ASD*Year 3		0.103	0.564*	0.138+	0.103	0.988***	0.139	0.121	-0.026	0.122
		(0.088)	(0.268)	(0.084)	(0.088)	(0.213)	(0.085)	(0.090)	(0.191)	(0.091)
Mediator				-0.063**			-0.037+			0.048**
				(0.021)			(0.019)			(0.018)
Adjusted R Squared		0.419	0.137	0.419	0.419	0.213	0.419	0.425	0.130	0.426
Observations		161642	161642	161642	161642	161642	161642	127836	127836	127836
iZone										
iZone*Year 1		0.146***	0.300+	0.155***	0.146***	0.436*	0.159***	0.154***	-0.037	0.156***
		(0.041)	(0.176)	(0.041)	(0.041)	(0.173)	(0.042)	(0.040)	(0.211)	(0.039)
iZone*Year 2		0.179***	0.028	0.180***	0.179***	0.156	0.183***	0.173***	0.496*	0.150***
		(0.039)	(0.147)	(0.039)	(0.039)	(0.191)	(0.039)	(0.041)	(0.214)	(0.040)
iZone*Year 3		0.165***	-0.160	0.161***	0.165***	0.147	0.170***	0.157**	0.712**	0.124*
		(0.048)	(0.183)	(0.048)	(0.048)	(0.286)	(0.050)	(0.053)	(0.236)	(0.049)
Mediator				-0.028			-0.029			0.046*
				(0.018)			(0.018)			(0.018)
Adjusted R Squared		0.421	0.065	0.421	0.421	0.150	0.421	0.427	0.143	0.428
Observations		213683	213683	213683	213683	213683	213683	169495	169495	169495
Student and School Level Covariates		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School and Grade Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note. Standard errors in parentheses clustered at the school level. + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 5. Difference-in-Differences Mediation Models: Student Characteristics

	Mediator:	Student Missed More than 10 Percent of Instructional Days (Chronically Absent)			Student is New to the School at the Beginning of the Year (Nonstructural Move)		
		(1)	(2)	(3)	(4)	(5)	(6)
	Outcome:	Test Scores	Mediator	Test Scores	Test Scores	Mediator	Test Scores
ASD							
ASD*Year 1		-0.071	0.077+	-0.059	-0.071	0.026	-0.069
		(0.075)	(0.040)	(0.075)	(0.075)	(0.037)	(0.074)
ASD*Year 2		0.064	0.036	0.070	0.063	-0.000	0.063
		(0.068)	(0.043)	(0.066)	(0.069)	(0.069)	(0.073)
ASD*Year 3		0.100	0.113+	0.118	0.100	-0.031	0.098
		(0.090)	(0.067)	(0.099)	(0.090)	(0.051)	(0.089)
Mediator				-0.156***			-0.070***
				(0.009)			(0.006)
Adjusted R Squared		0.421	0.020	0.425	0.421	0.026	0.422
Observations		159171	159171	159171	159190	159190	159190
iZone							
iZone*Year 1		0.145***	0.060*	0.155***	0.145***	0.028	0.146***
		(0.041)	(0.026)	(0.041)	(0.041)	(0.023)	(0.041)
iZone*Year 2		0.181***	0.051+	0.189***	0.181***	0.090*	0.186***
		(0.040)	(0.026)	(0.041)	(0.040)	(0.039)	(0.040)
iZone*Year 3		0.168***	0.055*	0.177***	0.168***	0.114**	0.175***
		(0.048)	(0.024)	(0.048)	(0.048)	(0.041)	(0.048)
Mediator				-0.169***			-0.060***
				(0.008)			(0.007)
Adjusted R Squared		0.424	0.020	0.428	0.424	0.026	0.425
Observations		210691	210691	210691	210706	210706	210706
Student and School Level Covariates		Yes	Yes	Yes	Yes	Yes	Yes
School and Grade Fixed Effects		Yes	Yes	Yes	Yes	Yes	Yes

Note. Standard errors in parentheses clustered at the school level. + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 6. Partially Standardized Indirect Effect Estimates and Bias-Corrected Bootstrap 95% Confidence Intervals

	Proportion of Teachers who are New to the School	Proportion of Teachers who are New to the School and have a TVAAS Score >= 4 (out of 5)	Proportion of Teachers who are New to the School and have Fewer than Three Years of Experience	Principal is New to the School	Principal Has Fewer than 3 Years of Experience as a Head Principal	Principal's Composite Observation Score	Student Missed More than 10 Percent of Instructional Days (Chronically Absent)	Student is New to the School at the Beginning of the Year (Nonstructural Move)
ASD Year 1	-0.11**	0.012	-0.087*	-0.032*	-0.015	-0.013	-0.012	-0.0019
	[-0.20, -0.040]	[-0.075, 0.068]	[-0.18, -0.020]	[-0.074, -0.0085]	[-0.051, 0.0017]	[-0.043, 0.00014]	[-0.025, 0.0014]	[-0.0071, 0.0041]
ASD Year 2	-0.076*	0.0026	-0.061*	-0.053*	-0.032	-0.0067	-0.0056	0.000015
	[-0.16, -0.032]	[-0.012, 0.036]	[-0.13, -0.018]	[-0.11, -0.016]	[-0.073, -0.0005]	[-0.040, 0.0079]	[-0.020, 0.0075]	[-0.010, 0.010]
ASD Year 3	-0.073*	0.0023	-0.057*	-0.036	-0.037	-0.0012	-0.018	0.0022
	[-0.16, -0.027]	[-0.014, 0.075]	[-0.14, -0.017]	[-0.10, 0.0038]	[-0.090, -0.0013]	[-0.027, 0.018]	[-0.039, 0.0027]	[-0.0044, 0.011]
iZone Year 1	-0.023	0.056*	-0.021	-0.0085	-0.013	-0.0017	-0.010*	-0.0016
	[-0.089, 0.026]	[0.016, 0.11]	[-0.054, 0.0034]	[-0.036, 0.00097]	[-0.037, -0.0001]	[-0.028, 0.016]	[-0.020, -0.0015]	[-0.0051, 0.0007]
iZone Year 2	-0.0024	-0.0073	-0.0023	-0.00081	-0.0046	0.023*	-0.0086	-0.0054*
	[-0.025, 0.0029]	[-0.024, 0.0049]	[-0.019, 0.0039]	[-0.016, 0.0064]	[-0.026, 0.0043]	[0.0026, 0.060]	[-0.017, 0.00029]	[-0.012, -0.0014]
iZone Year 3	-0.0014	0.004	0.00009	0.0045	-0.0043	0.033*	-0.0092*	-0.0068*
	[-0.029, 0.0062]	[-0.0097, 0.029]	[-0.014, 0.013]	[-0.0029, 0.029]	[-0.043, 0.010]	[0.0074, 0.080]	[-0.018, -0.0009]	[-0.014, -0.0025]

Note. Bias-correct Bootstrap 95% CI shown in brackets. + p<0.10 * p<0.05 ** p<0.01 *** p<0.001

Table 7. Average Mediation Effects Using a Potential Outcomes Mediational Framework Proposed by Imai, Keele, and Tingley (2010)

	Proportion of Teachers who are New to the School	Proportion of Teachers who are New to the School and have a TVAAS Score >= 4	Proportion of Teachers who are New to the School and have Fewer than Three Years of Experience	Principal is New to the School	Principal Has Fewer than 3 Years of Experience as a Head Principal	Principal's Composite Observation Score	Student Missed More than 10 Percent of Instructional Days (Chronically Absent)	Student is New to the School at the Beginning of the Year (Nonstructural Move)
ASD								
Year 1	-0.09**	0.01	-0.05	-0.04*	-0.01	-0.01	-0.01	0.0006
	[-0.15,-0.02]	[-0.05,0.08]	[-0.11,0.02]	[-0.08,-0.01]	[-0.05,0.01]	[-0.03,0.003]	[-0.03,0.004]	[-0.01,0.01]
Year 2	-0.06*	0.003	-0.04	-0.07**	-0.03	-0.005	-0.004	0.003
	[-0.12,-0.01]	[-0.01,0.02]	[-0.08,0.01]	[-0.12,-0.02]	[-0.07,0.01]	[-0.03,0.01]	[-0.02,0.01]	[-0.02,0.02]
Year 3	-0.06*	0.003	-0.04	-0.05+	-0.03	-0.0009	-0.02	0.01
	[-0.11,-0.01]	[-0.02,0.03]	[-0.08,0.01]	[-0.10,-0.01]	[-0.08,0.01]	[-0.02,0.01]	[-0.04,0.004]	[-0.005,0.02]
iZone								
Year 1	-0.04+	0.06*	-0.02+	-0.01	-0.02+	-0.005	-0.01*	-0.002
	[-0.08,0.01]	[0.02,0.12]	[-0.05,0.002]	[-0.04,0.003]	[-0.05,-0.001]	[-0.04,0.02]	[-0.02,-0.001]	[-0.007,0.002]
Year 2	-0.003	-0.005	-0.001	0.0003	-0.01	0.03+	-0.01+	-0.005+
	[-0.02,0.01]	[-0.02,0.01]	[-0.01,0.01]	[-0.01,0.01]	[-0.03,0.02]	[0.003,0.06]	[-0.02,0.001]	[-0.01,0.0004]
Year 3	-0.006	0.008	-0.004	0.01	-0.01	0.04*	-0.01*	-0.01*
	[-0.03,0.01]	[-0.01,0.03]	[-0.02,0.01]	[-0.01,0.03]	[-0.05,0.02]	[0.01,0.09]	[-0.02,-0.001]	[-0.01,-0.002]

Note. 95% confidence intervals in parentheses from 1000 quasi-Bayesian simulations. Average mediation effects estimated from simulation methods proposed by Imai et al. (2010).

Table 8. Sensitivity Parameters

	Proportion of Teachers who are New to the School	Proportion of Teachers who are New to the School and have a TVAAS Score >= 4	Proportion of Teachers who are New to the School and have Fewer than Three Years of Experience	Principal is New to the School	Principal Has Fewer than 3 Years of Experience as a Head Principal	Principal's Composite Observation Score	Student Missed More than 10 Percent of Instructional Days (Chronically Absent)	Student is New to the School at the Beginning of the Year (Nonstructural Move)
ASD								
ρ	-0.03540	0.00520	-0.01650	-0.03930	-0.00790	0.02780	-0.08380	-0.05410
$R_m^2 * R_Y^2$	0.00050	0.00001	0.00010	0.00070	0.00002	0.00040	0.00380	0.00160
iZone								
ρ	-0.01720	0.03200	-0.02280	-0.01670	-0.02710	0.04550	-0.08400	-0.03770
$R_m^2 * R_Y^2$	0.00010	0.00040	0.00030	0.00010	0.00040	0.00100	0.00380	0.00080

Note. Rho is the correlation between the error term in the mediator model with the error term in the outcome model where the average mediation effect is expected to equal 0. $R_m^2 * R_Y^2$ represents the product of the total variation explained for the mediator times the total variation explained for the outcome from a hypothetical unobserved confounder that would render the average mediation effect statistically insignificant. The outcome is always standardized student test scores and each mediator is tested separately.

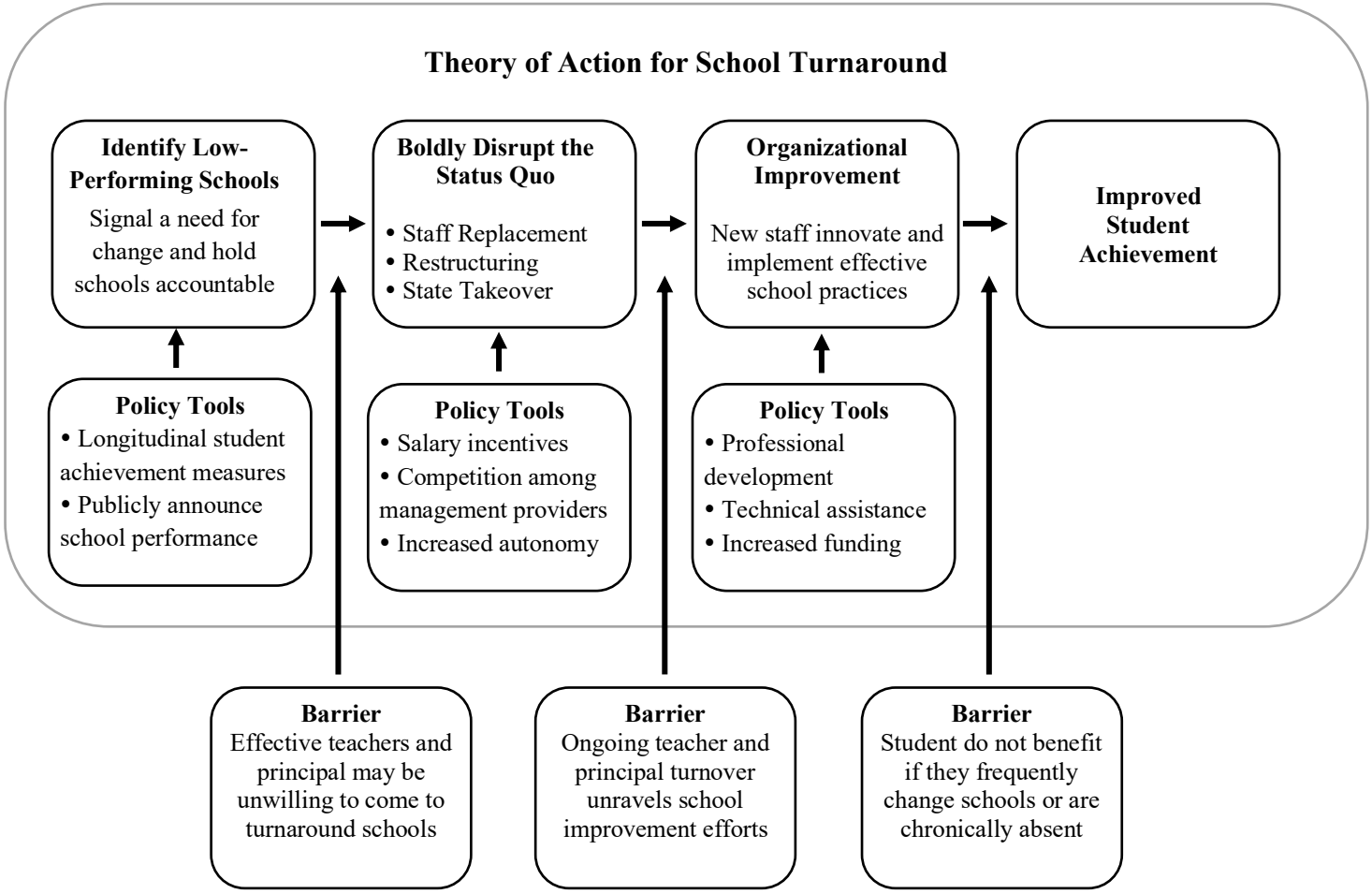

Figure 1. Theory of Action for School Turnaround and Unanticipated Obstacles that May Suppress Positive Effects of Reform Efforts

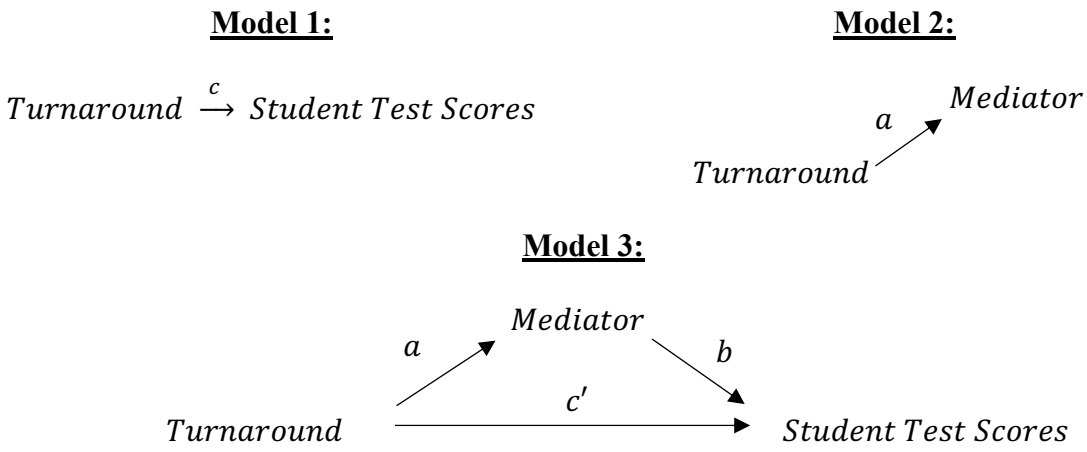

Figure 2. Mediational Framework

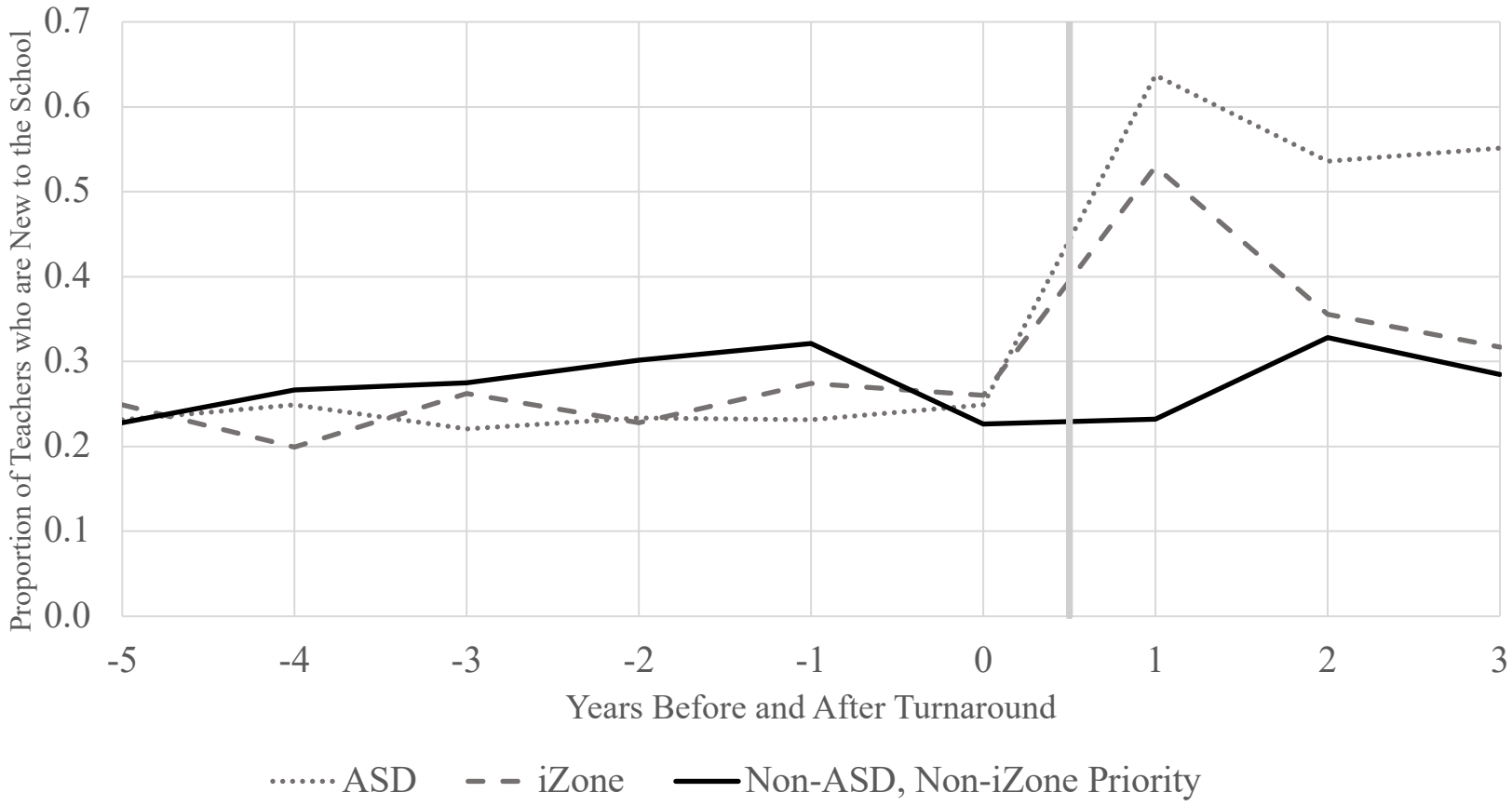

Figure 3. Trends in Teacher Turnover Rates for ASD, iZone, and Non-ASD, Non-iZone Priority Schools for Five Years Before and Three Years After Turnaround Reforms Began.

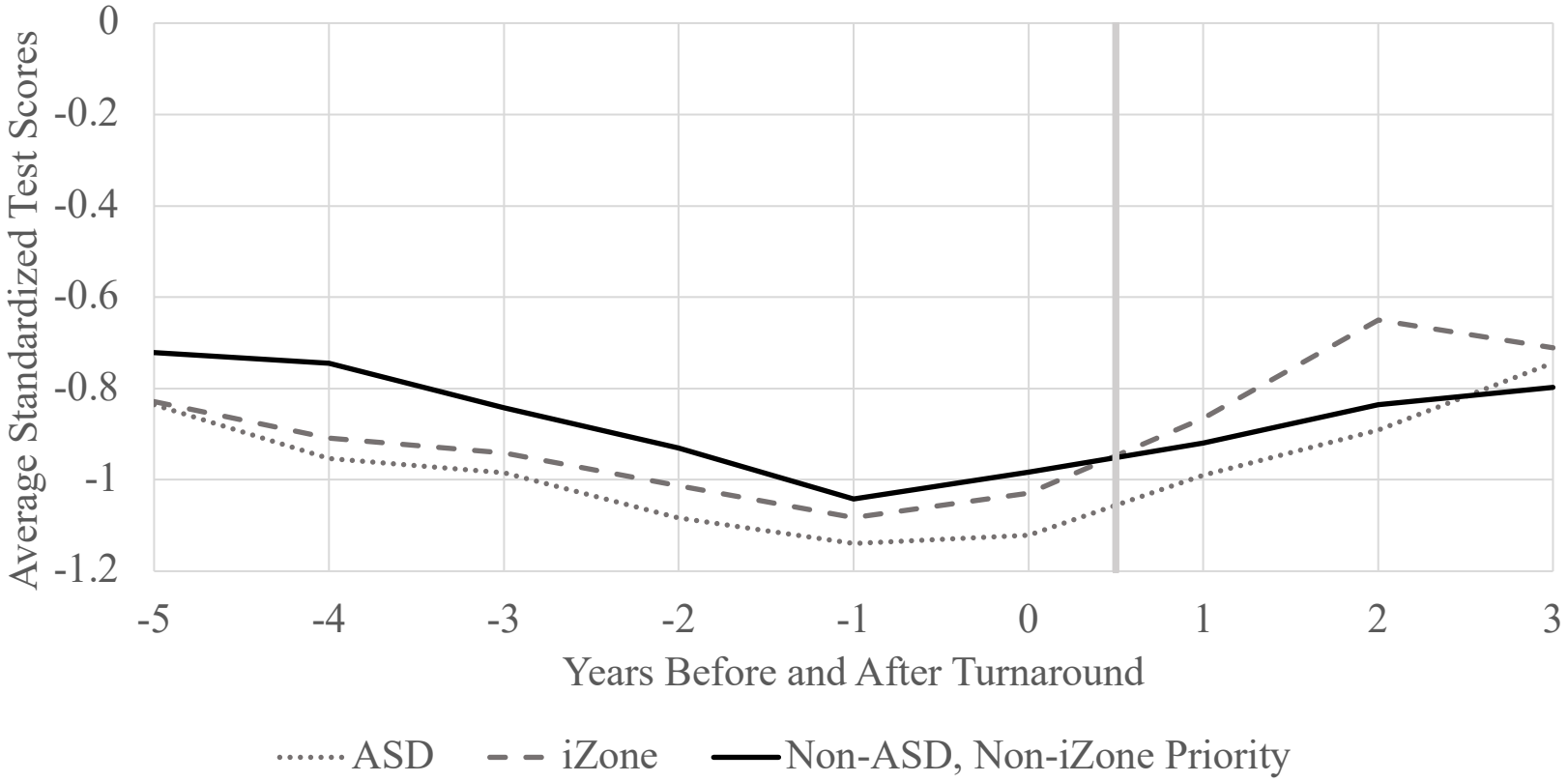

Figure 4. Trends in Average Standardized Test Scores at the School Level

Figure Captions

Figure 1. Theory of Action for School Turnaround and Unanticipated Obstacles that May Suppress Positive Effects of Reform Efforts

Figure 2. Mediational Framework

Figure 3. Trends in Teacher Turnover Rates for ASD, iZone, and Non-ASD, Non-iZone Priority Schools for Five Years Before and Three Years After Turnaround Reforms Began.

Figure 4. Trends in Average Standardized Test Scores at the School Level